Intra- and Inter-Cultural Usability in Computer-Supported Collaboration

Ravi Vatrapu
Assistant Professor
Center for Applied ICT
(CAICT)
Copenhagen Business School
Howitzvej 60, 2nd floor
Frederiksberg, DK-2000
Denmark
vatrapu@cbs.dk

Dan Suthers
Professor
Dept. of Information and
Communication Sciences
University of Hawai`i at
Manoa
POST 309,1680 East West RD
Honolulu, HI, 96826
USA
suthers@hawaii.edu

Abstract

In this paper, we argue for an increased scope of universal design to encompass usability and accessibility for not only users with physical disabilities but also for users from different cultures. Towards this end, we present an empirical evaluation of cultural usability in computer-supported collaboration. The premise of this research is that perception and appropriation of socio-technical affordances vary across cultures. In an experimental study with a computer-supported collaborative learning environment, pairs of participants from similar and different cultures (American-American, American-Chinese, and Chinese-Chinese) appropriated affordances and produced technological intersubjectivity. Cultural usability was analyzed through the use of performance and satisfaction measures. The results show a systemic variation in efficiency, effectiveness, and satisfaction between the two cultural groups. Implications of these findings for the research and practice of usability, in general, and cultural usability, in particular, are discussed in this paper.

Keywords

affordances, apperception/perception/appropriation of affordances, comparative informatics, culture, cultural usability, computer supported intercultural collaboration, computer supported collaborative learning, technological intersubjectivity

Usability Professionals' Association
Promoting usability concepts and techniques worldwide

Introduction

Usability, accessibility, and universal design have been three of the central concepts in the design of person-environment systems in general and human-computer interaction in particular (Iwarsson & Ståhl, 2003). Ron Mace (who coined the term *universal design*) said, "Universal design is the design of products and environments to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design"[1]. Based on this definition, in an increasingly multicultural world, universal design should encompass not only accessibility but also cultural usability. This would mean paying close and careful attention to cross-cultural variations in the design, development, deployment, and evaluation of Information and Communication Technologies (ICT). Such a broadening of the scope of the concept of universal design and ICT would require an examination of mono-cultural design and evaluation assumptions in the research and practice of usability. Towards this end, this paper presents an empirical evaluation of intra- and inter- cultural usability in computer-supported collaboration.

Cultural Usability

Cultural aspects of usability have been a topic of study in the field of human-computer interaction (HCI). Early research focused on localization and internationalization of user interfaces with respect to languages; colors; and conventions of data, time, and currency (Fernandes, 1995; Khaslavsky, 1998; Russo & Boor, 1993). Subsequent research investigated cultural influences on usability evaluation methods and usability processes. For example, cultural differences were found in the usability assessment methods of focus groups (Beu, Honold, & Yuan, 2000), think-aloud (Clemmensen, Hertzum, Hornbæk, Shi, & Yammiyavar, 2009; Yeo, 2001), questionnaires (Day & Evers, 1999), and structured interviews (Vatrapu & Pérez-Quiñones, 2006). Cultural differences with respect to usability processes were found in the understanding of metaphors and interface design (Day & Evers, 1999; Evers, 1998) and non-verbal cues (Yammiyavar, Clemmensen, & Kumar, 2008; Yammiyavar & Goel, 2006). Moreover, culture was found to affect web design (Marcus & Gould, 2000), objective and subjective measures of usability (Herman, 1996), and subjective perceptions and preferences in mobile devices (Wallace & Yu, 2009). Emerging findings also show that the understanding of the concept of usability and its associated constructs are culturally relative (Frandsen-Thorlacius, Hornbæk, Hertzum, & Clemmensen, 2009; Hertzum et al., 2007).

Culture, Collaboration, and Usability

Existing research in cultural usability has largely focused on aesthetic issues, methodological aspects, and practitioner concerns of stand-alone desktop applications and websites. Currently, there is little research on the cultural usability of computer-supported collaboration environments where members of different and similar cultures not only interact with the technology but also interact with each other through the technology. Given the social web (web 2.0) phenomenon that includes the participatory turn of the Internet, social sharing of media, cloud computing, and web services, there is a need to investigate the extent to which culture affects usability in socio-technical systems. In this paper, we present a usability analysis of an experimental study of intra- and inter- cultural computer-supported collaboration of Chinese and American participants.

Computer-Supported Intercultural Collaboration

Computer-supported intercultural collaboration (CSIC) is an emerging field of study centrally concerned with the iterative design, development, and evaluation of technologies that enhance and enrich effective intercultural communication and collaboration (Vatrapu & Suthers, 2009b). There are two interrelated aspects of interaction design in developing CSIC systems: (a) interacting with computers and (b) interacting with other persons using computers. Both these aspects of interaction can be influenced strongly by culture, given the strong empirical evidence documenting cultural differences in cognition (Nisbett & Norenzayan, 2002), communication (Hall, 1977), behavior (House, Hanges, Javidan, Dorfman, & Gupta, 2004), and interacting with computers (Vatrapu & Suthers, 2007). In line with the research program articulated in (Vatrapu, 2007), the research project discussed here originally focused on the influence of culture on (a) how participants appropriate affordances (Vatrapu, 2008; Vatrapu & Suthers,

---

[1] http://www.design.ncsu.edu/cud/about_ud/about_ud.htm

2009a) and (b) how participants relate to each other during and after computer-supported collaborative interaction (Vatrapu, 2008; Vatrapu & Suthers, 2009b). This paper focuses on usability aspects of the research project described in (Vatrapu, 2007). The analytical aim of this paper is to investigate subjective and objective aspects of cultural usability in computer-supported collaboration with conceptual representations. Before discussing the methodological aspects of the research project, the key definitions of socio-technical affordance, appropriation of affordance, and technological intersubjectivity are provided in the following sections.

Definition of Socio-Technical Affordance

In computer-supported collaboration, each actor is both a user of the system as well as a resource for the other users. *Technology affordances* (Gaver, 1991; Suthers, 2006) are action-taking possibilities and meaning-making opportunities in a user-technology system with reference to the actor. Similarly, *social affordances* (Bradner, 2001; Kreijns & Kirschner, 2001) are action-taking possibilities and meaning-making opportunities in a social system with reference to the competencies and capabilities of the social actor. In socio-technical systems that facilitate collaboration, technology affordances, and social affordances amalgamate into *socio-technical affordances* (Vatrapu, 2007, 2009b). For an outline of a theory of socio-technical interactions, see (Vatrapu, 2009b). Drawing upon foundational work in ecological psychology on the formal definition of affordances (Stoffregen, 2003; Turvey, 1992), the following definition is offered for socio-technical affordance.

Let $W = (T, S, O)$ be a socio-technical system (e.g., person-collaborating-with-another-person system) constituted by technology T (e.g., collaboration software), self-actor S, (e.g., artifact creator), and other-actor O (e.g., artifact editor). Let p be a property of T, q be a property of S, and r be a property of O. Let β be a relation between p, q and r, $p/q/r$ and β defines a higher order property (i.e., a property of the socio-technical system). Then β is said to be a socio-technical affordance with respect to W if and only if:

$W = (T, S, O)$ possesses β

None of T, S, O, (T, S), (T, O), or (S, O) possess β

The formal definition of socio-technical affordance provided above reflects the duality of individuals' perception with respect to the technology as well as other persons. The duality is essential and is present right in the middle of the tuple (T, S, O). Self-actor S needs technology T to interact with the other-actor O and vice-versa. Technology T should have the capabilities to support the interactional needs and dynamics of the actors S and O. That is, T should be able to support (by conscious or unconscious design) the interactional (communication and informational) needs and necessities of S and O. Moreover, the self-actor S needs to have the social as well as technical competencies and O needs to be in the intersubjective realm. Then and only then does an affordance qua affordance comes into being as an action-taking possibility and a meaning-making opportunity waiting for creative, generative, transformative, reformative, or simply repetitive appropriation. It is this appropriation of affordance that manifests as external action (see below).

The formal definition informed the design of an experimental study of computer-supported intra- and inter-cultural collaboration to be discussed shortly. In brief, the experimental design consisted of a systematic variation of two of the three elements—self (S) and other (O) with the technology (T) remaining invariant.

Definition of Appropriation of Affordances

Interactions in socio-technical environments are a dynamic interplay between ecological information as embodied in artifacts and individual actions grounded in cultural schemas. The essential mediation of all interaction is the central insight of socio-cultural theories of the mind (Wertsch, 1985, 1998). The conceptualization of interaction as being mutually "accountable" (observable and reportable) is the critical insight of ethnomethodology (Garfinkel, 1967). Following these two schools of thought, interactions in socio-technical systems are conceptualized as *accountable appropriation of socio-technical affordances*. Based on Stoffregen's (2003) discussion of behavior, appropriation is defined as "what happens at the conjunction of complementary affordances and intentions or goals" (p.125).

Research into social aspects of HCI (Reeves & Nass, 1996) has shown that even computer-literate users tend to use social rules and display social behavior in routine interactions with

computers. Social interaction is grounded strongly in culture as every person carries within patterns of thinking, feeling, behaving, and potential interacting. Thus, participants in computer-supported collaboration make culturally appropriate and socially sensitive choices and decisions in their actual appropriation of affordances. On these terms, the concept of appropriation employed here is similar to the notion of appropriation in adaptive structuration theory (DeSanctis & Poole, 1994; Orlikowski, 1992) inspired by Giddens (1986). Appropriation in adaptive structuration theory refers to the utilization of structural features of the system. Appropriation of affordances in our theory of socio-technical interactions refers to the intentional utilization of action-taking possibilities and meaning-making opportunities in a culturally-sensitive and context-dependent way. See (Vatrapu & Suthers, 2009a) for a report of how "representational guidance" (Suthers, Vatrapu, Medina, Joseph, & Dwyer, 2008) informed appropriation of affordances analysis of the study data.

Technological Intersubjectivity

Intersubjectivity is the fabric of our social lives and social becoming (Crossley, 1996). Information and Communication Technologies (ICT) and the Internet continue to transform our social relations with others and objects in fundamental ways. Our interactions with others and objects are increasingly informed by the operational logic of technology, hence technological intersubjectivity. Technological intersubjectivity is the production, projection, and ultimately the performance of intersubjectivity in socio-technical systems. Our psychological perception of and phenomenal relations with social others are being increasingly transformed by the advances in information and communication technologies and social software. For example, technology lets us assign distinct ring tones, images, or priorities to our significant others. Human beings are not only functional communicators but also hermeneutic actors. In technological intersubjectivity, technological mediation can sometimes (but not necessarily always) disappear like in Clarke's third law of technology (1962).

