

# SUPPORTING LEARNER SOCIAL RELATIONSHIPS WITH ENCULTURATED PEDAGOGICAL AGENTS

AMY OGAN

February 2011  
CMU-HCII-11-101

Human-Computer Interaction Institute  
School of Computer Science  
Carnegie Mellon University  
Pittsburgh, PA 15213

## THESIS COMMITTEE

Vincent Alevan, Carnegie Mellon University (Co-chair)  
Christopher M. Jones, Carnegie Mellon University (Co-chair)  
Sara Kiesler, Carnegie Mellon University  
Randall W. Hill, Jr., University of Southern California

*Submitted in partial fulfillment of the requirements  
for the Degree of Doctor of Philosophy.*

Copyright © 2011 Amy Ogan. All rights reserved.

This research has been sponsored by the U.S. Army Research, Development, and Engineering Command (RDECOM). Statements and opinions expressed do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred. It was supported in part by the Institute of Education Sciences grant R305B040063. The Pittsburgh Science of Learning Center, funded by National Science Foundation award number SBE-0354420, has also provided a research home.

**Keywords**

Virtual agents, pedagogical agents, social relationships, social dialog, dialog modeling, serious games, simulation-based learning, educational technology, intercultural competence, interpersonal skills training, human-computer interaction.

# Abstract

Embodied conversational agents put a “human” touch on intelligent tutoring systems by using conversation to support learning. When considering instruction in interpersonal domains, such as intercultural negotiation, the development of an interpersonal relationship with one’s pedagogical agent may play a significant role in learning. However, there is conflicting evidence in the literature both regarding the ability of agents to cultivate social relationships with humans, and their effect on learning. In this dissertation, I present a model of social dialog designed to affect learners’ interpersonal relations with virtual agents, a development process for creating social dialog, and empirical studies showing that this dialog has significant effects on learners’ perceptions of the agents and negotiation performance.

In early work, I explicitly prompted learners to have social goals for the interaction. I found that while students who reported social goals for interacting with the agents had significantly higher learning gains, explicit prompting was not effective at inducing these goals. I thus focused on implicit influence of learner goals, developing a model of social instructional dialog (SID) that integrates conversational strategies that are theorized to produce interpersonal effects on relationships. In two subsequent studies, an agent with the SID model engendered greater feelings of entitativity, shared perspective, and trust, suggesting that the model improved learner social relationships with the agent. Importantly, these effects transferred to other agents encountered later in the environment. The social dialog condition also made fewer errors and achieved more negotiation objectives in a subsequent negotiation than a control group, evidence that the improved social relationship lead to better negotiation performance. These findings regarding interpersonal relationships with agents contribute to the literature on learner-agent interactions, and can guide the future development of agents in social environments.

# Acknowledgments

The road to any dissertation is paved with the support of innumerable people along the way. I would like to thank my advisors, Vincent Aleven and Chris Jones, for their guidance, advice, and their patience. They have helped me integrate two very different perspectives to take a direction that was new and exciting to all of us. It's been a long road that we've traveled together, and it appears my journey at Carnegie Mellon will not end here!

My work has also benefitted greatly from discussion with my committee. Sara Kiesler has assisted in fine-tuning every detail, while Randy Hill has inspired me to consider the broader impact of the research.

The project could not have been undertaken without the support of the amazing team at the Institute for Creative Technologies. Julia Kim has been a tireless advocate and advisor, and a fierce debater over the meaning and role of culture. Eric Forbell, Kim LeMasters, Glenn Storm, Daniel Auerbach, Kelly Christoffersen, and many others at the ICT and Psychic Bunny have provided technical and design support that made Zahora come to life.

Queenie Kravitz, Jo Bodnar, Sharon Carver, Richard Scheines, Howard Seltman, Sue Fussell, Stacy Marsella, Ruth Wylie, Moira Burke, Ian Li, Scott Davidoff, H. Lane, Matt Easterday, Gary Hsieh, Giancarlo Dozzi, Turadg Aleahmad, Ido Roll, Jason Wiese, Stephen Oney, and all of the brilliant PIER, PSLC, and HCI students have provided many and varied contributions.

My family has pushed me every step of the way to be what I wanted to be, be as industrious as I could be, and always strive for my best. Words can't express the thanks they deserve for shaping my values and giving me everything.

Erin Walker, life wouldn't be the same without you.

Finally, this dissertation is dedicated to my fiancé, Chris Harrison. I can only hope to provide him with the unconditional love and support that he has shown me. Thank you for being exactly who you are.

# Table of Contents

1. INTRODUCTION.....	11
2. INTERCULTURAL TRAINING.....	17
2.1. DEFINING CULTURE AND CULTURAL GROUPS .....	18
2.1.1. CULTURAL GROUPS BEYOND NATIONALITIES.....	18
2.1.2. APPROACHES TO UNDERSTANDING CULTURE.....	19
2.2. UNIVERSALISMS AND CULTURAL SPECIFICS.....	20
2.2.1. SYSTEM OF VALUES PARADIGM.....	21
2.3. THE DEVELOPMENT OF INTERCULTURAL COMPETENCE .....	23
2.3.1. STAGES OF INTERCULTURAL COMPETENCE .....	24
2.3.2. SKILLS OF INTERCULTURAL COMPETENCE.....	25
2.4. CULTURAL PEDAGOGY AND INSTRUCTION .....	26
2.4.1. HISTORY OF INTERCULTURAL EDUCATION.....	27
2.4.2. INTRODUCTION OF TECHNOLOGY INTO TRAINING .....	29
2.4.3. EVALUATION AND ASSESSMENT .....	30
3. VIRTUAL LEARNING ENVIRONMENTS AND AGENTS FOR CULTURAL INSTRUCTION.....	33
3.1. VIRTUAL ENVIRONMENTS FOR CULTURAL INSTRUCTION .....	34
3.1.1. ADAPTIVE THINKING AND LEADERSHIP SYSTEM (ATL).....	35
3.1.2. CROQUELANDIA.....	37
3.1.3. SECOND CHINA .....	38
3.1.4. TACTICAL LANGUAGE AND CULTURE TRAINING SYSTEMS.....	40
3.1.5. VIRTUAL ENVIRONMENT CULTURAL TRAINING FOR OPERATIONAL READINESS.....	42

