Culture, Computer-Mediated Communication, and Survey Interviewing

Susan R. Fassell
Carnegie Mellon University, Pittsburgh, Pennsylvania

Qiping Zhang
Long Island University, Brooklyn, New York

Frederick G. Conrad
University of Michigan, Ann Arbor, Michigan

Michael F. Schober
New School for Social Research, New York, New York

Leslie D. Selick
Carnegie Mellon University, Pittsburgh, Pennsylvania

11.1 INTRODUCTION

As survey designers test and implement new interviewing technologies, the growing body of evidence on cultural differences in computer-mediated communication (CMC) is becoming increasingly relevant. People from different cultures can differ in patterns and styles of communication and interpretation, and this can affect how they interact with new technologies for communicating with human partners and with computer systems. This is likely to be the case for current and future interviewing systems. Although there is as yet little direct evidence from studies of survey interviews, studies of CMC help lay the groundwork for understanding and predicting effects of culture in technologically mediated survey interviews.

Envisioning the Survey Interview of the Future, Edited by Frederick G. Conrad and Michael F. Schober
Copyright © 2006 John Wiley & Sons, Inc.
Consider the sample dialogue in Table 11.1. These conversations come from pairs of American and Chinese students negotiating a jointly agreed upon order of priority for items in the Arctic Survival Task (Setlock et al., 2004; Stewart et al., in press) either face-to-face or over instant messaging (IM). In the selected excerpts, the pairs are trying to agree on the most important items in the set. When American pairs do this task, they discuss each item in a cursory manner (6-7 speaking turns) regardless of communication medium, and they are quick to acquiesce to their partners’ suggested rankings (see the last turn of each utterance). When Chinese pairs do this task face-to-face, they discuss each item in depth, asking each other questions and working through the survival scenario (e.g., “...the most important thing we need to fight in the coldness? Right?”). Discussion of a single item can take many speaking turns, 42, in this excerpt. The most striking aspect of these dialogues is the way in which the Chinese pairs’ conversation shifts when they talk over IM. These conversations are similar to those of the American pairs in terms of brevity and acquiescence, and quite unlike the lengthy discussions of the Chinese pairs in a face-to-face setting.

CMC studies such as the one that produced these dialogues suggest design considerations—though not yet prescriptions—for designers of new interviewing systems. Designers of future interviewing systems will be able to choose system features that are not available to designers of current surveys and these choices may differently affect the behavior of respondents from different cultures. For example, a designer might choose to display an interviewer’s or interviewing agent’s facial cues in the user interface and these cues may affect respondents from context-dependent cultures differently than respondents from context-independent cultures (see later discussion). Similarly, the dialect of the interviewer’s or interviewing agent’s voice seems likely to affect respondents who are speakers of that dialect differently than respondents from other linguistic communities. Such differences can potentially affect people’s willingness to participate in the interview, to provide thoughtful answers, to provide honest answers to sensitive questions, and the likelihood that they will complete the interview. But these effects will only be evident if the medium communicates an interviewer dialect. Such differences would not be evident in an IM interview, for example. In a globalizing world with increasing migration, survey interviews are increasingly intercultural (with interviewers and respondents coming from different cultural backgrounds), which heightens the need for understanding these issues.

A growing body of literature about the impact of culture on survey data now exists. A substantial component of this work addresses the practical issues of conducting cross-cultural survey research such as translating questionnaires (e.g., Hardaker, 2003). Closer to the concern of the current chapter, other work addresses cultural sources of measurement error, that is, the discrepancy between what a respondent reports and the true value of the answer. For example, Johnson and van de Vijver (2003) report that survey respondents from collectivist societies, that is, societies in which people prioritize the benefits of the larger group over their own benefit, are more likely to give socially desirable answers than their counterparts from individualistic societies, that is, societies in which these priorities are reversed, presumably because there is greater pressure to conform to social norms in the former than the latter type of society. Even closer to the topic of the current chapter, there is some evidence that...

<table>
<thead>
<tr>
<th>TABLE 11.1 Speech Between Americans and Chinese Pairs in Arctic Survival Task Carried Out Face-to-Face (Row 1) and Through Instant Messaging (Row 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Medium</strong></td>
</tr>
<tr>
<td>Face-to-Face</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>IM</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
cultural differences in socially desirable reporting are moderated by the mode of survey administration. Acquaviva (1994) found that the mode effect on reports of using drugs and alcohol (i.e., more use reported in self-administered questionnaires than in face-to-face interviews) was larger for African-American and Hispanic respondents than for whites. Johnson and van de Vijver suggest that this may be related to greater privacy concerns among members of vulnerable minority groups when they are asked to report socially undesirable behaviors than among members of the predominant cultural group.