Definition of Technological Intersubjectivity

Technological intersubjectivity (TI) refers to a technology supported interactional relationship between two or more actors. TI emerges from a dynamic interplay between the technological relationship of actors with artifacts and their social relationship with other actors.

TI is an emergent resulting from psychological-phenomenological nexus of the electronic self–other social relationship. Psychological intersubjectivity refers to a functional association between two or more human communicators. Phenomenological intersubjectivity refers to an empathetic social relationship between two or more human actors.

From a functional perspective, psychological intersubjectivity doesn't require two or more persons to have the same or similar subjective experience. Put differently, having a collective phenomenal experience is not a necessary condition for psychological intersubjectivity. In psychological intersubjectivity, the other human being is always an object of our attention and an object in our awareness. We observe the other person for communicative cues and informational structures relevant to the ongoing interaction. Unlike in phenomenological intersubjectivity there is no requirement for an emphatic relationship with the other person, and indeed intersubjectivity can be antagonistic (Matusov, 1996). However, in the emergent technological case, there is a dynamic interplay between these psychological and phenomenological aspects. In technological intersubjectivity, information processing entailed by computational support can enhance and enrich the communicative possibilities and communion potentials of two or more human beings. Socio-technical systems and online communities have potentials for both psychological and phenomenological intersubjective experiences without the requirement that interacting persons be co-present in the same place and interact at the same time. With reference to cultural usability, even though prior empirical research has shown cross-cultural differences in traditional face-to-face intersubjectivity, the cultural variation in the structures and functions of technological intersubjectivity have received little empirical attention (see Vatrapu & Suthers, 2009b for a report of the TI aspects of the experimental study).

Research Questions

Four separate lines of empirical research have demonstrated that culture influences social behavior, (House et al., 2004), communication (Hall, 1977), cognitive processes (Nisbett & Norenzayan, 2002), and interacting with computers (Vatrapu & Suthers, 2007). These four lines of empirical research were integrated into a conceptual framework, and an experimental study was designed to empirically evaluate the framework. The primary purpose of the study was to answer two basic research questions. The first research question asked "To what extent does culture influence the appropriation of socio-technical affordances?" The second research question asked "To what extent does culture influence technological intersubjectivity?" As mentioned before, an experimental study was originally designed and conducted to answer these two primary research questions. Several theoretical predictions were generated from prior empirical evidence (Hall, 1977; House et al., 2004; Nisbett & Norenzayan, 2002) warranting the claim that both the perception and appropriation of affordances vary across cultures and that interpersonal perceptions and relations also vary across cultures (see Vatrapu, 2007, 2008). This paper presents a usability analysis of the empirical data generated by the experimental study. Specifically, this paper seeks to answer the following two research questions on cultural usability in computer-supported collaboration settings:

- To what extent does culture influence objective measures of usability in computer-supported collaboration?
- To what extent does culture influence subjective measures of usability in computer-supported collaboration?

Methodology

The formal definition of socio-technical affordance $\beta = p/q/r$ in $W = (T, S, O)$ has two important elements: technology T and individual actors S and O. The definition of appropriation of affordances has two important elements: affordances and intentions. Based on these two definitions, an experimental study was designed that introduced a variation in the cultural background of individuals (by selecting participants from a nation-state based ethnically stratified random sampling frame) but kept invariant the technological interface T and interactional setting. Briefly, the experimental study investigated how pairs of participants from similar and different cultures (American-American, American-Chinese, and Chinese-Chinese) appropriated affordances in a quasi-asynchronous computer-supported collaborative learning environment with external representations in order to collaboratively solve a public health science problem. Usability analysis reported in this paper was conducted on the empirical data generated by this experimental study.

In the next five subsections experimental design, materials, research hypotheses, sampling, and procedure are discussed briefly.

Experimental Design
The experimental study design consisted of three independent groups of dyads from similar and different cultures (American, Chinese) doing collaborative problem solving in a knowledge-mapping learning environment (described below). The three experimental conditions were the Chinese-Chinese intra-cultural condition, the American-American intra-cultural condition, and the Chinese-American inter-cultural condition.

In all three experimental conditions, the collaborative dyads were given the same experimental task. All the collaborative dyads interacted in the same computer-supported collaborative learning environment after reading the same instructions, software tutorial, and demonstration. The same instruments were administered to all participants. Internal validity and external validity were actively considered when designing and conducting the experiment. Construct validity was addressed by using existing instruments with high validity and reliability (Bhawuk & Brislin, 1992; House et al., 2004; Schwartz et al., 2001; Suinn, Ahuna, & Khoo, 1992). Brief descriptions of the experimental study's software and topics are provided in the following sections.

Materials
The following sections discuss the software, protocol for workspace updates, and alternatives for action.

Software

The computer-supported collaborative learning environment used in this experimental study has an Information Viewer on the left of the screen (see Figure 1) in which materials relevant to the problem are displayed. This Information Viewer functions as a simple web browser, but the presentation of materials is constrained as discussed in the next section. The environment has a shared workspace or Information Organizer section on the right side of the screen in which participants can share and organize information they gather from the problem materials as well as their own interpretations and other ideas. The Discussion tool below the Information Viewer on the left enables participants to discuss their ideas in a threaded discussion format.

The Information Organizer workspace includes tools derived from Belvedere (Suthers et al., 2008) for constructing knowledge objects under a simple typology relevant to the task of identifying the cause of a phenomenon (e.g., a disease), including data (green rectangles, for empirical information) and hypotheses (pink rectangles, for postulated causes or other ideas). There are also linking tools for constructing consistency (For) and inconsistency (Against) relations between other objects, visualized as green links labeled **+** and red links labeled **-** respectively. Unspecified objects and Unknown links are also provided for flexibility. Finally, an embedded note object supports a simple linear (unthreaded) discussion that appears similar to a chat tool, except that a note is interactionally asynchronous and one can embed multiple notes in the knowledge-map and link them like any other object. In the Threaded Discussion section of the environment (see bottom left of Figure 1) participants can embed references to knowledge-map objects in the threaded discussion messages by selecting the relevant one or more graph object while composing the message. The references show up as small icons in the message. When the reader selects the icon, the corresponding object in the knowledge-map is highlighted, indicating the intended referent. Figure 1 displays a captioned screenshot of the environment used in the experimental study.

Figure 1. Screenshot from I3P1's session

Mutual awareness of participants' artifacts is supported in the software environment as follows: all knowledge-map nodes and threaded discussion messages carry the name of the participant

who first created it. The mutual awareness features of artifacts and of activity are shown in Figure 1, a screenshot taken from I3P1 (I stands for American-Chinese inter- cultural session, 3 stands for the number of experimental session in this condition, P1 stands for Participant 1). In Figure 1, the I3P1's screen name of "Teri" (screen name selected by participant) appears on the title bar of the application window and on knowledge-map nodes and message created by I3P1. Similarly, I3P2's screen name of "Sue" appears on artifacts created by her. Artifacts marked with a solid red triangle in the top-right corner are from I3P2 and are yet to be opened by I3P1. The yellow circle on the threaded discussion message of I3P1 in the lower-left region of Figure 1 indicates artifacts created by "Teri" (I3P1) but not yet read by the study partner, "Sue" (I3P2). Thus each participant is potentially aware of the new artifacts from the study partner as well as the artifacts not yet read by their study partner.

Protocol for workspace updates

To simulate asynchronous online interactions that are typical of many computer-supported collaboration settings in learning and work, the actions of each participant in the shared workspace were not displayed immediately in the other participant's workspace. As a person worked, the actions of that person were sent to the other participant's client application, but were queued rather than displayed. Participants were given a new report after playing the game of Tetris™. Tetris™ was chosen as it presents a different sensory-motor perceptual task than the primary experimental task of collaborative knowledge-map co-construction. The game simulates taking a break from the task in real-world asynchronous settings (Suthers et al., 2008). After the game of Tetris™, all of the queued actions on that client were displayed. Conflicts that might arise when both participants edited the same object were resolved through operational transformations (Sun, Jia, Zhang, Yang, & Chen, 1998). The delayed updating protocol simulates one aspect of the experience of asynchronous collaboration: a participant sees what one's partner has done upon returning to a workspace after a period of time. It excludes the possibility of synchronous conversation in which one participant posts a message in the workspace and receives an immediate reply. The Refresh feature of the software enables one to get all updates to that point in time.

Alternates for action

The software environment provides multiple alternatives for appropriation of affordances and multiple ways to relate to the social other (the study partner). For example, participants can discuss with each other using the threaded discussion tool or the embedded notes tool. Participants can also use the knowledge-map objects to discuss the task at hand or any other topic of interest. Participants can refer to artifacts by deictic referencing (this, that, etc.) or use the cross-referencing feature of the threaded discussion. Participants can externalize the perceived relations between their concepts by creating external evidential relations between objects in the knowledge-map, by spatial arrangement, or by mentioning them in discussion. Participants have multiple ways of sharing the information presented to them (threaded discussion, embedded notes, and knowledge-map).

The research strategy was to provide participants with a feature rich collaborative environment with multiple alternates for action. By incorporating systematic variation in the assignment of participants to the collaborative dyad based on their cultural background and gender, the experimental design measured and observed systemic differences in how participants used the tools and resources of the technology (research question 1, appropriation of affordances) and related to each other during and after their interaction (research question 2, technological intersubjectivity). This in turn helped us analyze the empirical data from a cultural usability perspective.

Task

The study presented participants with a "science challenge" problem that requires participants to identify the cause of a disease known as Amyotrophic Lateral Sclerosis-Parkinsonism/Dementia (ALS-PD). This disease is prevalent on the island of Guam and has been under investigation for over 60 years, in part because it shares symptoms with Alzheimer's and Parkinson's diseases. Only recently have investigators converged on both a plausible disease agent (a neurotoxic amino acid in the seed of the Cycad tree) and the vector for introduction of that agent into people (native Guamians' consumption of fruit bats that eat the seed). Over the

years numerous diverse hypotheses have been proposed and an even greater diversity of evidence of varying types and quality explored. These facts along with the relative obscurity, multiple plausible hypotheses, contradicting information, ambiguous data, and high interpretation make this a good experimental study task for measuring cultural effects on appropriation of affordances and on technological intersubjectivity. Cognition, social behavior, and communication are all involved in this particular task and software. For example, instead of having a socially awkward and culturally inappropriate (for social harmony maintenance) verbal disagreement with the other-actor about a data node contradicting a hypotheses node, the self-actor could choose to create an evidential "Against" link between the two nodes. For other rationale, see the hypotheses listed in (Vatrapu, 2008).