3.1.6. BiLAT.....	44
3.2. EMBODIED ENCULTURATED CONVERSATIONAL AGENTS .....	52
3.2.1. FRAMEWORKS FOR CULTURALLY-AWARE VIRTUAL AGENTS.....	52
3.2.2. USER PERCEPTIONS OF VIRTUAL AGENTS.....	55
<b>4. STUDY 1: EXPERT VALIDATION .....</b>	<b>57</b>
4.1. ARMY SUBJECT MATTER EXPERT INTERVIEWS .....	58
4.1.1. PARTICIPANTS.....	58
4.1.2. METHOD .....	59
4.1.3. OBSERVATIONS.....	60
4.2. IRAQI SUBJECT MATTER EXPERT INTERVIEWS.....	64
4.2.1. PARTICIPANTS.....	64
4.2.2. METHOD .....	65
4.2.3. OBSERVATIONS.....	66
4.3. DISCUSSION AND CONTRIBUTIONS .....	67
<b>5. STUDY 2: NOVICE EXPERIENCE OF GAMEPLAY .....</b>	<b>69</b>
5.1. SOCIAL CONSTRUCT: LOCUS OF CONTROL .....	70
5.2. PARTICIPANTS.....	71
5.3. MEASURES .....	71
5.4. METHOD .....	72
5.5. VERBAL DATA RESULTS.....	72
5.5.1. GAME-RELATED GOALS AND ATTRIBUTIONS .....	72
5.5.2. CULTURAL ATTRIBUTIONS.....	74
5.5.3. METACOGNITIVE MONITORING AND PERCEPTIONS OF CONTROL.....	75
5.6. LOCUS OF CONTROL RESULTS .....	77
5.7. DISCUSSION AND CONTRIBUTIONS .....	79
<b>6. STUDY 3: MANIPULATING LEARNERS' SOCIAL GOALS THROUGH EXPLICIT PROMPTS.....</b>	<b>82</b>
6.1. SOCIAL CONSTRUCT: SOCIAL GOALS .....	82

6.1.1. GOAL ORIENTATION.....	82
6.1.2. INDUCING SOCIAL GOALS.....	84
6.2. EXPERIMENTAL DESIGN.....	85
6.2.1. PARTICIPANTS .....	86
6.2.2. MEASURES.....	86
6.2.3. PROCEDURE.....	89
6.3. RESULTS .....	91
6.3.1. LEARNING GAINS .....	91
6.3.2. REPORTING OF SOCIAL GOALS .....	92
6.3.3. LEARNING GAINS BY REPORTED GOALS .....	93
6.3.4. GAMEPLAY RESULTS .....	94
6.3.5. GENDER GAP .....	95
6.4. DISCUSSION AND CONTRIBUTIONS.....	95
<b>7. SOCIAL INFORMATIONAL DIALOG DEVELOPMENT .....</b>	<b>99</b>
7.1. SOCIAL CONSTRUCTS: TRUST, ENTITATIVITY, AND SHARED PERSPECTIVE.....	100
7.1.1. TRUST .....	100
7.1.2. ENTITATIVITY .....	101
7.1.3. PERSPECTIVE .....	101
7.2. DIALOG DEVELOPMENT PROCESS.....	102
7.2.1. PREPARATION.....	104
7.2.2. DIALOG GENERATION .....	106
7.2.3. ITERATIVE DEVELOPMENT .....	109
7.2.4. VALIDATION WITH IRAQI NATIONALS .....	113
7.3. SOCIAL INSTRUCTIONAL DIALOG MODEL .....	114
7.3.1. NARRATIVE.....	116
7.3.2. AFFIRMATION.....	117
7.3.3. SELF-DISCLOSURE .....	118
7.4. TASK INFORMATIONAL DIALOG MODEL .....	119
7.5. CONTRIBUTIONS.....	120

<b>8. STUDY 4: LEARNERS' INTERPERSONAL RELATIONS WITH AN INSTRUCTIONAL AGENT .....</b>	<b>122</b>
8.1. PARTICIPANTS.....	122
8.2. METHOD .....	123
8.3. MEASURES .....	123
8.4. RESULTS.....	124
8.5. DISCUSSION AND CONTRIBUTIONS.....	127
<b>9. STUDY 5: COGNITIVE AND INTERPERSONAL EFFECTS OF SOCIAL INFORMATIONAL DIALOG .....</b>	<b>131</b>
9.1. SOCIAL CONSTRUCTS .....	132
9.2. EXPERIMENTAL DESIGN .....	133
9.3. PARTICIPANTS.....	134
9.4. MEASURES .....	134
9.4.1. LEARNER CHARACTERISTICS .....	134
9.4.2. ATTITUDES.....	136
9.4.3. MANIPULATION CHECK: PERCEPTIONS OF ACTIVITY AND DIALOG .....	136
9.4.4. SOCIAL AND INTERPERSONAL MEASURES.....	137
9.4.5. LEARNING .....	138
9.5. PROCEDURE.....	139
9.6. RESULTS.....	142
9.6.1. SCALE RELIABILITY AND RANDOMIZATION CHECKS.....	142
9.6.2. H1: SOCIAL AND INTERPERSONAL MEASURES .....	146
9.6.3. H2: SOCIAL GOALS.....	150
9.6.4. H3: LEARNING RESULTS.....	153
9.6.5. INFLUENCE OF LEARNER CHARACTERISTICS ON INTERPERSONAL AND LEARNING OUTCOMES .....	159
9.7. DISCUSSION AND CONTRIBUTIONS .....	162

<b>10. DISCUSSION, CONTRIBUTIONS, AND FUTURE DIRECTIONS .....</b>	<b>168</b>
10.1. LEARNING SCIENCES.....	169
10.1.1. EXPLICIT PROMPTING OF SOCIAL GOALS.....	169
10.1.2. LEARNING INTERPERSONAL SKILLS FROM A VIRTUAL ENVIRONMENT.....	170
10.1.3. INTERACTIONS IN THE LEARNING ENVIRONMENT .....	171
10.2. INTERCULTURAL COMPETENCE TRAINING .....	173
10.2.1. SUCCESSFUL STRATEGIES AND MISCONCEPTIONS.....	173
10.2.2. CONTRIBUTING CHARACTERISTICS TO INTERCULTURAL COMPETENCE	174
10.2.3. SIMULATION VALIDATION AND LEARNING GAINS.....	175
10.3. VIRTUAL AGENTS.....	177
10.3.1. DIALOG MODEL AND DEVELOPMENT PROCESS.....	177
10.3.2. SOCIAL EFFECTS IN LEARNER-AGENT INTERACTIONS.....	179
10.3.3. ABILITY OF AGENTS TO INFLUENCE LEARNER ATTITUDES.....	181
10.4. FUTURE DIRECTIONS .....	182
<b>11. REFERENCES.....</b>	<b>184</b>
<b>APPENDIX A.....</b>	<b>196</b>
SAMPLE ITEMS FROM THE CULTURE ASSIMILATOR ASSESSMENT .....	196
<b>APPENDIX B .....</b>	<b>199</b>
INSTRUCTIONS FOR THE SITUATIONAL JUDGMENT TEST .....	199
SAMPLE ITEMS FROM THE SITUATIONAL JUDGMENT TEST.....	199
<b>APPENDIX C.....</b>	<b>201</b>
REFACTORED SOCIAL GOALS QUESTIONNAIRE .....	201

# I. INTRODUCTION

Globalization has produced a dramatic increase in communication (and consequently miscommunication) between people from diverse cultures. The United States, which has always been a melting pot of cultures, is projected to become a “nation of minorities” by 2050 ([www.census.gov](http://www.census.gov)). Similarly, by 2031, one third of Canadian citizens are expected to have a mother tongue other than French or English ([www.statcan.gc.ca](http://www.statcan.gc.ca)). Technology has been a driving force behind these changes, by enabling greater international mobility and communication for everyone from rural children to Wall Street bankers. When everyday interactions were between people from the same geographic region, who shared a common ground in social norms and expectations, issues of miscommunication and friction resulting from cultural differences were of less concern. Living as a global citizen, however, necessitates widespread intercultural education. UNESCO defines the objectives of intercultural education as “cultural knowledge, attitudes and skills that enable [people] to contribute to respect, understanding and solidarity” (2007). While technology has provided a fantastic opportunity for intercultural exchange, with that opportunity comes certain obstacles to understanding. In my thesis, I argue that technology is an integral part to overcoming those obstacles, by demonstrating how intercultural competence skills can be improved using virtual learning environments.