These findings suggest that survey responses might be differently affected by mode across cultural groups, but the evidence is just suggestive. The definitive controlled studies have not been done. Moreover, there are no studies to our knowledge that investigate whether culture interacts with mode for cutting edge survey modes like video, text chatting, speech dialogue systems, and Web-questions with computerized animated agents. Nonetheless, we can derive predictions from studies of CMC and culture about how culture might affect survey interviews across different media. For example, much of the interaction that has been observed in survey interviews involves "paradigmatic" sequences (see Schaeffer and Maynard, Chapter 2 in this volume); that is, the interviewer asks the question, the respondent gives a problem-free answer, and the interviewer acknowledges this answer, sometimes by simply asking the next question. But the example exchanges in Table 11.1 suggest that belief among survey methodologists that paradigmatic sequences are the norm (Ongono, 2005) may be more culture specific than we have realized. Perhaps in Chinese or more generally East Asian interviews, it is typical for the face-to-face interaction between interviewer and respondent to involve more turns and more checking that the parties understand each other (grounding) than in American (or Western) interviews. If so, this seems likely to promote more accurate understanding and, as a result, accurate responding to the East Asian than Western interview questions as grounding has been shown to affect response accuracy (e.g., Conrad and Schöber, 2000; Schöber and Conrad, 2002; Schöber et al., 2004). However, the example also suggests that when interviews are conducted through a medium like IM, these cultural differences would go away and that both East Asians and Western interviews would be brief and more likely to follow the paradigmatic pattern. The reduction in grounding this would imply for Asian respondents could signal reduced comprehension accuracy, but at levels similar to the Western counterparts. While this is just one of our guesses about what might happen in the interview media, this is the kind of connection we will attempt to establish in this chapter: we will consider the implications of the results from CMC and culture studies for survey interviews across different media with respondents and interviewers from different cultures.

We encourage the reader to keep in mind survey interviewing through computers is similar to and different from the kind of communication in CMC studies. When a respondent completes a Web-based questionnaire, he/she communicates with the survey researchers through a computer, but this "conversation" does occur in real time and may never occur on an individual basis as it does in the collaborative tasks that characterize CMC research. When an interviewer is part of the data collection, the conversation is more individualized—the interviewer asks questions and records the respondent's answers; however, when the interviewer enters the responses into a computer, it is the interviewer, not the computer, that is the intermediary (Clark and Schenker, 1992). Despite these differences, we believe the interview is similar enough to make CMC tasks so that what is known about CMC and culture can at least stimulate thinking about the role of culture in the use of future interviewing technologies and the quality of the data they are used to collect.

One more caveat before we begin our discussion of CMC research. Culture is obviously a complex and nuanced construct (e.g., Miller, 2002). At least in the early stages of research on culture and communication, culture is operationalized with broad brush strokes (e.g., collective versus individualistic societies) that may feel overly simple to many readers. This is in part a result of conducting relatively small-scale laboratory studies in which simply as possible to enroll enough participants to span the range of cultural diversity that may be necessary to do justice to some distinctions. However, even relatively broad distinctions seem to have some measurable effects on the way participants communicate through different media as, at least a starting point, studies of culture at this level seem appropriate. Indeed, the early work on culture and survey responding mentioned earlier has proceeded with similarly broad distinctions. So despite the relatively high level at which culture is characterized in current CMC research, the effects observed in that literature may well transfer to communication through computational media in survey data collection tasks.

11.2 Introduction to CMC and Culture

A number of well-developed theories, based on evidence from Western participants, can be used to generate predictions about which media might work best for a given set of people performing a given set of tasks (e.g., Clark and Brennan, 1991; Doi and Longhi, 1984; Putnam, et al., 2002; Short et al., 1976; Witten, 1992, 1995), and a number of investigators have begun to examine cultural effects on CMC (e.g., Androutsos and Ellis, 2001; Kayan, et al., 2006a; Reineg and Mejía, 2003, 2004; Sheppard, et al., 2004; Zheng et al., 2006) in a variety of technologies and cultures, using a variety of research methods. The results to date suggest that people's use of CMC tools is influenced by their cultural background.

Adding issues of culture into the CMC mix complicates matters in interesting and important ways. Cultures vary along a number of dimensions that may affect group processes and outcomes, such as individualism versus collectivism (e.g., Hofstede, 1983; Triandis, 1995), low versus high context of communication (how much contextual information is required for communication; Hall, 1976), and task versus relationship orientation (whether people focus on getting work done or on establishing rapport with their partners; e.g., Triandis, 1995). These and other cultural dimensions may interact with features of media, such as the availability of visual cues, to create different effects on interaction and data quality in interviews and self-administered interviews.

We first present a conceptual framework to investigate how culture and CMC shape communication processes and task outcomes in general. Then we review research on each component of this framework, highlighting findings that we believe
broad implications for adopting new technologies in surveys. Note, however, that the findings thus far are rarely from studies carried out in interviewing settings but come from other cultures, and our primary contribution is in raising important questions that deserve future research in the survey context.

11.3 BACKGROUND

The theoretical framework we use to examine relationships between culture and CMC is an Input-Process-Output (I-P-O) model (Hackman, 1987), shown in Fig. 11.1. Here, culture and media are inputs that people bring to collaboration. These inputs, both allow and in interaction, influence communication processes and, in turn, subjective and objective outcomes. There are also a number of moderating variables that may influence relationships between inputs and processes and between processes and outcomes. Although the I-P-O framework is a simplification, it can help us conceptualize how culture and CMC interact by explicating relationships between inputs, process, outputs, and moderating variables. In the survey interview we conceive of the I-P-O sequence at the level of individual questions-answer exchanges, where the input includes the question as well as culture and particular modality process is the dialogue between respondent and interviewer (or interviewing system); and the outcome is the response, which in turn becomes one of the inputs to the process of answering the next question.