All experimental study materials were in English. All participants began with a mission statement that provided the problem description and task information. Four mission statements corresponding to the four participant assignment configurations (Chinese vs. American x P1 vs. P2) were administered (http://lilt.ics.hawaii.edu/culturalreps/materials/). Due to the distribution of conflicting evidence, sharing of information across participants and study sessions was needed to expose the weakness of the genetics explanation as well as to construct the more complex explanation involving bats and cycad seeds. Given the nature of the information distribution between the two collaborating participants, working out that the consumption of bats as an optimal hypothesis as a cause for contracting ALS-PD involved making these cross-report collaborative connections and also considering and rejecting other probable factors. The study task and task materials were designed to highlight the social division of cognitive labor between the collaborating dyad. The experimental study encouraged participants to interact with each other by including the following reinforcing task instruction on each report (set of four articles): "Please share and discuss this information with you colleague..."

Participants

Participants were recruited from the graduate student community at the University of Hawaiʻi at Mānoa. Each participant was offered a payment of $75 (U.S.) for participating in the study. Participant selection and treatment assignment are discussed in the following sections.

Sampling

There is a tendency in cross-cultural, computer-mediated communication research to use cultural models bounded by modern nation-states. Nationality based stratified sampling frames remain a methodologically convenient way to select participants provided that cultural homogeneity of the participants is not to be assumed but empirically measured. We used the Portrait Value Questionnaire (PVQ) individual values survey (Schwartz et al., 2001) and the Global Leadership and Organizational Behavior Effectiveness (GLOBE) instrument (House et al., 2004) to empirically assess differences in the two participant groups at the individual and group levels respectively.

Selection

Participants were selected based on the sampling frame consisting of two randomly stratified subject pools of graduate students either from mainland USA or from the People's Republic of China and the Republic of China (Taiwan). That is, all the American participants in the study self-reported growing up in mainland USA and similarly, all the Chinese participants self-reported growing up in China or Taiwan. We used the stratified random sampling frame to control for the context effects of Hawaii. For detailed rationale, please see (Vatrapu, 2007, pp. 129-130).

Assignment

Participants were randomly assigned to either the intra- or the inter-cultural profiles and the same or different gender profiles. Excluding six pilot studies, a total of 33 experimental sessions involving 66 pairs of participants were conducted. Data from three experimental sessions was discarded due to issues of a missing screen recording, a software crash, and a disqualification. As a result, there were 10 pairs of participants for each of the three treatment groups: Chinese-Chinese intra-cultural, American-American intra-cultural, and American-Chinese inter-cultural. All the three conditions were gender-balanced because gender can substantially influence social

interaction (Tannen, 1996). Each treatment group included three female-female, three male-male, and four female-male dyads.

Instruments

The following sections discuss the demographic questionnaire, the self-perception PVQ, the GLOBE cultural dimensions instrument, the individual essays, the peer-perception PVQ, the acculturation SL-ASIA questionnaire, intercultural sensitivity, and the user satisfaction QUIS questionnaire.

Demographic questionnaire

A demographic questionnaire (Vatrapu, 2007, pp. 275-276) was administered to collect participants' familiarity with each other, with online learning environments, with usability evaluation studies as well as data about age, gender, ethnic background, duration of stay in the USA, and duration of stay in the state of Hawai'i. All participants were requested to make a self-report of their Cumulative Grade Point Average (CGPA) and also assign a release form for obtaining official records of their CGPA, Graduate Record Examination (GRE) scores, and the Test of English as a Foreign Language (TOEFL; Chinese participants only).

Self-perception: Portrait value questionnaire (PVQ)

The 40 item version of the PVQ instrument (Vatrapu, 2007, pp. 277-279) recommended for intercultural contexts (S.H. Schwartz, personal communication, 2006) was used in the study. The PVQ scale measured cultural values at the individual level. Cronbach's "alpha measures of internal consistency range from .37 (tradition) to .79 (hedonism) for the PVQ (median, .55)" (Schwartz et al., 2001, p.532). Gender specific versions of the self perception PVQ scale were administered.

GLOBE cultural dimensions instrument

The GLOBE instrument (House et al., 2004) was used to measure cultural values at the group level (Vatrapu, 2007, pp. 280-293). Section 1 ("The way things are in your society") and Section 3 ("The way things generally should be in your society") of the original GLOBE instrument were used in this study. Section 1 of the GLOBE instrument measures a responder's perceptions of their society. Section 3 of the GLOBE instrument measures a responder's preferences for their society. According to the *Guidelines for the Use of GLOBE Culture and Leadership Scales[2]* (2004) "the construct validity of the culture scales was confirmed by examining the correlations between the GLOBE scales with independent sources (e.g., Hofstede's culture dimensions, Schwartz's value scales, World Values Survey, and unobtrusive measures)" (p.5). Phrasing of "this country" has been changed to "my home society" to remove possible ambiguity for Chinese graduate students who might rate Hawai'i, USA instead of the society they grew up in.

Individual essays

At the end of the collaborative science problem-solving session, each participant wrote an essay. This constituted the immediate post-test. Identical essay writing instructions were provided to all participants. The instructions asked the participants to (a) state the hypotheses they considered, (b) whether and how their hypotheses differed from those of their study partners, and (c) their final conclusion.

Peer-perception: PVQ

Technological intersubjectivity after interaction was measured by the second immediate post-investigative test. This was the administration of the PVQ (Schwartz et al., 2001) instrument with a reversal of the direction of assessment (Vatrapu, 2007, pp. 304-306). This time instead of assessing themselves, participants assessed their collaborative partners. Based on their

---

[2]http://www.thunderbird.edu/wwwfiles/sites/globe/pdf/GLOBE_Culture_and_Leadership_Scales_Guidelines.pdf

collaborative interactions, each participant rated his/her impressions of the study partner on the PVQ.

Acculturation: Suinn-Lew Asian Self Identity Acculturation (SL-ASIA) questionnaire

Acculturation is a process that occurs when members of two or more cultures interact together. This becomes an external variable in cross-cultural research conducted with participants from an immigrant culture in a host culture (in our case, Chinese participants in Hawaiʻi, USA). This external variable can be controlled by measuring the acculturation level of the participants belonging to the minority immigrant culture (Triandis, Kashima, Shimada, & Villareal, 1986). Participants with a high level of acculturation can be best used as members of the majority host culture or not included in the study (Triandis et al., 1986). This research project used the SL-ASIA scale (Suinn et al., 1992) to measure the acculturation levels of the Chinese participants (Vatrapu, 2007, pp. 307-311). This scale was chosen as it is specifically designed for Asians. Suinn et al. (Suinn et al., 1992) reported an internal-consistency estimate of .91 for the SL-ASIA instrument.

Intercultural sensitivity: Intercultural sensitivity instrument

Intercultural sensitivity is a vital skill for intercultural collaborations (Bhawuk & Brislin, 1992). The SL-ASIA scale provided a measure of Chinese participants assimilation to the USA. The intercultural sensitivity instrument (ICSI(Bhawuk & Brislin, 1992) was used to measure American participants self-assessment of intercultural sensitivity (Vatrapu, 2007, pp. 312-315). Bhawuk and Brislin (1992) report that "the ICSI was validated in conjunction with intercultural experts at the East-West Center with an international sample (n=93)" (p. 423). The word "Japan" in the original ICSI scale was changed to "China" to fit the context of Chinese-American collaboration setting of the experiment. Part three of the original ICSI instrument was not used, as pilot studies indicated that it was irrelevant to the purposes of this experimental study.

User Satisfaction: Questionnaire for User Interaction Satisfaction (QUIS)

The QUIS 7.0 questionnaire (Harper, Slaughter, & Norman, 2006) was administered to collect the participants subjective perceptions and preferences of the learning. The QUIS has high reliability, Cronbach's alpha = 0.95 and high construct validity (alpha = 0.86), (Harper et al., 2006).

Procedure

Two students participated in each session. Experimental sessions lasted about 3.5 hours on average. Informed consent was obtained from all participants for both the pilot studies and the experimental studies. After signing the informed consent forms, participants completed a demographic survey. They were then given a CGPA/GRE/TOEFL score release form, a self-perception PVQ (Schwartz et al., 2001), and the GLOBE instrument (House et al., 2004). After completing these three forms, participants were brought into a common room. Participants were then introduced to the software and the structure of the experimental study through an identical set of instructions and demonstrations across all three conditions.

After the software demonstration, the two participants were led back to their respective workstations in two different rooms. They were then instructed to begin work on the study task. Participants had up to 90 minutes to work on the information available for this problem. The update protocol described in (Suthers et al., 2008) was used to synchronize the workspaces of the two participants. At the conclusion of the investigative session, each participant was given up to 30 minutes to write an individual essay. The CSIC environment remained available to each participant during the essay writing, but the participants were requested not to engage in any further communication. After each participant had finished writing the individual essay, the other-perception PVQ instrument (Schwartz et al., 2001) and the QUIS instrument (Harper et al., 2006) were administered. This concluded the experimental session. Participants then completed the payment forms and were debriefed.

Results

Results are grouped under the following five subsections: Demographics, Culture Measures, Objective Usability Measures, and Subjective Usability Measures. The empirical data generated by the experimental study were analyzed at four levels: culture (American, Chinese), gender (female, male), dyadic culture (American-American, American-Chinese, Chinese-Chinese), and dyadic gender (female-female, female-male, male-female).