Virtual learning environments provide new and unique ways in which to convey cultural knowledge and develop intercultural communication skills. Examples of existing game-based instructional systems with social learning components

include FearNot (Hall, Vala, Hall, Webster, Woods, Gordon, & Aylett, 2006), PeaceMaker (Burak, Keylor, & Sweeney, 2005), and BiLAT (Kim, Hill, Durlach, Lane, Forbell, Core, et al., 2010). In classroom-based instruction on cultural knowledge and communication, these skills are often taught through methods like role-playing exercises and tutoring, which are very resource-intensive (Landis, Bennett, & Bennett, 2003). Thus, simulation-based instructional systems are increasingly being used to facilitate learning social or interpersonal skills such as conflict resolution, by simulating human behavior with virtual characters (Raybourn & Waern, 2004). The successful development of virtual environments for teaching intercultural skills is only now becoming a reality, and their effects on learning are still unclear (Ogan & Lane, 2010). However, recent advances in artificial intelligence (AI) and cognitive modeling now permit rich modeling of emotions, language, tasks, and more (e.g., Swartout, Gratch, Hill, Hovy, Marsella, Rickel, & Traum, 2006; also see Rehm, 2010 for more information on cultural modeling for training and embodied conversational agents). When built with cultural accuracy, these models - and the virtual humans who utilize them - open new avenues for teaching the cognitive and interpersonal aspects of other cultures.

These systems have particular advantages over traditional instruction of cultural skills. High-fidelity graphics, sound, and animation make it possible to simulate many tangible aspects of a specific culture, such as buildings, streets, art, dress, speech, gestures, and more. These physical artifacts can enable the provision of more authentic computer-based practice environments than is feasible using traditional classroom role-play and media-based approaches, which may not have the same level of detail or realism. In addition, agents allow students to practice culturally-appropriate behaviors with less risk of social embarrassment in front of peers, which is a common source of reticence in language classrooms (Horwitz, Horwitz, & Cope, 1986). Agents have the advantage of being able to replicate specific behaviors repeatedly and precisely, or modifying

their behavior to be consistent with changing cultural landscapes. The use of an agent system also enables tracking of the learner's behavior and progress, which allows the system to present personal feedback by contrasting the learner's behavior with appropriate target culture behaviors. Once developed, these simulations offer a major advantage for social learning by providing an easily deployable solution to a wide audience.

If it is true, as Reeves and Nass' Media Equation suggests, that people respond to computers as if they were humans (Reeves & Nass, 1996), then it is likely that people also form social relationships with the virtual personalities that exist in these environments. There have been several recent studies, however, that show that this social phenomenon might be more nuanced than previously believed. Rosé and Torrey (2005) found that students displayed more productive learning behaviors when they believed there was a human driving the responses behind the instructional dialog system they were using. Additionally, Okita, Bailenson, and Schwartz (2008) found that students who believed there was a human behind an avatar in a virtual environment exhibited better learning, more attention, and higher arousal. These studies suggest that the kinds of relationships students form with partners they believe have social agency might in fact be most beneficial for learning.

One explanation for the discrepancies between these results is that learning is not an automatic process like emotion but requires attention and processing (Cohen, Ivry, & Keele, 1990). If students believe they are taking a socially relevant action when they interact with a human, they pay more attention and feel more accountable. Many researchers therefore focus on developing increasingly sophisticated social models to drive the behaviors of virtual agents in the hopes that greater realism will increase students' learning (Tomlinson, 2005). For this tactic to work, it is important to investigate which aspects of agent realism affect student learning, and examine the mechanisms behind these effects. My dissertation focuses on two key questions: what benefits are there

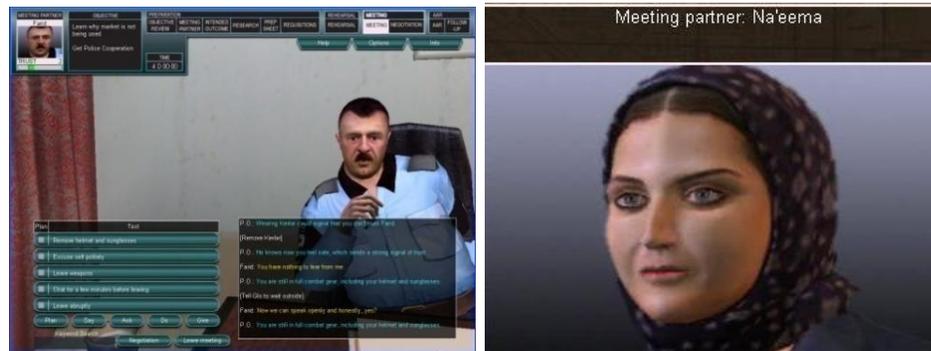


Figure I.1. At left, a meeting in BiLAT with police captain Farid with the goal of solving a problem with a market in an Iraqi town. At right, meeting partner Na'eema, a doctor.

to interacting with virtual agents in realistically social ways? If there are benefits, what features of agents encourage learners to perceive them socially?

I explore this research agenda within BiLAT, a virtual environment for practicing bilateral negotiation in a cross-cultural context (see Figure I.1). In BiLAT, the learner is put into the role of a U.S. Army officer tasked with meeting with Iraqi townspeople (simulated by virtual agents) to accomplish peace-keeping and rebuilding missions. In each meeting, learners are given concrete, negotiation task-related goals, and must interact with a different virtual agent using menu-based communication. The virtual agents each have an underlying model of culture and personality that drives their responses to the learner, as well as their gestures, gaze, and posture (see Hill, Belanich, Lane, Core, et al., 2006). The aim of BiLAT is not simply to guide students towards obtaining the most lucrative negotiation outcomes, but rather to successfully build lasting relationships through cultural understanding. By rewarding students who take actions that conform to appropriate cultural behaviors, BiLAT addresses cultural concepts like polychronicity (differences in the meaning of time), the value of social relationships and building trust, and face-saving, along with basic etiquette concepts like greetings, giving gifts, and leave-taking. This dissertation focuses on how to design the virtual environment to improve the acquisition of cultural concepts and skills, by examining the role of

students' social perceptions of the virtual agents in acquiring this knowledge, and then investigating at how those perceptions are influenced by the features of the agents in the environment.