11.4 DIMENSIONS OF CULTURAL VARIABILITY

In cross-cultural research (both on intracellular and intercultural interaction), there has been substantial debate about the definition of culture, as well as about the number, size, and significance of dimensions along which cultures vary (e.g., Hofstede, 1983; Oyserman et al., 2000; Schwartz, 1992; Triandis, 1995). For the purposes of this chapter, we define culture as a set of norms, roles, and values

emphasized by a culture and adopted, to greater or lesser degrees, by members of that culture through such processes as imitation and teaching. We focus on three cultural dimensions—individualism-collectivism, high vs. low context of communication, and task versus relationship focus—that affect processes central to collaborative work (of which interviewing is a special case). These dimensions are not intended as an exhaustive description of how cultures differ but rather as a way of focusing our investigation on those dimensions most likely to influence what happens in new interviewing technologies.

11.4.1 Individualism-Collectivism

While all anthropological culture theories distinguish between individualistic cultures, in which people tend to identify themselves as individuals and focus on their own personal goals, and collectivistic cultures, in which people identify themselves as members of a collective and focus on the betterment of that collective (e.g., Hofstede, 2001; Triandis, 1995), Nisbett (2003) describes a wide range of cognitive processes affected by membership in individualistic versus collectivistic cultures, including reasoning styles and memory processes. Hofstede’s (2001) analysis of survey responses from a global sample of IBM employees show how individualism/collectivism is associated with preferences for business practices, child-rearing, and many other aspects of culture. Markus and Kitayama (1991) show that individualism/collectivism is associated with people’s concept of themselves as either independent of or interdependent with other individuals.

11.4.2 High Versus Low Context of Communication

Hall (1976) proposed that cultures vary in how much contextual information is required for communication. Low context, typically Western communication is verbally explicit, to the point, with relatively little attempt to mask one’s feelings. In contrast, high context, typically Eastern communication is indirect, often ambiguous, and sensitive to the context in which it occurs (e.g., the relationship between speaker and addressee, nuances of facial expressions or tone of voice). Much of the research on communication styles has used a self-report methodology, in which people respond to questions such as “Teach on to what others means when they do not say it directly” or “My speech tends to be very picturesque” (Gudykunst and Ting-Toomey, 1988; Gudykunst et al., 1996). These studies typically show that while people in all cultures use both styles, low context communication is preferred in individualistic societies and high context communication is preferred in collectivistic societies (Gudykunst and Ting-Toomey, 1988; Gudykunst et al., 1996). As we will discuss further later, cultural differences in use of indirectness in speech have been further supported through analyses of actual conversations in face-to-face and mediated settings.

11.4.3 Task Versus Relationship Focus

A third dimension of cultural variation pertinent to new forms of interviewing is task versus relationship orientation (Triandis, 1995). Task-oriented cultures such
CULTURE, COMPUTER-MEDIATED COMMUNICATION

as the United States, Canada, and Australia focus on getting work done, whereas relationship-oriented cultures such as Japan, Korea, and China focus on establishing rapport with one’s partners. The task versus relationship focus is only quasi-independent of the other dimensions: cultures identified by Triandis as task-oriented overlap substantially with those categorized as individualistic (Hofstede, 2001) and those described as low context communication (Hall, 1976; Gudykunst and Ting-Toomey, 1988). Similarly, cultures identified by Triandis as relationship-oriented overlap with those identified by other researchers as collectivistic and high context communicators.

Dimensions like individualism-collectivism are often applied at the national level (Hofstede, 2001; Triandis, 1988), but it has become increasingly obvious that to understand links between culture and communication, it is necessary to examine how national values are related to individual’s personal values (Schwartz, 1992) and their conceptions of themselves as interdependent versus independent (Markus and Kitayama, 1991; Singelis and Brown 1995). Both individual values and individuals’ self-concepts are influenced by national culture but not entirely determined by it. Gudykunst et al. (1990) and Oetzel (1998), among others, have shown that national values, individual values, and self-concepts each have an impact on self-reported communication behavior. Generally, these studies have looked at the impact of these variables in the abstract, independent of any given communicative domain. Thus, the question of how the results would pertain to the interviewing context remains open for investigation.

11.5 AFFORDANCES OF MEDIA

To build a theoretical model of how culture interacts with features of communication media, it is essential to characterize media at the right level of analysis. A number of theories distinguish media along a single dimension such as media richness (Daft and Leong, 1984). For our purposes, single-dimension theories do not differentiate clearly enough among media. Instead, we draw on Clark and Brennan’s (1991) influential theory of media affordances, which provides a fine-grained analysis of the resources media provide for communication (see Table 11.2). For example, telephone calls and videoconferencing provide modality, and thus afford the use of speech, whereas E-mail does not. In this framework, communication across different media will entail different costs for producing messages, receiving and understanding messages, changing speakers, and repairing misunderstandings.