Demographics

The age of the participants (n=60) ranged from a minimum of 22.00 years to a maximum of 45.00 years. The average age of the participants was 28.20 years (SD = 4.6, SE = 0.60). There was no age difference at any of the four levels of analysis (culture, gender, dyadic culture, dyadic gender). As expected, American participants reported to have spent significantly more time in the United States of America than the Chinese participants. On the other hand, the time spent by the participants in Hawaiʻi with respect to culture and gender was not statistically significant. Of the 60 participants, 30 reported being doctoral students and the other 30 participants reported being masters students. Of the 30 Chinese participants 14 were doctoral students and 16 were masters students. Of the 30 American participants, 16 were doctoral students and 14 were masters students. Participants belonged to 30 different departments at the University of Hawaiʻi at Mānoa. There were no significant differences at any of the four levels of analysis for prior experience with experimental studies, prior knowledge about the experimental task, and partner familiarity.

Culture Measures

As mentioned before, a PVQ (Schwartz et al., 2001) was used to measure culture at the individual level. The GLOBE instrument (House et al., 2004) was used to measure culture at the group level.

Ten individual values were measured by the PVQ (Schwartz et al., 2001). Statistical analysis showed that at the level of culture the following PVQ values were significant: Conformity, $F(1,56)=7.71$, $p=0.008$; Benevolence, $F(1,56)=5.60$, $p=0.02$; Universalism, $F(1,56)=6.66$, $p=0.01$; Self-Direction, $F(1,56)=7.48$, $p=0.01$; Stimulation, $F(1,56)=10.02$, $p=0.003$; and Security, $F(1,56)= 30.76$, $p<0.0001$.

Significant differences were observed on both sections of the GLOBE instrument. For the "AS IS" section, significant differences between the American and Chinese groups were observed for the following: Institutional Collectivism, $F(1,56)=43.55$, $p<0.01$; In-Group Collectivism, $F(1,56)=102.43$, $p<0.01$; and Assertiveness, $F(1,56)=28.57$, $p<0.01$. For the "SHOULD BE" section of the GLOBE instrument, statistically significant differences were found for the following: Uncertainty Avoidance, $F(1,56)=49.65$, $p<0.01$; Assertiveness, $F(1,56)=4.20$, $p=0.04$; Future Orientation, $F(1,56)=14. 23$, $p=0.01$; Humane Orientation, $F(1,56)= 7.90$, $p=0.007$; and Gender Egalitarianism, $F(1,56)=4.89$, $p=0.03$.

In summary, there is necessary and sufficient evidence to conclude that Chinese and American participants significantly differ on specific PVQ individual values as well as GLOBE cultural dimensions. Even though a nation-state based stratified random sampling frame was utilized, systemic variation between the two participant groups was empirically documented and not stereotypically assumed or dogmatically asserted.

Objective Usability Measures

Objective usability measures consisted of the efficiency (total task time in minutes) and effectiveness (usage of certain features of interest). A multivariate analysis of variance with the independent variable of dyadic culture (American-American, American-Chinese, and Chinese-Chinese) and the dependent variables of task time, cross-referencing, shared workspace refresh, threaded discussion messages, embedded discussion notes, evidential relation links (for+against+unknown), data nodes, hypotheses nodes, and unspecified nodes was statistically significant, Roy's Largest Root=0.767, $F(9, 50)=4.259$, $p<0.001$. Each of these dependent variables is discussed below.

With respect to the independent variable of culture (American, Chinese), a multivariate analysis of variance with the same set of dependent variables mentioned above yielded significant results, Roy's Largest Root=0.586, $F(9, 50)=3.758$, $p=0.003$. Each of the dependent variables is discussed below.

Efficiency

On average, task time was greater for Chinese participants (*M*=156.07 minutes, *SD*=19.22) than the American participants (*M*=144.96, *SD*=25.14). On average, female participants' task time (*M*=155.58, *SD*=20.88) was greater than the male participants (*M*=145.44, *SD*=24.00) in the study. A two-way ANOVA showed marginal main effects for culture, $F(1,56)= 3.77$, $p=0.06$ and gender, $F(1,56)=3.14$, $p=0.08$. On the other hand, total task time varied significantly between the intra- and inter-cultural conditions of the experimental study, $F(2,51)=5.17$, $p=0.009$. A Bonferroni post-hoc comparison showed that the American intra-cultural group had significantly lower task time than the Chinese intra-cultural group and the American-Chinese inter-cultural group. No significant differences were observed at the dyadic gender level.

Effectiveness

Effectiveness measures counted the number of uses of the software features of structural and functional significance to computer-supported collaborative learning. Each measure is introduced, briefly discussed, and then empirical results are presented in the following sections.

Shared workspace refresh

As discussed in the Software section in the Methodology section of this paper, the shared workspace (information organizer + discussion) could be refreshed (a) automatically after returning from the game or (b) on demand when the participant clicked on the Refresh button (see top-right in Figure 1). There were four reports and a final page. Participants had to play and quit the game in order to receive the next report. All the participants played and quit the game at least four times and therefore, received all four reports. However, the refresh count varied due to the differences in the number of on-demand refreshes of the shared workspace. There was no significant main effect for refresh count with respect to culture or dyadic culture. Figure 2 presents the effectiveness results (minus the cross-referencing usage measure) at the culture level of analysis. Figure 3 presents the effectiveness results at the dyadic culture level of analysis.
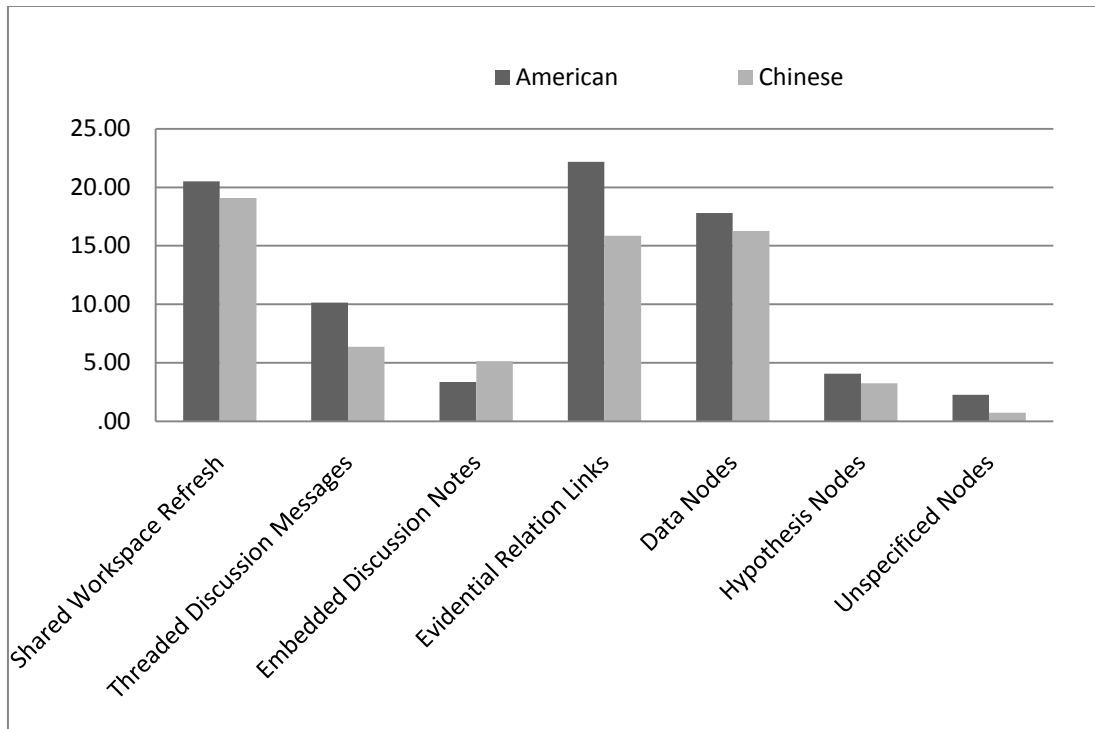
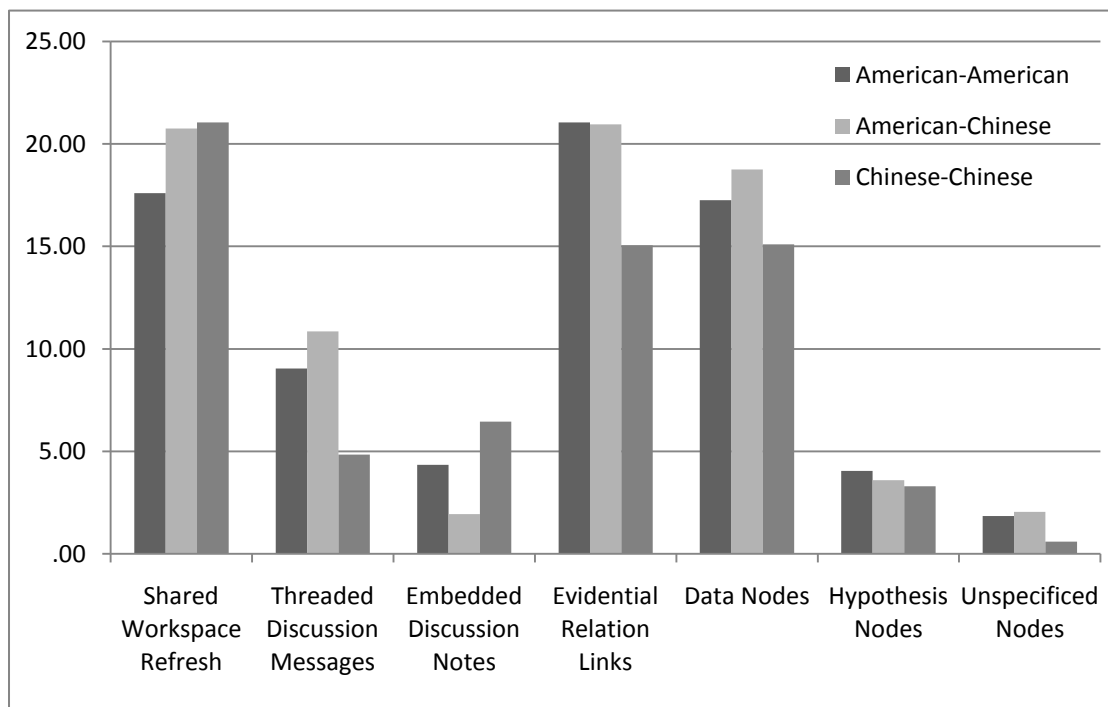**Figure 2.** Effectiveness with respect to culture

Figure 3. Effectiveness with respect to dyadic culture

Cross-referencing

Video analysis of the screen recordings of participant sessions was done to obtain the counts for cross-referencing (graphical objects embedded in the threaded discussion messages, see Figure 1). Even though the empirical trend was that on average, American participants used cross-referencing more than the Chinese participants, no statistically significant differences were found at any of the four levels of analysis (culture, gender, dyadic culture, dyadic gender).