The research reported in this dissertation is composed of exploratory work with students and content experts from two fields, the iterative development of a model of social informational dialog, and three controlled experiments investigating social and learning outcomes. Throughout the course of the research, I explored a number of social constructs that may influence learners' interactions with social simulations and virtual agents. Each relevant construct is described in the chapter where it first appears, under the subsection "Social construct:." The dissertation is organized as follows.

In Chapters 2 and 3, I review related work in intercultural competence and agent-based virtual learning environments. Chapters 4 and 5 describe my preliminary work in these areas. First, I conducted interviews with content experts within the Army and the Iraqi communities, using BiLAT as the context for understanding their experiences in intercultural situations. Additionally, I conducted a think-aloud study while learners played with the BiLAT environment, to understand how students interact with the game and the virtual characters. In this study, I explored the *locus of control* construct and how beliefs about social agency relate to success in the environment.

Based on the results of this preliminary work, Chapter 6 describes a study investigating learners' *social goals*. Within BiLAT, I developed an intervention to explicitly scaffold learners' goals for interaction with the virtual agents. This study hypothesized that students who hold more social goals both interact more socially with the virtual agents in the simulation, and benefit more from the interaction.

Given that this study found that holding social goals was beneficial for learning, Chapter 7 describes the process of development of a model for social

instructional dialog which focuses on implicitly manipulating learners' social interactions. This model links conversational strategies available to agent dialog systems, with three social effects: *trust*, *entitativity*, and *shared perspective*. These social constructs are taken from human-human communication literature and are hypothesized to influence learning.

Chapters 8 and 9 present the findings from two studies using an agent developed with this dialog model. In the first study, I examined learners' interactions with the agent in isolation from the rest of the learning experience, in order to determine the effects of the dialog on the three social constructs. In the second study, I integrated the agent into BiLAT to understand these interactions, and how they affect learning, in the context of an educational experience. Also in this chapter, I describe how social constructs are influenced by and interact with learner characteristics such as cultural intelligence (Ang, Van Dyne, Koh, Ng, Templer, Tay, & Chandrasekar, 2007), social intelligence (Silvera, Martinussen, & Dahl, 2001), and personality traits such as extroversion (Wiggins, 1979).

In the final chapter I discuss the central findings and contributions of this work in three areas: learning sciences, intercultural training, and virtual agent literature. With respect to learning sciences, this work increases our understanding of how holding a social orientation influences learners' interactions with virtual agents, and finds that it can be promoted through social agent dialog in a way that is beneficial to learning. Contributions to intercultural training include enumerating successful strategies and misconceptions held by experts and novices, exploring what learner characteristics are associated with intercultural competence, and validating the content and investigating learning gains from an intercultural simulation. Contributions to understanding human-agent interactions include formalizing a dialog development process and creating a model for social informational dialog, investigating the social effects that this model has on learner-agent

interactions, and uncovering the attitude change associated with differing perceptions of the agents. As a result of this dissertation, I demonstrated that social human-agent relationships matter in virtual environments for interpersonal learning, and through careful agent design, can be engendered in the learner.

## 2. INTERCULTURAL TRAINING

Cultural understanding is critical in many contexts, from language classrooms to business negotiations and international relations or service abroad (Landis, Bennett, & Bennett, 2003). This section defines culture and intercultural competence within the framework of this dissertation, reviews current theory on the development of the skills associated with intercultural competence, and describes classroom pedagogy and instructional interventions aimed at promoting these skills.

### 2.1. DEFINING CULTURE AND CULTURAL GROUPS

#### 2.1.1. CULTURAL GROUPS BEYOND NATIONALITIES

Before discussing the nature of culture itself, it is essential to define the notion of a cultural group. Cultural groups are a coherent and stable ensemble of individuals to which a culture can be associated (Blanchard, Mizoguchi, & Lajoie., 2010). The concept of cultural group is frequently simplified to mean large human groups such as countries or religions. However, numerous sub-groups which do not fall under this definition are useful for explaining individuals' behavioral and cognitive characteristics. Hence, any group of individuals coherent enough to develop a specific set of such characteristics can be considered a cultural group. This includes, for example, businesses, communities of interest or practices (e.g., sport fans and carpenters, respectively; for a more complete overview, see Lave & Wenger (1991)). Most

individuals today are subject to multiple cultural influences, sometimes referred to as layers of cultural identity (Rehm, 2010; Reinecke, Schenkel, & Bernstein, 2010).

### 2.1.2. APPROACHES TO UNDERSTANDING CULTURE

A very large number of definitions for culture have been proposed, which are often strongly influenced by the interests of a particular discipline. In cross-cultural psychology, Kashima (2000) states that two schools of thought exist, defining culture either as “a process of production and reproduction of meanings in particular actors’ concrete practices or actions or activities in particular contexts in time and space”, or as “a relatively stable system of shared meanings, a repository of meaningful symbols, which provides structure to experience”.

Cooper and Denner (1998) present several theoretical approaches to the study of culture in human and social sciences. Their focus varies in order to consider, among other things, (a) core cultural ideas and the key role of shared social values in shaping individuals’ cognitive, affective and social processes, (b) the interpretation of individuals’ characteristics with regards to their surrounding social and material context, (c) the consequences of differences in social position among cultural subgroups in historical and cultural context, or (d) how individuals develop and claim membership in specific socio-cultural group(s) and its implications for intergroup relations.

It should be noted that in the approaches presented above, the notion of culture is considered in terms of cognitive and behavioral implications. However, this approach may be too restrictive from the viewpoint of cultural training. For instance, other domains, such as archaeology or anthropology, heavily consider physical cultural artifacts. Such information is highly relevant

when designing a virtual cultural environment or when looking for concrete examples to include as pedagogical resources.

## 2.2. UNIVERSALISMS AND CULTURAL SPECIFICS

Two main approaches in cultural studies that are useful for instruction consist of identifying universalisms and group specifics. Universalisms are intrinsic characteristics of human beings and as such, are supposedly shared by a wide cluster of cultural groups (if not all). Group specifics are characteristics specific to cultural groups in that they are understood or endorsed by an important portion of insiders, and unknown or considered external by outsiders. Discussing universalisms or group specifics is equivalent to eliciting cultural aspects that unite all human groups, versus those that distinguish each of them. Oversimplification is a key concern when addressing group specifics: a given characteristic of a cultural group is rarely shared by all its members (Scharifian, 2003). In order to discriminate between cultural groups, scholars frequently suggest attaching a pool of common characteristics to a cultural group (Scharifian, 2003) rather than referring to one unique characteristic.

Universalisms have been posited in many aspects of human life including facial expressions of emotions (Ekman, 1972), motivation (Ryan & Deci, 2000), and politeness (Brown & Levinson, 1987), to cite but a few. Cultural specifics are similarly reported along many dimensions, including cognitive (e.g., core cultural ideas, interpretations, beliefs), behavioral (e.g., body language, rituals, good practices), and physical (e.g., artifacts) (see Blanchard, Mizoguchi, & Lajoie, 2010). Although frequently presented as universals, empirical research has demonstrated group specifics in such aspects of human life as basic emotions (Mesquita, Frijda, & Scherer, 1997), frequency of personality profiles (Allik & McCrae, 2004), basic wellbeing needs (Hofstede, 1984), and cognitive processing (Nisbett & Norenzayan, 2002).