Newer modes for survey interviewing will certainly require extensions of the above framework. For example, in Web-based surveys and audio-CASI, the conversational partner is explicitly present during the interaction but the dynamics of the medium may enable the survey researcher to impose a sense of presence. This kind of effect has been demonstrated by Braden and Mark (2001), who found social presence effects were as strong in a computer-based math task when people used an application-sharing tool as when they were observed via two-way video (even though they could not be sure anyone was really watching). In these cases there might be a gray "Y" in the co-temporality row of the table.

11.6 APPLYING THE FRAMEWORK TO INTERVIEWING

Starting from Clark and Brennan’s model, dimensions of cultural variability may alter the perceived importance of affordances such as modality, visibility, and co-presence. For example, Gudykunst and Kim (1997) suggest that in several cases may be more important for communication in high context cultures because the meanings of messages resides in the situation context, not in the words themselves. Thus, we might anticipate that visibility will be more important for successful communication among members of high context cultures than for members of low context cultures. If this is the case, then one should expect that new interviewing technologies that affect visibility of partners—the respondent being able to see the interviewer, the interviewer being able to see the respondent, during question asking, during answers—should matter differently for members of different cultural groups. In fact, as we will see later, there are a few pieces of evidence from current interviewing methods that are consistent with this idea. Let us first examine how culture and media have been shown to influence conversational processes and team outcomes in multireviewing settings.

11.7 COMMUNICATION PROCESSES

Culture may, alone or in interaction with features of media, influence group processes—particularly processes of communication. Here, we focus on two aspects of mediated communication that we view as essential for successful understanding in interviews—conversational grounding and relational communication—and review prior work on the ways in which culture and media affect these two communication processes.
11.7.1 Conversational Grounding: The Basis for Question Comprehension

Questioning in survey interviews consists of words that respondents need to interpret, and the cognitive and interactive resources that respondents use are those that they use to understand what their partners say more generally (Scholte, 1999). Conversational grounding refers to the interactive process by which communication exchange evidence in order to reach mutual understanding (Clark and Brennan, 1991; Clark and Schopler, 1992; Clark and Wilkes-Gibbs, 1986). Speakers and listeners work together by asking questions, providing clarifications, and other procedures to ensure that messages are understood as intended. Grounding is easier, and conversation more efficient, when collaborators share common ground—mutual knowledge, beliefs, and so on (Clark and Marshall, 1981). This common ground can arise from membership in social groups (e.g., Russell and Krauss, 1992; Guare and Clark, 1987), through the process of exchanging messages (linguistic copresence), or by sharing a physical setting (physical copresence).

In Clark and Brennan's (1991) framework, affordances of media influence the strategies people use to ground their utterances. For example, face-to-face settings afford visibility and physical co-presence, so speakers can use gestures to refer efficiently to task objects (e.g., Bickmore et al., 1995; Clark and Krych, 2004). In media that lack visibility and physical co-presence, speakers must use longer verbal descriptions of the same objects (e.g., Delhery-Staddon et al., 1997; Kraut et al., 2003). A substantial body of research supports the conjecture that features of media influence grounding. For example, conversation is more efficient when technology provides a shared view of the workspace (e.g., Gergle et al., 2004; Kraut et al., 2003) and when tools allow people to gesture in that workspace (Russell et al., 2004; Kirk and Stammer-Fuerer, 2006).

How these findings apply to interviews is an important question, as in interviews it is rare that the questions refer to what is in the immediate physical environment or shared workspace. Nonetheless, the evidence from survey interviews and telephone interviews can indeed lead to more accurate question interpretation (Conrad & Scholte, 2002; Scholte and Conrad, 1997; Scholte et al., 2004); when interpretation can be clarified the interaction is less efficient (takes longer) but can lead to better answers and thus better data quality. The evidence is also that the effects of being able to ground can extend to both text-based self-administered interviewing (Conrad et al., 2007) and to speech-based self-administered interviews (Ehren, Scholte, and Conrad, in press). Visibility seems at least indirectly related to grounding in survey interviews. In a comparison of face-to-face and telephone interviews (Conrad et al., 2007) respondents provided more coherently articulated responses provided spoken cases of comprehension difficulty (e.g., 'are you sure?') more often on the phone than face-to-face, presumably to compensate for the absence of visual cues of uncertainty (e.g., facial evidence that the respondent is confused like a furrowed brow or looking away from the interviewer while answering). So how culture and media interact to affect the ability to ground understanding is particularly relevant when considering adopting new interviewing technologies.

Communication Processes

And there is indeed some evidence that cultures can vary in their strategies for grounding meaning in conversation (Li, 1999a,b). Hall (1976) proposed that availability and visibility may be more important for grounding in high context cultures than in low context cultures, because awareness of how others are reacting to one's messages is an important aspect of high context communication. This notion is supported indirectly by Voinov and Colleagues (1999), who examined how well pairs could perform a map-based task in which one person gave directions and the other had to draw the identical route on his/hers own map. Voinov and colleagues found that some native English speakers, many of whom were Asian, benefited from video over audio conferencing, whereas native English speakers did not. They infer that the former cases to mutual understanding provided by visibility (e.g., visual clues, hand gestures) were especially valuable for nonnative speakers. However, this study confounded native language with intercultural communication, so we don't know which factor accounts for the result.