Threaded discussion messages

Counts for *discourse usage* were obtained from the software logs of participant sessions. For threaded discussion messages, American participants created more threaded discussion messages than the Chinese participants, and the difference was statistically significant, $F(1,56)=8.88$, $p=0.004$. Significant differences were also observed at the dyadic culture level of analysis, $F(2,57)=8.84$, $p<0.001$. Participants in the Chinese-American inter-cultural condition created the highest number of threaded discussion messages followed by the American intra-cultural group and the Chinese intra-cultural group.

Embedded discussion notes

For the embedded discussion notes, no statistically significant differences were found. However, the observed empirical trend was that Chinese participants created more embedded discussion notes than the American participants. With respect to the dyadic culture level of analysis, significant differences were observed between the three experimental conditions, $F(2,57)=4.76$, $p=0.012$. Post-hoc comparisons showed that participants in the American intra-cultural group created more embedded discussion notes than those in the Chinese intra-cultural group while participants in the American-Chinese inter-cultural group created the lowest number of embedded discussion notes.

Evidential relation links

Counts for *evidential relation links* were obtained from the software logs of participant sessions. The total number of evidential links created was the sum of the Against links, the For links, and the Unknown links created. On average, American participants created significantly more evidential relation links compared to the Chinese participants of the experimental study, $F(1,56)=5.54$, $p=0.02$. However, no significant differences were observed at the dyadic culture level of analysis.

Knowledge-map nodes

No significant differences were observed between the Chinese and American participants in the number of data and hypotheses nodes created. However, Chinese participants created fewer unspecified nodes than the American participants, $F(1, 56)=5.76$, $p=0.02$. At the gender level of analysis, female participants created significantly more hypothesis nodes than the male participants, $F(1, 56)=4.68$, $p=0.035$. No significant differences were observed at the dyadic gender level of analysis.

Subjective Usability Measures

As mentioned earlier, a validated usability instrument, the QUIS questionnaire (Chin, Diehl, & Norman, 1988), was administered to collect the participants' subjective perceptions and preferences of the learning environment. In addition to various system measures, the QUIS 7.0 instrument also measured participants' subjective satisfaction with the instructions and the software tutorial. The coding key for the QUIS instrument was used for the quantitative analysis of the data (http://lap.umd.edu/QUIS/QuantQUIS.htm).

Overall system user satisfaction

On average, the overall user satisfaction for the Chinese participants was higher than that for the Americans (see Figure 4). However, no significant differences were found at any of the four levels of analysis (culture, gender, dyadic culture, dyadic gender). A consistent empirical trend at the dyadic level of analysis was that the mean subjective rating for the inter-cultural group was situated between the mean ratings for the two intra-cultural groups (Figure 5).

Information display

Significant differences were observed between the Chinese and American participants on the QUIS section for information display on the screen, $F(1,56)=8.00$, $p=0.01$. Chinese participants' subjective satisfaction scores for the screen information display were lower than the American participants.

System terminology

System Terminology and System Information sections of the QUIS instrument received significantly lower ratings from the Chinese participants, $F(1,56)=4.84$, $p=0.03$. Significant differences were observed at the dyadic culture level of analysis, $F(2,57)=3.30$, $p=0.04$. Chinese intra-cultural participants had the least subjective satisfaction with the system terminology followed the inter-cultural group and the American intra-cultural group.

System capabilities

No significant differences were observed despite lower scores by Chinese participants compared to the American participants of the experimental study. No significant differences were observed at the dyadic level of analysis.

Ease of learning of the system

Results for the Learning section of the QUIS instrument showed a marginally significant difference on the ease of learning measure at the level of culture. No significant differences were observed at the dyadic level of analysis.

Software demonstration and tutorial

Results for the Tutorial section of the QUIS instrument showed no significant difference for participants' subjective evaluation of the software demo and experimental instructions at any of the four levels of analysis. Therefore, experimenter bias and "demand characteristics" (Orne, 1962) are ruled out as confounding variables in the study.

Another interesting empirical trend was that the average subjective ratings in the inter-cultural condition were always between those of the subjective ratings for the two intra-cultural conditions.
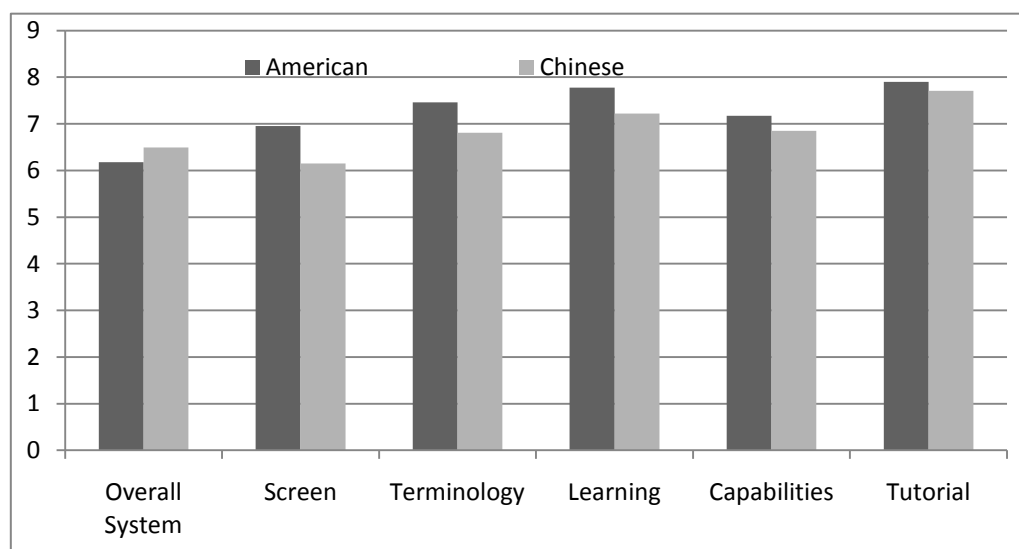

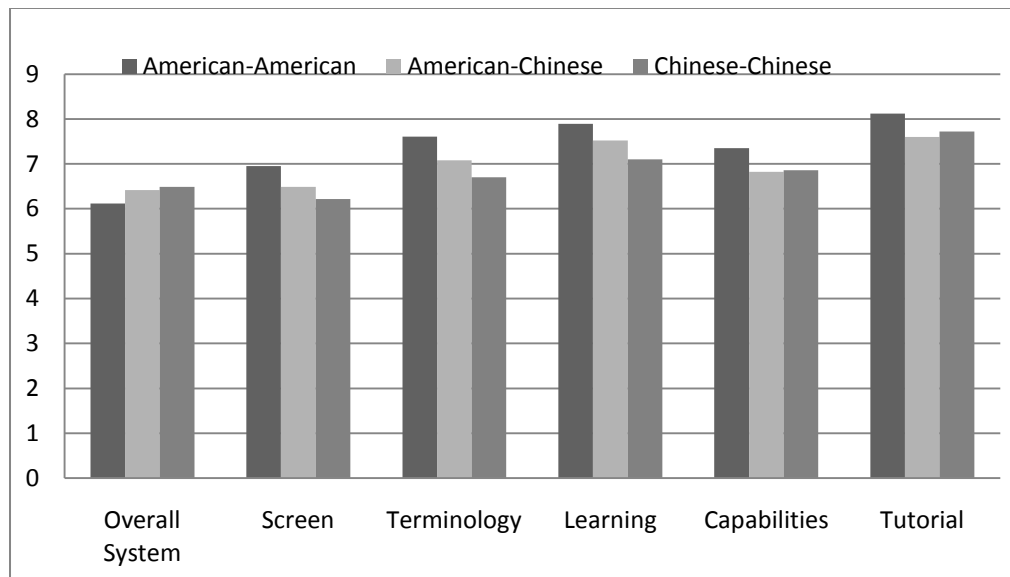**Figure 4.** QUIS ratings with respect to culture

Figure 5. QUIS ratings with respect to dyadic culture

In summary, there was a discrepancy between Chinese participants higher overall satisfaction ratings and their lower satisfaction ratings for the specific components of the system (Screen and Terminology & System Information). Similarly, American participants reported lower overall reaction ratings but higher satisfaction ratings to the specific components of the system. Figure 4 presents a summary of the QUIS results with respect to culture. Figure 5 presents the summary of the QUIS results with respect to the dyadic culture. It must be emphasized that while the mean difference in the overall ratings between the three collaboration conditions (Chinese-Chinese, American-Chinese, and American-American) and the two cultural groups (Chinese, American) is small and statistically significant. However, a qualitative analysis of the user comments on the QUIS questionnaire tells a different story as discussed in the following section.

Analysis of comments

The QUIS instrument includes an open-ended comments solicitation at the end of each of the six sections. The user comments were transcribed. A few illustrative user comments are included below:

C2P2[3] (Chinese, Female): "The link function is very helpful and I'll expect a drag and drop from the organizer to the message panel."

C6P2 (Chinese, Male): "Having a zoom in/out feature might help."

I3P1 (Chinese, Female): "I like the screen. But the text boxes are somehow difficult to organize, if there are too many of them."

I3P2 (American, Female): "I like the idea and that can link data boxes. However, a function that would put everything into a condensed list or a short outline to see all data at once should be helpful b/c sometimes too much information is displayed at once to work with in a rational manner."

---

[3] After both the participants for an experimental session have read and signed the informed consent form, each participant was assigned a unique participant ID of the form NxPy, where N refers to the treatment condition—one of the three cultural profiles, N=C for Chinese-Chinese condition, A for American-American condition, I for American-Chinese condition; x refers to the experimental session number within that condition (1-12); and Py refers to information distribution assignment (P1 or P2).

I8P1 (American, Female): "I found that performing an operation did not always lead to a predictable result. Sometimes, I was unable to move the text over or the wrong copied text appeared in the box."

I8P2 (Chinese, Male): "In general the system speed is satisfactory, and it's reliable, but it needs more on other functions such as undo, correcting typo, etc."