## 2.2.1. SYSTEM OF VALUES PARADIGM

Some approaches to understanding cultures include both universalism and group-specific considerations. System of values is a practical approach to describing cultures that emerged decades ago. It consists of identifying universal dimensions of the major orientations of cultural groups (their behavioral and cognitive tendencies) in order to develop group-specific models, thus providing an easy method for cross-cultural comparisons and assessments, and for potentially explaining cultural specifics. At present, the most popular system of values results from the analysis of a cross national survey of more than 100,000 people by Hofstede (2001; 2010). It characterizes more than 70 national cultures by computing their numeric scores for the following five dimensions: a) power distance (PDI: “the extent to which the less powerful members of organizations and institutions (like the family) accept and expect that power is distributed unequally”), b) individualism/collectivism (IDV: “the degree to which individuals are integrated into groups”), c) masculinity/femininity (MAS: “the distribution of roles between the genders”), d) uncertainty avoidance (UAI: “a society’s tolerance for uncertainty and ambiguity”), and e) long term orientation (LTO: a more recently added dimension referring to a general interest for “virtue regardless of truth”, where a high LTO would indicate greater respect for tradition and social obligations). Table 2.1 presents scores of Hofstede’s dimensions for a limited set of nations.

	<b>PDI</b>	<b>IDV</b>	<b>MAS</b>	<b>UAI</b>	<b>LTO</b>
<b>Brazil</b>	69	38	49	76	65
<b>Canada</b>	39	80	52	48	23
<b>India</b>	77	48	56	40	61
<b>Japan</b>	54	46	95	92	80
<b>USA</b>	40	91	62	46	29
<b>Arab World</b>	80	38	52	68	N/A

Table 2.1. Hofstede’s scores for six different nations, taken from Hofstede (2010).

Hofstede describes the Arab World, which includes Egypt, Iraq, Kuwait, Lebanon, Libya, Saudi Arabia, and the United Arab Emirates, according to the scores that deviate most from the average across all countries ([www.geert-hofstede.com](http://www.geert-hofstede.com)):

The lowest Hofstede Dimension for the Arab World is the Individualism (IDV) ranking at 38, compared to a world average ranking of 64. This translates into a Collectivist society as compared to Individualist culture and is manifested in a close long-term commitment to the member 'group', that being a family, extended family, or extended relationships. Loyalty in a collectivist culture is paramount, and over-rides most other societal rules.

The combination of these two high scores (UAI) and (PDI) create societies that are highly rule-oriented with laws, rules, regulations, and controls in order to reduce the amount of uncertainty, while inequalities of power and wealth have been allowed to grow within the society. These cultures are more likely to follow a caste system that does not allow significant upward mobility of its citizens.

When these two Dimensions are combined, it creates a situation where leaders have virtually ultimate power and authority, and the rules, laws and regulations developed by those in power, reinforce their own leadership and control. It is not unusual for new leadership to arise from armed insurrection – the ultimate power, rather than from diplomatic or democratic change.

Numerous studies have used Hofstede's framework in different contexts and disciplines (see Kirkman, Lowe, & Gibson, 2006). However Hofstede's method of analysis has been strongly criticized (McSweeney, 2002), and competing systems of values have emerged. Of particular note is GLOBE (House, Hanges, Javidan, Dorfman, & Gupta, 2004), which has garnered considerable attention in recent years and has the advantage of proposing dimensions that discuss cultural issues at both group and individual levels.

However the system of values paradigm is not to be considered a perfect solution. Indeed, many researchers continue to argue about its insufficient consideration of individual variations within cultural groups, as well as risks of

over-simplification that may result in cultural stereotyping (McSweeney, 2002). Furthermore, systems of values are mainly developed in business-related research fields. Consequently, dimensions that have been identified may not be well adapted for cross-cultural research in other domains (Blanchard, 2009; Stewart & Chakraborty, 2010). Indeed, a lack of education-focused studies, especially those based on student sampling, can be identified in Kirkman's listing of Hofstede's related studies (Kirkman, Lowe, & Gibson, 2006).

### 2.3. THE DEVELOPMENT OF INTERCULTURAL COMPETENCE

The domain of interest in this work is intercultural competence, that is, the development of skills that lead to improved understanding and interactions across cultures. In *Standards for Foreign Language Learning in the 21st Century*, The American Council on the Teaching of Foreign Languages has established standards for what learners should know and be able to do (ACTFL, 1996). A significant number of these focus on cultural understanding, e.g.:

1. Standard 3.2: Students acquire information and recognize the distinctive viewpoints that are only available through the foreign language and its cultures.
2. Standard 4.2: Students demonstrate understanding of the concept of culture through comparisons of the cultures studied and their own.

The document stresses the importance of going beyond a simple knowledge of cultural practices to reflecting and gaining insight on native perspectives, opinions, and values. For example, the *Cultura* project (Furstenberg, Levet, English, & Maillet, 2001) found that "The word *individualisme/individualism*, is a prime example where highly positive connotations of words such as 'freedom,' 'creativity,' and 'personal expression' appear on the American side, while the French side is replete with such negative notions as '*égoïsme*,' '*égoïsme*,' '*solitude*.'" Students do not make absolute conclusions about cultural issues (nor does this work encourage them to develop one single interpretation of

cultural phenomena), but rather use instructional materials to generate possible perspectives. In the frame of this work, this intelligence encompasses an ability to understand another's actions and thoughts with regards to his or her cultural specifics, to undertake actions in order to optimize positive as well as limit negative interactions with foreigners, and to develop culturally-informed perspectives of a socio-cultural environment (Kramsch, 1993).

### 2.3.1. STAGES OF INTERCULTURAL COMPETENCE

Although intercultural training is often motivated by immediate need, such as international travel, researchers and educators are also interested on its long-term impact on cultural learning and development. There is widespread agreement that intercultural growth occurs in stages and can take many years (Savicki, 2008). Whether it be a student studying abroad, or a business executive starting a new branch in a foreign country, the assumption that people acclimate gradually is both intuitive and generally supported by psychometric measures of cognitive, affective, and emotional change (Paige, 2004). Decades of research in cognitive psychology also lends support to this idea, as it is reported that development of expertise can take up to 10 years of study and practice in many domains (Ericsson, Charness, Feltovich, & Hoffman, 2006).