In one of our own studies (Seidloch et al., 2004), we compared American, Chinese, and mixed American-Chinese dyads performing scenario-based negotiation tasks face-to-face or via IM. The goal of these tasks is to rank salvaged items from a crash in order of importance. Pairs first rank the items individually, then negotiate until they come to agreement on a joint ranking. We hypothesized that the lack of visual cues in IM would make it poorly suited for communication among members of high context cultures but not affect communication among members of low context cultures. Consistent with this hypothesis, we found no difference between media in terms of how much grounding American pairs required to complete the task, but a large impact of medium for Chinese pairs who spoke much more face-to-face (see the example interactions in Table 11.1). This culture by medium interaction is displayed in Fig. 11.2.

![Figure 11.2: Mean ranking time per task by culture group and medium (AA = American only, AC = mixed American-Chinese, CC = Chinese only).](image-url)
As we showed in Table 11.2, face-to-face interaction has many affordances not present in IM, including audibility and visibility. To assess which of these two affordances was more important, Sedlick et al. (2007) compared the same cultural groups interacting via audio or video conferencing. No main effect of culture, nor a culture by medium interaction, was found. Thus, the increased speech in face-to-face interactions for Chinese dyads seems to result from the presence of auditory information but is not further benefitted by adding visual cues via video conferencing. These results conflict with those of Veach et al. (1999) and suggest the need for a more detailed examination of factors that differ across the two studies (e.g., tasks, specific cultural backgrounds of participants). Possibly the most relevant difference between the two studies for our current purposes is whether a physical artifact was part of the task (i.e., a map). In this respect the verbally based negotiation task used by Sedlick and her colleagues is more similar to a typical interviewing situation where the questions refer to events and attitudes that are not present or visible. Thus, it would seem that telephone interviews should increase conversational grounding for Chinese respondents relative to IM interviews (or interviews in any mode with Westerners) but that face-to-face interviews are unlikely to lead to more grounding.

11.7.2 Relational Communication and Rapport

In addition to the cognitive aspects of survey interviews, socioemotional aspects also matter, affecting not only respondents' willingness to participate, but their motivation to provide thoughtful and accurate answers, to answer sensitive questions, and to complete interviews. These aspects of interviewing are related to the larger literature on relational aspects of communication, which are concerned not only with what information is conveyed but with how that information is conveyed and what it indicates about the relationship between the speaker and addressee(s). Much research on relational communication has focused on nonverbal cues such as eye gaze (e.g., Argyle and Cook, 1976), facial expressions (e.g., Ekman, 1982), and posture (e.g., Mahlen, 1967), which can be used to indicate intimacy, trust, and attraction. In addition, messages can be formulated in different ways to establish, maintain, and/or build closeness with a partner. In Table 11.3, we list some of the ways that nonverbal, paralinguistic, and verbal cues can add socioemotional meaning to people's messages. A key aspect of relational communication is face maintenance, or ensuring that one does not cause another person to lose respect (Goffman, 1967). Linguistic politeness

<table>
<thead>
<tr>
<th>TABLE 11.3 Basic Categories of Relational Communication</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Category</strong></td>
</tr>
<tr>
<td>Nonverbal behavior</td>
</tr>
</tbody>
</table>
| Paralinguistic behavior | Intonation patterns, pitch, pauses, 
| Verbal behavior    | Forms of address (e.g., *John, Mr. Jones*), 
|                   | pronouns (e.g., *we*), hedges (e.g., *sort of*), intensifiers (e.g., *very*), 
|                   | indirect requests (e.g., *would you mind...*), 
|                   | uses words |

refers to a range of strategies by which people demonstrate concern for their own and others' faces (Brown and Levinson, 1978). For example, indirect requests such as, "Could you close the door?" are more polite than directives such as, "Close the door!" (Holmes, 1997). Similarly, hedging an opinion (e.g., "I think you might be wrong") is more polite than directly stating that opinion (e.g., "You are wrong"). Misunderstandings can arise when one partner places less emphasis on relational concerns than another partner.

Features of media have been shown to affect relational communication (e.g., Herzig, 1994; Kiesler et al., 1983; see Whitman, 2003 for a review). In some cases, greater negative emotion and "flaming" has been found in text communication (e.g., Kiesler et al., 1983), a finding that was attributed to Kiesler and colleagues to the lack of social context cues in text communication but which other researchers have attributed to the difficulty of producing polite forms in typed discourse (e.g., Furnham and Ghauri, 1999). Other studies have found that relational communication in face-to-face settings is less found in text communication (e.g., Hitlin and Turndorf, 1978). Early studies also suggest that relational aspects of communication are reduced when conversations take place over the phone versus face-to-face (Ruthe, 1987; Stephenson et al., 1976). Less research has compared relational communication in audio versus video conferencing, although many media theorists (e.g., Duff and Long, 1994; Short et al., 1976) suggest that video will better support relational communication.