I11P2 (Chinese, Male): "It is good if there is a 'undo' and 'redo' (ctrl+z or ctrl+y)."

I12P2 (Chinese, Female): "I think it would be better to separate the links sent by partner and mine. It seems all the links are put together and inconvenient [*sic*] to read."

A3P1 (American, Female): "When new text boxes appear from the partner, they should be in a separate section so it is easy to see them and sort them out from mine and older ones. They should be color coded differently until read. The size of screens should be adjustable to allow more."

A7P2 (American, Male): "Instructions were well laid out and easy to use. Messages sometimes appear overlapping, difficult to see everything that way."

A8P1 (American, Female): "Overall fairly clear & easy to navigate."

A9P2 (American, Male): "In the boxes, the word 'text' should be eliminated in a click. It shouldn't need deleting."

Qualitative analysis of the comments shows that undo, copy + paste, zooming, and color coding of contributions were the most frequent usability suggestions. Usability problems mentioned included scrolling issues, font size, default text, etc. Negative comments were mainly about screen clutter.

The coding scheme developed in (Vatrapu & Pérez-Quiñones, 2006) was adapted for the content analysis of the comments. The modified coding scheme is described below:

- Usability problem (U): Interaction design flaw or a user difficulty that is directly associated with an interface/interaction design flaw.
- Suggestion (S): subjective preference of the participant to the implemented design choice or tradeoff.
- Positive comment (P): participant's subjective approval of a design choice or tradeoff.
- Negative comment (N): participant's subjective disapproval of a design choice or tradeoff.
- Other comment (O): User comment that couldn't be categorized under one of the above categories.

Total comments = usability problems (U) + suggestions (S) + negative comments (N) + positive comments (P)

Figure 6 presents the results for the aggregate count of user comments with respect to culture. Figure 7 presents the aggregate count of user comments with respect to the dyadic culture.

Even though the Chinese participants made more usability suggestions, more positive comments, and less negative comments than the American participants, no significant differences were observed at any of the four levels of analysis. Another interesting but statistically insignificant empirical trend was that more usability suggestions, negative comments, and positive comments were made in the inter-cultural condition compared to the two intra-cultural conditions.
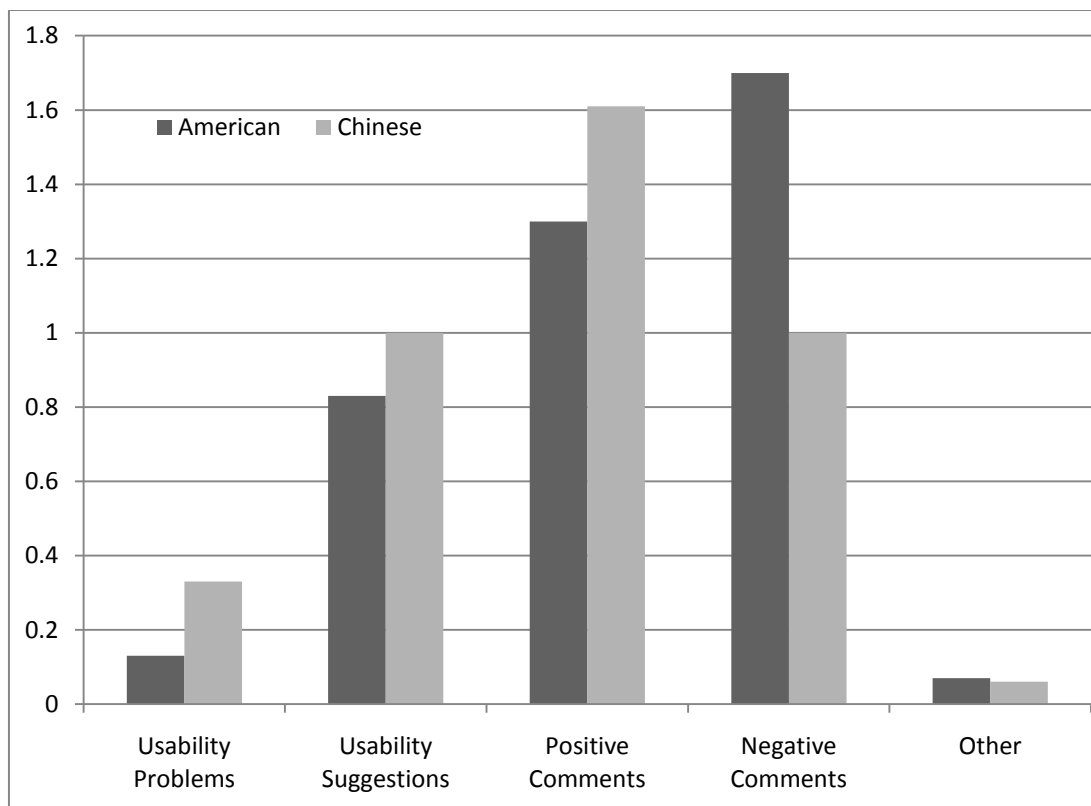
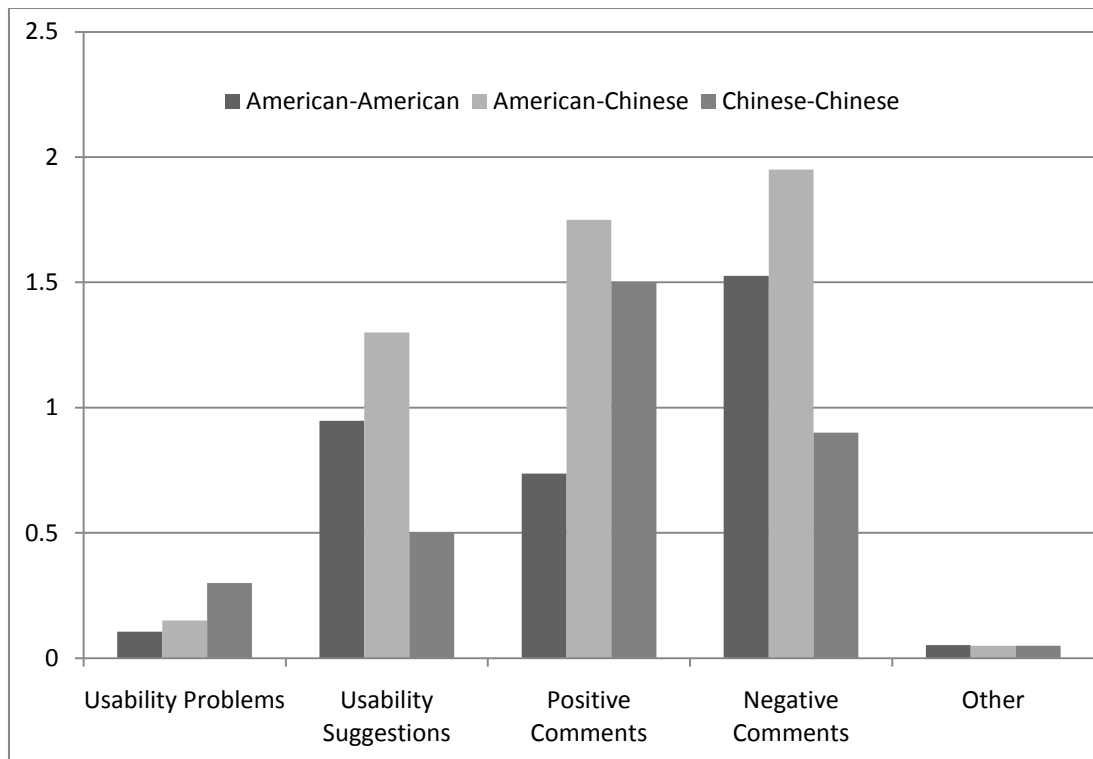**Figure 6.** Aggregate count of user comments with respect to culture

Figure 7. Aggregate count of user comments with respect to dyadic culture

Discussion

The empirical findings of the cultural usability analysis can be summarized as follows: Culturally different participants were found to be engaged at different levels in different collaborative activities such as shared workspace refresh, cross-referencing of knowledge-map objects in threaded discussion. The participants also created different quantities of artifacts such as evidential relation links and knowledge-map nodes. Further, participants from the two cultural groups gave different user interface satisfaction ratings besides different quality and quantity of comments. The results of our cultural usability analysis add to the empirical findings that document cultural effects on usability evaluation processes and products discussed in the Introduction section of this paper.

Specifically, Chinese participants, on average, reported higher overall user interface satisfaction scores but gave significantly lower ratings for information display and terminology aspects of the system (see Figure 4). This finding doesn't completely agree with the cultural-cognitive difference of holistic vs. analytical reasoning (Nisbett & Norenzayan, 2002; Nisbett, Peng, Choi, & Norenzayan, 2001). Following (Nisbett & Norenzayan, 2002; Nisbett et al., 2001)., participants from holistic thinking style cultures (such as Chinese participants of this study) are expected not to offer more specific user interface satisfaction ratings for the individual components of the system than the overall system. In other words, following (Nisbett & Norenzayan, 2002; Nisbett et al., 2001), in usability evaluation settings, it is to be expected that the holistic thinking style participants would be less discriminatory in their ratings of the constituent parts of the system compared to the overall system. But the empirical trend we found doesn't agree with these expectations. Chinese participants rated the components of the socio-technical system with greater specificity than the overall system. On the other hand, American participants in the study gave lower overall system ratings and higher constituent parts ratings. More research is needed on this empirical finding of discrepancies in ratings for whole vs. parts of the system between culturally different participants.

A possible explanation for the significant differences on participant ratings for the Screen section of the QUIS questionnaire could be the cultural-cognitive differences in field-dependent vs. field-independent modes of perception between the Chinese and American participants respectively (Nisbett & Miyamoto, 2005; Nisbett & Norenzayan, 2002; Witkin, 1967; Witkin & Goodenough, 1977). Briefly, field-dependent participants rely more on external cues than field-independent participants with respect to perceptual organization.

Even though not statistically significant, an interesting empirical trend was observed with the Chinese participants making more positive comments and less negative comments than the American participants of the experimental study. Chinese participants' cultural concerns about deference and harmony-maintenance might account for the higher number of positive comments than the negative comments. However, Chinese participants gave lower ratings to the parts of the system compared to the American participants. Moreover, Chinese participants found more usability problems and made more usability suggestions. From this empirical trend it appears that for measuring user satisfaction, it might be beneficial to solicit open-ended comments in addition to seeking Likert-type ratings on a questionnaire.