To understand the long-term development of intercultural competence, one of the more studied models is Bennett's Developmental Model of Intercultural Sensitivity (DMIS; Bennett, 1993). The DMIS rests on the assertion that as one's ability to construe cultural differences evolves, intercultural competence also increases. According to Bennett, "it is the construction of reality as increasingly capable of accommodating cultural difference that constitutes development" (p.24). The DMIS posits two broad worldview orientations: ethnocentrism and ethnorelativism, which refers to the positioning of one's

own culture in relation to others. An ethnocentric orientation implies that one views all other cultures relative to his or her own, whereas an ethnorelative perspective implies that one's own culture is understood in the context of others. Three sub-stages are included within each orientation that describe common cognitive and affective states that evolve during development. For Bennett, the goal of intercultural training is to promote gradual movement through the stages and deliver appropriate training given the learner's stage. If, for example, behavioral change is rushed, the learner may develop an impoverished understanding of the new culture. As with learning in most domains, it is important to avoid shallow learning and to develop an underlying conceptual understanding with better chances for retention and transfer.

Typically, the goal in intercultural development is to induce changes in knowledge, skills, and/or attitudes (Mendenhall, Stahl, Ehnert, Oddou, Osland, & Külmann, 2004). Knowledge includes basic facts about a new culture, such as common values and beliefs, preferences for physical contact, or typical eating and drinking patterns. Skills usually refer to the learner's ability to interact with someone from the new culture, including communicating desires and interpreting the behaviors of others. Finally, attitudes have to do with basic beliefs a learner has about people of a different culture and whether a positive, neutral, or negative disposition exists towards them. Development of intercultural competence should eventually lead students towards better integration in cultural contexts, e.g., by exhibiting fewer stereotypes and misconceptions, and employing better communication (Savicki, 2008).

### 2.3.2. SKILLS OF INTERCULTURAL COMPETENCE

More specifically, this dissertation focuses on two of the main skills of intercultural competence, which have been described in Byram's *savoirs*, or "competences" (1997). *Savoir-être* refers to the ability to approach intercultural

learning with curiosity, openness and reflexivity. It is described as an affective capacity to relinquish ethnocentric attitudes towards and perceptions of otherness, and relates strongly to the quality of empathy. *Savoir-faire* indicates the ability to interact in culturally appropriate ways by making appropriate cultural explanations for behaviors. Practicing these skills differs from simply knowing a few facts about a culture (e.g., that the French standard criterion for completing high school is the *Baccalauréat* exam) in that students must be able to produce appropriate cultural behaviors and take into account multiple points of view. These skills are difficult to acquire, in part because people often instinctively interpret events from their home culture's perspective, and much of their knowledge about culture is tacit. The DMIS model notes how rare it is to reach the final stage of complete intercultural competence (1993).

## 2.4. CULTURAL PEDAGOGY AND INSTRUCTION

Development of the skills of intercultural competence may be aided by the introduction of intercultural education. UNESCO (2007) has highlighted several objectives for this type of training:

1. To respect “the cultural identity of the learner through the provision of culturally appropriate and responsive quality education for all”,
2. To provide “every learner with the cultural knowledge, attitudes and skills necessary to achieve active and full participation in society”,
3. To provide “all learners with cultural knowledge, attitudes and skills that enable them to contribute to respect, understanding and solidarity among individuals, ethnic, social, cultural and religious groups and nations”.

These guidelines provide teachers and technologists with a focus for the development of future educational tools, those that incorporate two important facets of cultural intelligence (Earley & Mosakowski 2004). In line with UNESCO Guideline 1, instruction and technology should respect the culture of the user (i.e., make culturally-intelligent adaptations) by providing examples and

utilizing communication schemas that are more familiar to the user. Not only will this provide a more user-centered experience, but may also make learning more efficient. Second, following UNESCO Guidelines 2 and 3, educational tools should strive to provide opportunities to increase the learner's cultural intelligence. The following sections describe how they can be used to effectively introduce cultural knowledge, skills, and attitudes.

### 2.4.1. HISTORY OF INTERCULTURAL EDUCATION

Cultural training programs have evolved substantially in the last six decades. The earliest examples began to surface after World War II when international travel and collaboration became more prevalent in business and government work. As the need for these programs became more evident, scientific interest in creating theories of intercultural growth, identifying underlying cognitive processes, and demonstrating their effectiveness also grew. The field of intercultural training is highly interdisciplinary, attracting researchers from a variety of fields, including anthropology, cognitive psychology, social science, business, and more. Surprisingly, very little of this work leverages state of the art computing technology. The usual structure of intercultural training programs includes a blend of didactic and experiential components, including methods such as lectures, discussion, film, case study, and role playing (see Landis, Bennett, & Bennett, 2003 for a review of many types of cultural instruction). Many of these methods are based on a classroom instruction model and seek to leverage peer interaction and debate to engage learners.

A common approach in classroom instruction is to show multimedia in the form of videos, commercials, or other media, and then ask students to reflect and discuss in a classroom setting. Kitajima and Lyman-Hager (1998) suggest that "students can watch silent videos and discuss similarities and differences between, say, a train ride in their own country and in the target language community." In one such project at the University of Hawaii, students view

Japanese commercials to identify cultural stereotypes and then role-play, in discussion groups, versions of these commercials created from their observations. This classroom discussion format creates a synergy that can supersede the reflective capability of individual students. In another study (Herron & Dubreil, 2000), students watched French video as an 'advance organizational tool' to gather information that was shown to be helpful in answering cultural assessment questions. In a different take on cultural perspectives, constructionist research such as the Cultura project (Furstenberg, Levet, English, & Maillet, 2001) invites students to construct their own knowledge of cultural values and attitudes. This construction is accomplished, largely without instructor intervention, by having students answer questionnaires about their own culture and then communicate with a class in France to assess the authentic cultural descriptions provided by the other class's questionnaires. This method of cross-cultural learning is motivational and deeply informative, but involves a great deal of overhead in linking classes across continents.

Of the methods that provide experiential training, role-play is the most popular format for allowing students to practice the skills of intercultural communication. According to Landis, Bennett, & Bennett (2003), "done carefully and skillfully, role playing provides a dress rehearsal in a safe and supportive environment, with feedback and a chance to replay the situation or see others do it differently." Students gain decision-making abilities and must be adaptable to unpredictable changes in circumstances. However, one drawback is that students may be too shy or self-conscious to participate. Also, this instructional technique is resource-intensive; it requires a skilled trainer, often on a one-to-one basis, who is prepared to deal with unexpected behaviors and strong emotions from the learner.

#### 2.4.2. INTRODUCTION OF TECHNOLOGY INTO TRAINING

Fowler and Blohm (2004) suggest: "...if the goal is intercultural effectiveness, performance in multicultural settings, sensitivity, and establishing cross-cultural relationships, these cannot be done by a computer." (p. 40). On the contrary, there has been a small but innovative history of using technology to support intercultural interactions. "A 'culturally-aware system' refers to any system where culture-related information has/had some impact on its design, runtime or internal processes, structures, and/or objectives" (Blanchard, Mizoguchi, & Lajoie, 2010). There is a growing community of researchers who investigate cultural issues with regards to educational technology.