The cultural theories outlined earlier suggest that high-context, relationship-oriented cultures place more emphasis on relational communication than do low-context, task-oriented cultures (Ting-Toomey et al., 1991). This hypothesis has been supported in studies of conversational indirectness, where high-context cultures such as China and Korea use more indirectness than low-context ones such as the United States (Amabile et al., 1996; Holmgreen, 1997). Chinese speakers are also more likely to use "we" pronouns and social language than American speakers, both face-to-face and via IM (Sedlick et al., 2000). Further support comes from cross-cultural research on negotiation, which has shown that relational strategies (e.g., compromising) are favored by high-context negotiators whereas informational strategies (e.g., dominating the conversation) are favored by low-context cultures (e.g., Adair and Brint, 2002; Adair et al., 2001; Ting-Toomey et al., 1991). Such differences have been attributed to cultural variation in concerns for one's own face versus the other person's face (Ting-Toomey, 1983).

Although it has not to our knowledge been directly tested, we anticipate interactions between the cultural background of communicators and the affordances of a medium on the amount and valence of relational communication. High context communicators may especially rely on facial expressions and tone of voice when producing and interpreting relational cues, whereas low context communicators may find verbal substitutes such as hedges and indirect requests to be adequate substitutes. This leads to a set of tentative predictions for new interviewing technologies: high context interviewees should be likely to build rapport with interviewers using nonverbal media that support visibility and audibility, whereas low context respondents' motivation and satisfaction should be less affected by visual and auditory affordances. For example, we might expect reduced bias or even a bias from A-CASS for participants.
from high cultures versus low context cultures. Respondents from low context (i.e., Western) cultures seem to feel more private answering questions posed via A-CASI than by face-to-face interviews; based on the increase in reports of sensitive behavior in the former versus the latter mode (e.g., Tourangeau and Smith, 1996; Turner et al., 1995). However, for high context cultures, the fear of a human voice is displayed under A-CASI may overwhelm the sense of privacy and lead to more socially desirable (i.e., less candid) responding. Similarly, an anonymous and disengaged agent may increase rapport with low context cultures, but it may lead to an overinterpretation of affective cues by high context cultures—for example, if the interviewing agent's smile is poorly timed or inappropriate, this might distract high context respondents or lead them to feel they are not performing adequately (see Persson, D'Mello, and Oleson, Chapter 10 in this volume, for a discussion of affective agents and rapport).

Furthermore, one can imagine that without some way to ground interviewer affect in a neutral medium such as DM (e.g., the use of "smileys" like :D), high context respondents may ascribe affect to the interviewer that is not warranted, much like leaving the interpretation of context undefined in standardized interviews leads to more variation in how the terms are interpreted, including unintended meanings (see Scherer and Conrad, 2002; Suszerech et al., 2005).

11.8 OUTCOME MEASURES

In our F-P-O model in Fig. 11.1, inputs (culture and features of technology) impact communication processes, which in turn impact a range of outcome measures. Here, we touch briefly on several outcome measures that are especially important in intercultural teamwork and that are likely to affect the quality of survey data: subjective evaluation of one's partner, communication, trust, and objective performance measures. This set of outcome measures, common to CMC research, maps only partially to the outcomes of interest in survey research. For example, in CMC studies, subjective outcomes like partner perception and trust are typically measured at the end of the conversation, under the assumption that such outcomes will have bearing on future interactions between the same individuals. In survey research, future interactions between the surveyor and respondent may be quite unlikely. Instead, we might conceptualize partner perception and trust as outcome measures at the end of each question-answer pairing, outcomes that build up over the course of the survey interview. In addition, CMC studies have rarely if ever considered whether the answers people provide are valid or reliable, so additional research will be needed to understand how the inputs and processes in our F-P-O model affect such outcomes.

11.8.1 Perception of Partners

Features of media can affect collaborators' impressions of each other. For example, early studies found greater liking for partners when using video versus audio alone (Short et al., 1976; Williams, 1977). Hancock and Dunham (2001) suggest that the lack of social cues in text CMC creates ambiguity that affects one's impression of a partner. When less is known about a remote collaborator's immediate experiences, problems such as delays and awkward expressions are more likely to be attributed to internal, dispositional factors (e.g., moods) rather than external causes (e.g., network delays) (Cronin, 2001). Such effects may be especially strong in intercultural interaction, in which people share less initial background knowledge. Consistent with this, Setlock et al. (2004, 2007) found that members of intercultural dyads rated each other more negatively than homogeneous American or Chinese dyads. However, the tendency to attribute behavior to dispositional factors in part culturally specific: East Asians are more likely to consider situational explanations for behaviors than Westerners (Choi et al., 1999; Morris and Peng, 1994). We anticipate that media that reduce or eliminate visual and auditory cues would have a greater impact on partner perceptions in high context than low context cultures. Thus, as social cues are reduced across interviewing modes (e.g., face-to-face vs telephone vs IM vs internal Web-based questionnaires), respondents from high context cultures may increasingly attribute lack of interest or disapproval to the interview (or interviewing system). The danger in this kind of perception could lead high context respondents to terminate the interview, whereas low context respondents may experience some of this.

11.8.2 Persuasion

The relatively high response rates in face-to-face interviews (versus telephone interviews versus Web survey participation) are often attributed to the interviewer's persuasive abilities, which are more effectively applied when the interviewer is physically present and the respondent cannot make the interviewer disappear by hanging up the phone or ignoring an e-mail invitation. How might this differ cross-culturally, especially with new and emerging interview technologies? The relevant studies have not been conducted yet, but the CMC literature is again instructive.