In prior work, we have proposed a design evaluation framework of usability, sociability, and learnability for computer-supported collaborative learning (CSCL) environments (Vatrapu, Suthers, & Medina, 2008). Despite the differences in usability reported in this paper and sociability reported elsewhere (Vatrapu, 2008; Vatrapu & Suthers, 2009b), a preliminary analysis of individual learning outcomes on the essays show no significant differences (Vatrapu, 2008; Vatrapu & Suthers, 2009a). So, even though the subjective perception of the ease of tool-learning was marginally significant, domain-learning outcomes were not statistically different. This corroborates prior findings in HCI and human factors that show a discrepancy between objective measures of performance and subjective indicators of preference or what we have termed as the *performance preference paradox* (Vatrapu et al., 2008).

Research Limitations: Linguistic Relativity and Cultural Relativity
English language comprehension and articulation remains a potential mediating variable in the performance of the Chinese participants. To empirically evaluate this mediating variable, session verbosity (total words individually produced by a participant in the collaborative session) and essay verbosity (total words produced by a participant in the individually written essay) were calculated. A two way analysis of variance for session verbosity showed significant main effects for both culture, $F(1,56)=4.46$, $p=0.04$, and gender, $F(1,56)=6.70$, $p=0.01$. On average, American participants produced more words in the collaborative session than the Chinese participants. Female participants produced more words in the collaborative session than the male participants. Similar results were obtained for the essay verbosity measure. It should be noted that copy + pasted information from the source materials is included in both verbosity measures. Also, given the nature of the experimental task, there were no strong theoretical reasons to control for copy + pasted words in the verbosity. Future studies using Chinese language materials and instruments should help disambiguate the language effect. Having said that, the experimental results are relevant to computer-supported collaborative settings with English as the medium of interaction.

Implications for Cultural Usability Research
Honold (2000, p. 341) identified eight factors to be taken into account for investigating product usage across cultural contexts: objectives of the users, characteristics of the users, environment, infrastructure, division of labor, organization of work, mental modes based on previous experience, and tools. Objectives of users have been the focus of recent work in cultural usability that focuses on evaluator-participant interpersonal relations (Shi & Clemmensen, 2007; Vatrapu & Pérez-Quiñones, 2006). Cultural usability researchers should carefully re-consider the debate around "demand characteristics" (Orne, 1962) in experimental psychology. Orne (1962) defined demand characteristics as "totality of cues that convey an experimental hypothesis to subjects [and] become significant determinants of subject's behavior" (p.779). In the context of cultural usability in particular, and usability evaluation in general, demand characteristics refer to the study expectations, evaluator-participant relationships, and cultural norms for appraisal and attribution. The experimental study reported in this paper was designed taking into account the Contact Hypothesis (Allport, 1954) that beneficial intergroup impressions and outcomes would result if equal status, equal institutional support, equal incentives, and a super-ordinate goal are provided for group interactions.

Regarding characteristics of user, the concept of culture is often used but rarely operationalized in experimental studies. Cultural characteristics are attributed to individuals by virtue of ethnic affiliation or nation-state membership. Given the "fading quality of culture" (Ross, 2004) the homogenous assumptions of culture are unwarranted. We conducted a manipulation check to verify that cultural characteristics of participants were in fact different. As stated in the culture measures in the Results section, cultural characteristics of participants in this experimental study are assessed, evaluated, and documented at level of individual values as well at the level of group cultural dimensions.

In our opinion, cultural usability research needs to move beyond documenting cultural variation in usability assessment methods and outcomes. We need to move toward an empirically informed theoretical understanding, explanation, and prediction of cultural HCI phenomena. Computer-supported collaboration presents some unique challenges that do not entirely fall under the purview of usability typically conceived of as efficiency, effectiveness, and satisfaction. Research in the established HCI research field of computer-supported cooperative work (CSCW), the emerging HCI field of human-information interaction (HII(Jones et al., 2006), the technology enhanced learning paradigm of computer-supported collaborative learning (CSCL; (Stahl, Koschmann, & Suthers, 2006), and the emerging field of computer-supported intercultural collaboration (CSIC; (Vatrapu & Suthers, 2009b) is investigating how users interact with each other as well as with the computers. Currently, the Internet is undergoing a profound shift towards a participatory mode of interaction. With the advent of fundamentally social software such as social networking sites (Orkut, Facebook, MySpace, Mixi etc.), cultural usability needs to be expanded to include technological intersubjectivity. We need a richer understanding of the phenomenon at the human-computer interactional level. In our opinion, there is a real need for a real-time and real-space interactional account of cultural cognition for the field of HCI. A first attempt at a socio-technical interactional theory of culture could be found in (Vatrapu, 2009a). We believe that an interactional understanding of cultural variance at the level of perception and appropriation of affordances and structures and functions of technological intersubjectivity can better inform the design, development, and evaluation of computer-supported collaboration systems than the current homeostatic conceptions of cultures as dimensional models or typologies. This is not to say that cultural dimensional models are not useful nor is it to say that they are not insightful. Rather, it is to say that cultural HCI phenomena should also be theorized and empirically evaluated at the human-computer interactional level.

Practitioner's Take Away

The following are the key points of this paper:

- Demand characteristics, in the context of cultural usability, refer to the study of expectations, evaluator-participant relationships, and cultural norms for appraisal and attribution. Demand characteristics should be given careful consideration in the design of cultural usability evaluations and in the interpretation of the results.
- For usability evaluations involving participants from social collectivist cultures with interactional concerns for deference and harmony-maintenance, it might be beneficial to use questionnaires that solicit both Likert-type ratings and open-ended comments.
- Participants from highly collectivistic cultures (such as the Chinese participants of this experimental study) with a high-context communication style and a field-dependent perceptual style might offer higher overall ratings to a system despite offering lower subjective ratings for the constituent parts of the systems.
- Participants from highly individualistic cultures (such as the American participants of this experimental study) with a low-context communication style and a field-independent perceptual style might offer lower overall ratings to a system despite offering higher subjective ratings for the constituent parts of the systems.
- Participants from highly collectivistic cultures with a greater emphasis on deference and harmony-maintenance (the Chinese participants in the context of this study) might make a higher number of positive comments compared to negative

comments during usability evaluation despite higher negative ratings on Likert-type questionnaires.

- Participants from highly individualistic cultures with a lesser emphasis on deference and harmony-maintenance (the American participants in the context of this study) might make fewer positive comments and more negative comments during usability evaluation despite higher positive ratings on Likert-type questionnaires.

- Efficacy of open-ended comments might be higher for inter-cultural user groups when compared to intra-cultural user groups.

- Despite significant differences in objective measures of usability, there might be no significant differences in performance or achievement of the primary task. In other words, there is a possibility of a "performance preference paradox" in the usability evaluation of computer-supported collaboration systems.

- It might be more productive and informative to conceive of cultural variation at the level of human-computer interaction as a variation in the perception and appropriation of action-taking possibilities and meaning-making opportunities relate to actor competencies and system capabilities and the differences in social relationships and expectations rather than homeostatic cultural dimensional models or stereotypical typologies.

- Design the technological capabilities of the system taking into account the competencies of the users. Design should focus on what action-taking possibilities and meaning-making opportunities users can psychologically perceive and sociologically appropriate. Design for sustaining, transforming, and creating traditional and novel configurations of social relationships and interactions in socio-technical systems.

Acknowledgements

References

Allport, G. (1954). *The nature of prejudice*: Addison-Wesley.

Beu, A., Honold, P., & Yuan, X. (2000). How to Build Up an Infrastructure for Intercultural Usability Engineering. *The International Journal of Human-Computer Interaction, 12*(3&4), 347-358.

Bhawuk, D. P. S., & Brislin, R. W. (1992). The measurement of intercultural sensitivity using the concepts of individualism and collectivism. *International Journal of Intercultural Relations, 16*(4), 413-436.

Bradner, E. (2001). Social Affordances of Computer-Mediated Communication Technology: Understanding Adoption. *Extended Abstracts of the Conference on Human Factors in Computing Systems. CHI, Seattle, Washington.*

Chin, J. P., Diehl, V. A., & Norman, K. (1988). Development of an instrument measuring user satisfaction of the human-computer interface *SIGCHI '88* (pp. 213-218). New York: ACM.

Clarke, A. C. (1962). *Profiles of the future: an inquiry into the limits of the possible*: Harper & Row.

Clemmensen, T., Hertzum, M., Hornbæk, K., Shi, Q., & Yammiyavar, P. (2009). Cultural cognition in usability evaluation. *Interacting with Computers, 21*(3), 212-220.

Crossley, N. (1996). *Intersubjectivity: The Fabric of Social Becoming*. London: Sage.

Day, D., & Evers, V. (1999). *Questionnaire Development for Multicultural Data Collection*. Paper presented at the International Workshop on Internationalisation of Products and Systems (IWIPS-99).

DeSanctis, G., & Poole, M. S. (1994). Capturing the Complexity in Advanced Technology Use: Adaptive Structuration Theory. *Organization Science, 5*(2), 121-147.

Evers, V. (1998). *Cross-cultural Understanding of Metaphors in Interface Design.* Paper presented at the Attitudes toward Technology and Communication, London.

Fernandes, T. (1995). *Global Interface Design*: Academic Press.

Frandsen-Thorlacius, O., Hornbæk, K., Hertzum, M., & Clemmensen, T. (2009). *Non-universal usability?: a survey of how usability is understood by Chinese and Danish users.* Paper presented at the 27th international conference on Human factors in computing systems, Boston, MA, USA.

Garfinkel, H. (1967). *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.

Gaver, W. (1991). Technology Affordances. *Conference on Human Factors in Computing Systems (CHI'91). New Orleans, USA.*, 79-84.

Giddens, A. (1986). *The Constitution of Society: Outline of Theory of Structuration*: University of California Press.

Hall, E. (1977). *Beyond Culture*. New York: Anchor Books.

Harper, B., Slaughter, L., & Norman, K. (2006). Questionnaire administration via the WWW: A validation & reliability study for a user satisfaction questionnaire., www.lap.umd.edu/webnet/paper.html. Last retrived on 31-July-2010.