An early example of technology-enhanced instruction is text-based multi-user domains (MUDs). In these environments, there are no artificially intelligent agents. Instead, language students interact with each other online in an imaginary world where they can test their language skills with others and practice interacting in culturally influenced ways (Bruckman, 1995; Falsetti & Schweitzer 1995, etc). Another influential system was *A la rencontre de Philippe* (Furstenberg, 1993), a game in which students play as a French journalist using cultural knowledge to interact with the environment through branching storylines. Student journalists were tasked with helping a broken-hearted French man find a new apartment after being dumped by his girlfriend. These systems were the first to introduce the instructional methods of role-play to technology, facilitating cultural skill development by letting users experience situations in someone else's shoes. However, these systems offered little in the way of guidance, instead relying on experience-driven learning.

Intelligent tutoring systems (ITS) have been a successful way to introduce technology-enhanced guidance in well-defined domains such as algebra and physics (e.g., Alevan & Koedinger, 2002; Koedinger, Anderson, Hadley, & Mark, 2007; Anderson, Corbett, Koedinger, & Pelletier, 1995). However, their effectiveness in the domain of intercultural competence, and indeed in ill-

defined domains in general, has received less attention (see Lynch, Ashley, Alevan, & Pinkwart (2006) for a review of current work in these domains). Ogan, Alevan, and Jones (2010) describe how ITS principles might be adapted for learning in this domain. These principles can be used to develop interactive systems that help students examine cultural artifacts such as feature films or commercials. These systems cover cultural knowledge, analysis of cultural values and behaviors, and may have also focus on developing perspective-taking skills. ICCAT (InterCultural Competence Attention-focusing Tutor) is a tutor that enhances an existing classroom model for the development of intercultural competence (Ogan, Alevan, & Jones, 2010). Its on-line environment employs clips of feature films from a target culture intertwined with a set of attention-focusing techniques called pause-predict-ponder. A study in two French Online classrooms, comparing ICCAT versions with and without these techniques, found that the addition of pause-predict-ponder seemed to guide students in acquiring cultural knowledge and significantly increased students' ability to reason from an intercultural perspective. Further analysis of the posttest and students' online discussion found that students in the experimental condition were significantly assisted by making predictions of behavior, and were able to maintain a high quality of discussion over time.

Described in Chapter 3, virtual learning environments combined with intelligent tutoring systems may be the next evolution in intercultural training. VLEs have the potential to (1) promote positive movement through stage-based models of intercultural competence, possibly helping learners achieve greater levels of intercultural competence than traditional methods alone, and (2) act as an assessment tool to gauge movement through these stages.

### 2.4.3. EVALUATION AND ASSESSMENT

Computer-based training for intercultural competence is relatively new, but is beginning to enter more widespread use, especially with military audiences.

Next major steps are to evaluate more rigorously their effects on students' skills and attitudes, as well as to systematically compare their effectiveness against less expensive alternatives. Research with most of these systems has only just begun to look at fundamental questions of learning and acquiring intercultural competence. Experimental studies on these approaches suggest they may have a positive impact on knowledge: learners who take cross-cultural training are generally able to describe aspects of a different culture, accurately analyze behavior of someone from a different culture, and identify cultural differences (e.g., Ogan, Alevan, & Jones, 2010). Less of a general effect has been found on behavior, which looks at the problem solving ability and communicative choices of learners in intercultural settings (but see Hays, Lane, Auerbach, Core, Gomboc, & Rosenberg, 2009).

Evaluation of cultural skills (and language skills) is difficult and has been infrequently attempted with computer-based intercultural training. However, the research community could turn to the intercultural competence community who have developed a wide range of instruments to examine these questions (Paige, 2004). For example, the Intercultural Development Inventory focuses on determining one's place on a cultural worldview spectrum (from ethnocentric to ethnorelative) and has been validated extensively (Hammer, Bennett, & Wiseman, 2003). In general, there is no universal measure that determines the effectiveness of training programs; rather, the field has produced a large library of tools that seek to provide insights on how learners acquire cultural knowledge and intercultural skills. These instruments range from self-report questionnaires such as the Cultural Intelligence Scale (Ang et al., 2007) to choosing culturally appropriate explanations for events as in the Culture Assimilator (see Appendix B, Cushner & Brislin, 1995).

Other methodologies of evaluation have also been proposed. Johnson and Wu (2008) propose adoption-based research, which looks at the number of users and institutions choosing to use a system over time. Here, satisfaction is

implied by a growing user base, which may be more meaningful evidence than self-report data. User satisfaction outcomes, when checked, are also generally high (e.g., Raybourn, Roberts, Diller, & Dubow, 2008), but this comes as no surprise given the current novelty of such systems.

A recent meta-analysis suggests that many training programs have been found to be effective at teaching cultural knowledge and generating learner satisfaction, but generally fall short in skill acquisition and attitude change (Mendenhall, Stahl, Ehnert, Oddou, Osland, & Külmann, 2004). Determining appropriate methods for the evaluation of immersive cultural learning environments is a first step, although this will often depend on the context of use of the systems and most likely involves the use of multiple methods to triangulate students' developmental trajectories. Once the effectiveness of these environments has been shown, they have the potential to be used as testbeds for research into what leads to successful learning in these ill-defined, interpersonal domains. The next chapter reviews virtual learning environments including several systems designed specifically to teach skills of intercultural competence.

### 3. VIRTUAL LEARNING ENVIRONMENTS AND AGENTS FOR CULTURAL INSTRUCTION

According to situated learning research, virtual learning environments that allow students to immerse themselves in different socio-cultural contexts should provide good opportunities for cultural learning. Such virtual environments can also be augmented by embedding additional resources that can provide explanations about cultural specifics. High-fidelity graphics, sound, and animation make it possible for them to simulate many tangible aspects of a specific culture. The experience is greatly enriched by populating the virtual environment with embodied virtual agents that represent the local population.

Embodied agents are an emerging technology aimed at fostering human-computer interaction. Agents can exhibit a wide range of emotions, and can be built with underlying models of culture, personality, and affect (e.g., Gratch & Marsella, 2001; Cassell, 1999). They can produce varied verbal and non-verbal social cues; for example, they can perform gestures and body postures, both of which may be imbued with cultural meaning. Their abilities in simulating facets of human behavior endorse the development of embodied agents with cultural models for training purposes.

The chapter gives a summary of currently known VLEs for cultural training with adults, including BiLAT, the environment in which this dissertation work is set. The next sections describe cultural frameworks for virtual agent development and studies that have investigated user perceptions of such agents. These systems represent a growing trend recognizing the potential of immersive virtual environments for teaching social, interpersonal, and cultural domains.

### 3.1. VIRTUAL ENVIRONMENTS FOR CULTURAL INSTRUCTION

The broad appeal of games like *World of Warcraft*, *The Sims*, and *Second Life* show that immersive environments have become mainstream. Concurrently, their popularity as virtual learning environments has also increased. Some environments allow students to walk around a world and talk to its residents, such as *EcoMUVE* and *River City*, in which learners explore the physical environment and make scientific observations (Metcalf, Clarke, & Dede, 2009). Some are designed around quests or problem-solving adventures. *Crystal Island*, developed by Lester and colleagues, (McQuiggan, Rowe, Lee, & Lester, 2008), requires students to solve a mysterious outbreak of disease among a simulated island's residents. Bridging the commercial and academic worlds is *Whyville* ([www.whyville.net](http://www.whyville.net)), an online community where young learners can engage in an extensive set of educational activities to earn virtual currency. Given their natural focus on interactions with the agents who inhabit them, VLEs may be particularly well suited to providing instruction in social domains. For example, *FearNot!* is a project that helps children learn appropriate actions for dealing with bullying behaviors (Hall, Vala, Hall, et al., 2006). Patient-practitioner dialog is featured in a system that Hubal and colleagues (2000) have developed to allow medical students to practice interviewing patients.