Persuasion in CMC refers to the extent to which one team member can convince others that his/her viewpoint is correct. Early studies indicated that persuasion varied as a function of medium (e.g., Clark and Hagey 1976, Oudshoorn and Coddington, 2002). At the end of the conversation, under the assumption that such outcomes will have bearing on future interactions between the same individuals. In survey research, future interactions between the surveyor and respondent may be quite unlikely. Instead, we might conceptualize partner perception and trust as outcome measures at the end of each question-answer pairing, outcomes that build up over the course of the survey interview. In addition, CMC studies have rarely if ever considered whether the answers people provide are valid or reliable, so additional research will be needed to understand how the inputs and processes in our F-P-O model affect such outcomes.
lead to the following hypothesis: interviewers might be more effective in persuading potential respondents to participate if the respondent can see the interviewer, as in a video introduction (see Fuchs, Chapter 4 in this volume, for a discussion of video introductions in mobile Web surveys) and that this might be particularly effective for high context groups.

11.8.3 Trust
Trust is an individual's confidence in the goodwill of others and the expectation that others will reciprocate if one cooperates (e.g., Ring and Van de Ven's, 1994). McAllister (1995) differentiates two broad foundations for trust in organizational settings: cognitive and affective. Cognitive trust is built on people's intelligence, competence, and reliability, whereas affective trust is built on people's emotional bond and relationship. Several studies suggest that establishing trust is more difficult in remote collaborations than in face-to-face, and more difficult with lesser status media than with richer media like audio or video conferencing (Box et al., 2002), although having initial face-to-face interaction before working at a distance seems to help (Jessen et al., 2000; Recom, 1998). We would expect affective trust to be weighted more heavily in high context, relationship-oriented cultures and cognitive trust to be weighted more heavily in low context, task-oriented cultures. For high context cultures, establishing affective trust in lesser media may be especially difficult. Consistent with this, Zhang, Olson, and Olson (2004) found that Chinese pairs showed higher affective trust when negotiating by video conferencing than by audio conferencing, whereas American pairs showed no differences on either affective or cognitive trust in both media.

In the domain of survey interviewing, trustworthiness of an interviewer can strongly affect respondents' willingness to report sensitive or less socially desirable behaviors, and their tendency to overreport socially desirable behaviors. There is a small body of evidence suggesting that culture and social media interact on this score (e.g., Acquisto, 1994; Johnson and van de Vijver, 2003). One can extend this to emerging technologies such as automated interviewing agents. Imagine an agent that communicates warmth (e.g., it smiles and is polite) but lacks conversational ability (e.g., cannot clarify the questions it asks) and imagine an agent with the opposite characteristics (i.e., lacks warmth but possesses the ability to clarify question meaning). The CMC evidence on trust and culture just discussed would lead us to expect that high context respondents would trust the first interviewing agent more than the second, answering sensitive questions asked by the former with greater candor than the same questions asked by the latter. In contrast, respondents from low context cultures might be more likely to disclose sensitive facts to the second agent.

11.8.4 Objective Performance
The 1-P-O framework ultimately concerns task performance. In surveys, there are various straightforward indicators of objective performance, which align with the CMC research to varying degrees. Response accuracy and reliability are the most obvious indicators. The extent to which answers are influenced by question wording, question ordering, or response options are another possible indicator, with the assumption that less influenced answers are more likely to be accurate and reliable. There are various pieces of evidence on how media affordances affect people's performance on joint tasks in noninterviewing domains though, again, relatively little in surveys. We turn to the CMC results first.

Unexpectedly, the number of utterances spoken during a task is often significantly correlated with task completion time (Gergle et al., 2004; Krue et al., 2003). Thus, performance times are generally shorter when a medium allows for more efficient communication. When performance is measured in other ways, however, the effects of media are less clear. For example, Doherty-Sneddon et al. (1997) found no differences in the accuracy of map routes described in video versus audio conferencing; Jackson et al. (2000) found no effect of video frame rate on the quality of paper designs; and Stoney and McGraw (1994) found no performance differences between text-based and face-to-face interaction on idea generation, interactive, or judgment tasks. Theories of CMC suggest that tasks involving negotiation and persuasion will be more influenced by communication medium than tasks requiring less interpersonal finesse (e.g., Dalt et al. 1987; Short et al., 1976). Given so many studies using these sorts of tasks have found no differences in performance quality (e.g., Hill et al., 1986; see Whitaker and O'Connor, 1997, for a review), few studies have looked at how culture influences performance quality. Li (1996a, b) found no differences in accuracy of information transmission in Canadian and Chinese dyads, but significantly poorer transmission in mixed Canadian-Chinese dyads. Others (e.g., Addie et al. 2001) have found that intracultural teams perform more poorly on negotiation tasks.