Herman, L. (1996). *Towards Effective Usability Evaluation in Asia: Cross-Cultural Differences.* Paper presented at the 6th Australian Conference on Computer-Human Interaction (OZCHI'96), Hamilton, New Zealand.

Hertzum, M., Clemmensen, T., Hornbæk, K., Kumar, J., Shi, Q., & Yammiyavar, P. (2007). Usability Constructs: A Cross-Cultural Study of How Users and Developers Experience Their Use of Information Systems. *Lecture Notes in Computer Science, 4559*, 317.

Honold, P. (2000). Culture and Context: An Empirical Study for the Development of a Framework for the Elicitation of Cultural Influence in Product Usage. *International Journal of Human-Computer Interaction, 12*(3&4), 327 - 345.

House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., & Gupta, V. (2004). *Culture, Leadership and Organizations: The GLOBE study of 62 societies*. Newbury Park, CA: Sage Publications.

Iwarsson, S., & Ståhl, A. (2003). Accessibility, usability and universal design-positioning and definition of concepts describing person-environment relationships. *Disability & Rehabilitation, 25*(2), 57-66.

Jones, W., Pirolli, P., Card, S. K., Fidel, R., Gershon, N., Morville, P. et al. (2006). *"It's about the information stupid!": why we need a separate field of human-information interaction.* Paper presented at the Conference on Human factors in computing systems (CHI'06), Montreal, Quebec, Canada.

Khaslavsky, J. (1998). *Integrating Culture into Interface Design.* Paper presented at the Conference on Human factors in computing systems (CHI'98), Los Angeles, USA.

Kreijns, K., & Kirschner, P. (2001). *The social affordances of computer-supported collaborative learning environments.* Paper presented at the 31st Annual Frontiers in Education Conference, Reno, USA.

Marcus, A., & Gould, E. M. (2000). Cultural Dimensions and Global Web User-Interface Design. *Interactions, 7*(4), 32 - 46.

Matusov, E. (1996). Intersubjectivity without Agreement. *Mind, Culture, and Activity, 3*(1), 25-45.

Nisbett, R., & Miyamoto, Y. (2005). The influence of culture: holistic versus analytic perception. *TRENDS in Cognitive Sciences, 9*(10).

Nisbett, R., & Norenzayan, A. (2002). Culture and Cognition. In D. L. Medin (Ed.), *Stevens' Handbook of Experimental Psychology* (3rd ed., pp. 561–597).

Nisbett, R., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review, 108*(2), 291-310.

Orlikowski, W. J. (1992). The Duality of Technology: Rethinking the Concept of Technology in Organizations. *Organization Science, 3*(3), 398-427.

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist, 15*, 776-783.

Reeves, B., & Nass, C. (1996). *The media equation: how people treat computers, television, and new media like real people and places*: Cambridge University Press New York, NY, USA.

Ross, N. (2004). *Culture and Cognition: Implications for Theory and Method*: Sage Publications.

Russo, P., & Boor, S. (1993). *How Fluent is your Interface? Designing for International Users.* Paper presented at the Conference on Human factors in computing systems (CHI'93), Amsterdam, The Netherlands.

Schwartz, S. H., Melech, G., Lehmann, A., Burgess, S., Harris, M., & Owens, V. (2001). Extending the cross-cultural validity of the theory of basic human values with a different method of measurement. *Journal of Cross-Cultural Psychology, 32*(5), 519–542.

Shi, Q., & Clemmensen, T. (2007). Relationship Model in Cultural Usability Testing. In N. Aykin (Ed.), *Usability and Internationalization - Second International Conference on Usability and Internationalization, UI-HCII 2007, Held as Part of HCI International 2007, Beijing, China, July 22-27, 2007, Proceedings, Part I* (Vol. 4559, pp. 422-431). Heidelberg: Springer.

Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-Supported Collabortive Learning. In R. K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 409-425). Cambridge, UK: Cambridge University Press.

Stoffregen, T. A. (2003). Affordances as Properties of the Animal-Environment System. *Ecological Psychology, 15*(2), 115-134.

Suinn, R. M., Ahuna, C., & Khoo, G. (1992). The Suinn-Lew Asian Self-Identity Acculturation Scale: Concurrent and Factorial Validation. *Educational & Psychological Measurement, 52*(4), 1041-1046.

Sun, C., Jia, X., Zhang, Y., Yang, Y., & Chen, D. (1998). Achieving convergence, causality preservation, and intention preservation in real-time cooperative editing systems. *ACM Transactions on Computer-Human Interaction (TOCHI), 5*(1), 63-108.

Suthers, D. (2006). Technology affordances for intersubjective meaning-making: A research agenda for CSCL. *International Journal of Computers Supported Collaborative Learning, 1*(3), 315-337.

Suthers, D., Vatrapu, R., Medina, R., Joseph, S., & Dwyer, N. (2008). Beyond Threaded Discussion: Representational Guidance in Asynchronous Collaborative Learning Environments. *Computers and Education, 50*(4), 1103-1127.

Tannen, D. (1996). *Gender and Discourse*: Oxford University Press.

Triandis, H. C., Kashima, Y., Shimada, E., & Villareal, M. (1986). Acculturation Indices as a Means of Confirming Cultural Differences. *International Journal of Psychology, 21*, 43-70.

Turvey, M. T. (1992). Affordances and Prospective Control: An Outline of the Ontology. *Ecological Psychology, 4*(3), 173-187.

Vatrapu, R. (2007). *Technological Intersubjectivity and Appropriation of Affordances in Computer Supported Collaboration.* Unpublished Doctoral Dissertation, University of Hawaii at Manoa, Honolulu: Available at http://www.itu.dk/people/rkva/docs/Vatrapu-Dissertation.pdf.

Vatrapu, R. (2008). Cultural Considerations in Computer Supported Collaborative Learning. *Research and Practice in Technology Enhanced Learning, 3*(2), 159-201.

Vatrapu, R. (2009a). Toward a Socio-Technical Theory of Culture. *Paper presented at the Culture and Technologies for Social Interaction Workshop at the 12th IFIP TC13 Conference in Human-Computer Interaction (INTERACT 2009), Uppsala, Sweden*, Available at http://www.itu.dk/people/rkva/docs/2009-INTERACT-SocioTechnical-Theory-of-Culture.pdf.

Vatrapu, R. (2009b). Toward a Theory of Socio-Technical Interactions in Technology Enhanced Learning Environments. In U. Cress, V. Dimitrova & M. Specht (Eds.), *EC-TEL 2009, Lecture Notes in Computer Science (LNCS) 5794* (pp. 694-699). Berlin Heidelberg Springer-Verlag.

Vatrapu, R., & Pérez-Quiñones, M. (2006). Culture and Usability Evaluation: The Effects of Culture in Structured Interviews. *Journal of Usability Studies, 1*(4), 156-170.

Vatrapu, R., & Suthers, D. (2007). Culture and Computers: A Review of the Concept of Culture and Implications for Intercultural Collaborative Online Learning. In T. Ishida, S. R. Fussell & P. T. J. M. Vossen (Eds.), *Intercultural Collaboration I : Lecture Notes in Computer Science* (pp. 260-275): Springer-Verlag

Vatrapu, R., & Suthers, D. (2009a). Is Representational Guidance Culturally Relative? In C. O'Malley, D. Suthers, P. Reimann & A. Dimitracopoulou (Eds.), *Computer Supported Collaborative Learning Practices: CSCL2009 Conference Proceedings, June 8-13, 2009* (pp. 542-551). Rhodes, Greece: International Society of the Learning Sciences.

Vatrapu, R., & Suthers, D. (2009b). Technological Intersubjectivity in Computer Supported Intercultural Collaboration *Proceeding of the 2009 international Workshop on intercultural Collaboration (Palo Alto, California, USA, February 20 - 21, 2009). IWIC '09* (pp. 155-164). New York, NY: ACM.

Vatrapu, R., Suthers, D., & Medina, R. (2008). Usability, Sociability, and Learnability: A CSCL Design Evaluation Framework. *Proceedings of the 16th International Conference on Computers in Education (ICCE 2008)*, (CD-ROM).

Wallace, S., & Yu, H. (2009). The Effect of Culture on Usability: Comparing the Perceptions and Performance of Taiwanese and North American MP3 Player Users. *Journal of usability Studies, 4*(3), 136-146.

Wertsch, J. (1985). *Vygotsky and the social formation of mind*. Cambridge, MA, USA: Harvard University Press.

Wertsch, J. (1998). *Mind as action*. New York: Oxford University Press.

Witkin, H. A. (1967). A Cognitive-Style Approach to Cross-Cultural Research. *International Journal of Psychology, 2*(4), 233-250.

Witkin, H. A., & Goodenough, D. (1977). Field dependence and interpersonal behavior. *Psychological Bulletin, 84*(4), 661-689.

Yammiyavar, P., Clemmensen, T., & Kumar, J. (2008). Influence of cultural background on non-verbal communication in a usability testing situation. *International Journal of Design, 2*(2), 31-40.

Yammiyavar, P., & Goel, K. M. (2006, May 14-15). *Emphasis on non-verbal cues for interpreting cognitive processes in protocol analysis.* Paper presented at the Indo-Danish HCI Research Symposium Guwahati, India.

Yeo, W. A. (2001). Global-Software Development Life Cycle: An Exploratory Study. *Proceedings of the SIG-CHI Conference on Human Factors in Computing Systems*, 04 -11.

About the Authors

Ravi Vatrapu
Ravi Vatrapu is an assistant professor at the Center for Applied ICT of the Copenhagen Business School. His research interests are in comparative informatics, computer-supported intercultural collaboration, computer-supported collaborative learning, and digital democracy. His current focus is on conducting theory-based empirical studies of socio-technical affordances and developing an empirically-informed theory of technological intersubjectivity. He holds a PhD in Communication and Information Sciences from the University of Hawaii at Manoa.

Dan Suthers
Dan Suthers is professor in the Department of Information and Computer Sciences at the University of Hawai`i at Manoa, where he directs the Laboratory for Interactive Learning Technologies and is Chair of the interdisciplinary Communication and Information Sciences program. His current focus is on social affordances—how software interfaces influence collaborative learning and problem solving processes in small groups and the formation of relationships and social capital in online communities. He holds a PhD in computer science from the University of Massachusetts-Amherst.