There really is little evidence on how culture and media interact to affect performance in interviews but there are several demonstrations that culture can affect answers. For example, Ji, Schwartz, and Nisbett (2000) found that a well-known response scale effect—observed with Western respondents—only partly replicates with Chinese respondents. The original finding (e.g., Memon et al., 1995) was that there is a tendency for respondents to treat the middle scale value as reflecting the average frequency in the population and position themselves relatively close to what they see as the "average." However, when Ji, Schwartz, and Nisbett (2000) replicated the study with both American and Chinese respondents, the Americans reproduced the earlier finding but the Chinese did not. The Chinese only showed the effect for unobservable behaviors such as having nightmares but not for observable behaviors such as coming to class late. The authors suggest that in an independent culture such as Chinese culture, respondents are far more attuned to the behavior of others than in a dependent culture such as American and German culture. As a result, Chinese respondents know a great deal about the frequency of behaviors they can observe and are less tempted to consult the response scales for distributional information than are their American counterparts.

These findings come from self-administered questionnaires. It is possible that introducing an interviewer whose appearance provides clues relative to the frequency of the relevant behavior (e.g., the interviewer asks about frequency of exercise and looks relatively fit) might have more of an impact for American than Chinese respondents.
11.9 NEW INTERVIEWING TECHNOLOGIES THAT CONSIDER CULTURE

Considering culture in developing new interviewing systems raises a large number of new, untried questions, and new interviewing technologies will require answers about basic behavioral questions on culture and media in order to make the necessary design choices. Say it turns out to be true that visual cues are particularly helpful for increasing data quality, trust, and interview satisfaction for one cultural group, while not mattering much or even hurting data quality for another group. This would suggest that allowing respondents to have visual access to the interviewer (via the screen) to today's video conferencing, video Skype, or (Chat technology) should be a central consideration in designing the interviewing system, and perhaps that visual access should be a feature that can be turned on and off for different respondents—either at their request, or based on the survey system's diagnosis of the respondent's cultural background.

The same questions turn out to be relevant for self-administered interviewing systems. With current interview-administered surveys, various aspects of the interviewer are unalterable, and the only place to choose an interviewer's acceptability (on all unalterable dimensions) is in hiring, and (on any alterable dimensions) in training and monitoring. But imagine, for example, designing a new interviewing medium that includes spoken interaction with an interviewing system. A number of culturally relevant choices need to be made, including:

- What accent and pronunciation should the recorded or text-to-speech "interviewer" use? Should the accent be the same as the respondent's? Should the accent and voice tone adjust to the cultural background of the respondent, in order to increase trust and comfort with the interviewer?
- What kinds of interruption in the dialogue should be allowed? Is there substantial evidence on cultural variability in interruption styles (Schifrin; Thoman), with certain high context cultures that not only allow but expect substantial overlapping speech as signs of interest and attentiveness, and other low context cultures that find overlapping speech offensive and intrusive? Should an interviewing system ever interrupt a respondent? Should a respondent be allowed to interrupt the interviewer?
- If the effects of facial cues on data quality are shown to be high for respondents from a particular culture but low for others, should the system add a face for the respondents who will be helped by one?

REFERENCES

For example, a new interviewing tool might automatically modify messages to be more appropriate for the recipient's cultural background, similar to real-time translation software (Nomisubawa and Isakawa, 2006). Alternatively, interviewing tools might seek to educate interviewers about respondents' cultural backgrounds, or respondents about interviewers' backgrounds, for example, by informing the sender of a message so to why it might be inappropriate given the recipient's culture.

11.10 CONCLUSION

In this chapter we highlight some of the important considerations for cultural and behavioral differences in survey research. As survey designers consider the design of their communication systems, they should be aware of cultural and behavioral differences and how these may affect the design of their communication systems. More specifically, people from cultures that emphasize nonverbal and contextual aspects of communication are more affected by the visual and auditory affordances of communication media than are people from cultures that emphasize the verbal aspects of communication. Designers of interviewing systems can be affected by the use of different communication modalities and should be aware of the potential impact on the design of their communication systems.

REFERENCES


Envisioning the Survey Interview of the Future

Edited by
FREDERICK G. CONRAD
Institute for Social Research
University of Michigan
Ann Arbor, Michigan

MICHAEL F. SCHÖBER
New School for Social Research
New York, New York

Each generation has its unique needs and aspirations. When Charles Wiley first opened his small printing shop in lower Manhattan in 1807, it was a generation of boundless potential searching for an identity. And we were there, helping to define a new American literary tradition. Over half a century later, in the midst of the Second Industrial Revolution, it was a generation focused on building the future. Once again, we were there, supplying the critical scientific, technical, and engineering knowledge that helped frame the world. Throughout the 20th Century, and into the new millennium, nations began to reach out beyond their own borders and a new international community was born. Wiley was there, expanding its operations around the world to enable a global exchange of ideas, opinions, and know-how.

For 200 years, Wiley has been an integral part of each generation’s journey, enabling the flow of information and understanding necessary to meet their needs and fulfill their aspirations. Today, bold new technologies are changing the way we live and learn. Wiley will be there, providing you the must-have knowledge you need to imagine new worlds, new possibilities, and new opportunities.

Generations come and go, but you can always count on Wiley to provide you the knowledge you need, when and where you need it.

William J. Feakes
President and Chief Executive Officer

Peter Booth Wiley
Chairman of the Board