Over the past several decades, intercultural training has also begun to take advantage of immersive technologies. Modern virtual systems often utilize a preexisting commercial technology (e.g., *Second Life*, *Unreal Tournament Engine*) to create a simulated representation of another culture complete with architectural features and ambient sounds. They typically integrate a set of embodied enculturated conversational agents (EECAs) who utilize a model of cultural behavior. Interaction with these agents facilitates the practice of communicative skills in the new culture, from making appropriate gestures of greeting to conversing in culturally appropriate ways. These systems cover a range of cultures (e.g., Spanish, Chinese, Iraqi, Dari, Pashto, and French), and

exist for various training purposes, ranging from language classrooms to military or business contexts. Systems can also be developed that go beyond national cultures. For instance, Rothwell suggests using culturally-aware educational technology for strengthening a cross-institution and cross-nation culture of nuclear safety (Rothwell, 2010).

Researchers now have the ability to combine these training systems with artificial intelligence-based scaffolding such as intelligent tutoring systems (ITS). For example, in a virtual environment, all of the students' communicative actions can be linked to a detailed representation of learning objectives which is managed by an ITS (Lane, Core, Gomboc, Karnavat, & Rosenberg, 2007). Such an ITS coaching component can provide guidance and feedback during face-to-face meetings with a virtual character from the target culture (Lane, Hays, Core, Gomboc, Forbell, et al., 2008) or provide an after-action review following each meeting. These systems may also integrate other learning activities, such as multimedia resources, quizzes, and part-task training exercises (e.g., see Second China below).

The next sections summarize six virtual cultural learning environments that have been developed for adult learners. They are analyzed in terms of their focus on knowledge, skills, and/or attitudes, their interactions, learning activities, and underlying models are described, and finally I report any empirical findings about their effectiveness.

### 3.1.1. ADAPTIVE THINKING AND LEADERSHIP SYSTEM (ATL)

The Adaptive Thinking and Leadership (ATL) system is a networked, 3D role-playing environment that focuses on teamwork, intercultural communication, and adaptive thinking (Raybourn, Deagle, Mendini, & Heneghan, 2005). The goals of the system are to improve the players' abilities to make decisions under stress, balance lethal and nonlethal aspects of conflict, and apply

communication and intercultural skills. Using a role-play paradigm, learners take the role of a U.S. Army Special Forces Soldier, an indigenous Iraqi citizen, or an invisible evaluator. A screenshot of characters communicating with one another in ATL appears on the left of Figure 3.1. A related system, Ambush! NK (“non-kinetic”), was developed with the addition of simulated crowds and scenario authoring tools (Raybourn, Roberts, Diller, & Dubow, 2008).

In ATL, careful attention is paid to tangible cultural elements, such as appearance, buildings, dress, and so on. Players interact with one another via speech and can move around freely in the virtual world. An instructor station allows for control over the simulation, such as the triggering of story events and the introduction of “curveballs” that create additional stress and surprise (such as an explosive device going off). The instructor also assigns evaluation tasks to observers so that they can assess the performance of their peers’ interactions within the virtual world. Typically, learners are rotated in and out of this role so that after a session, they have both role-played and evaluated the abilities of others.



Figure 3.1. Adaptive Thinking and Leadership System for intercultural communication and leadership training. (© 2006, Sandia National Laboratories)

Intercultural communication knowledge in ATL is derived primarily from scenarios and backstories that are given to the learners before they begin. As

students interact with one another, they are expected to take the perspective of their character, adopt appropriate desires and attitudes, and act in ways that they believe the fictional character would. ATL applies a perspective-taking approach, a well-established technique in the field of intercultural communication (Kramsch, 1993). Peer assessments in ATL can be used to rate participants in their ability to role-play in this fashion. A preliminary self-report study was conducted with 85 officers on the usability of the system and their perceived learning (Raybourn, et al., 2005). The results suggested that the officers felt the scenario was realistic and that they believed they learned about their own strengths and weaknesses by participating in the simulation.

### 3.1.2. CROQUELANDIA

Croquelandia is a VLE designed for learning Spanish pragmatics, as part of an instructional sequence with college learners. Pragmatics is defined as the various ways in which meaning is communicated and interpreted in interaction (Sykes, Wendland, & Moore, 2008). In particular, Croquelandia focuses on the skills of making culturally appropriate requests. Politeness in requests varies based on the social distance between participants, the size of the request, and other circumstances.



Figure 3.2. Croquelandia for second language learning. (© 2009, University of Minnesota)

Learners in Croquelandia interact in three virtual spaces: their host family's house, a central marketplace (seen on the left in Figure 3.2), and the office of a university professor. Players complete a series of quests related to specific requests that might arise in these contexts (an example appears on the right in Figure 3.2), such as requesting an extension on a paper, or asking the host father to hold a party. Completing these quests successfully requires interaction with one or more of the non-player characters (NPCs) present in each space. NPC interactions were created from role-play conversations with native speakers, which were developed into conversation trees from which the learner chooses. Additionally, learners can collaborate synchronously with NPCs or other students in their class using voice or written chat. Asynchronously, learners can create their own game content to share or leave messages for classmates on a discussion board. Assistance is provided in the form of various tips and ideas that are hidden throughout the environment.

A first evaluation study of the system was completed in the context of an advanced Spanish course. 25 students completed a discourse completion task (Cohen, Paige, Shively, Emert, & Hoff, 2005) and an interview, which was coded for instances of participants' perception of learning (Sykes, in preparation). Sykes reports that the data from the request scenarios show little change from pre- to posttest. Unsurprisingly, most participants showed a preference for communicative strategies from their native language. However, both the interview data and a set of in-class presentations indicate an improved awareness of pragmatic issues. Participants demonstrated meta-linguistic knowledge that might eventually transfer to production skills.

### 3.1.3. SECOND CHINA

The Second China project focuses on “preparing the learner to behave with a level of sophistication that communicates respect and understanding of the target culture” (Henderson, Fishwick, Fresh, Futterknecht, & Hamilton, 2008).

Its two components, accessed from a central web portal, work in tandem: a web-based text and multimedia repository, and a Second Life island that mimics cultural and visual aspects of China. Links from the web content go to relevant locations and scenarios in Second Life, while scattered throughout the 3D environment are links to relevant web-based readings and video (seen on the right in Figure 3.3). Thus the learner determines a personally relevant path through the content, while receiving guidance on appropriate routes to take. Government personnel conducting foreign operations are the target audience.

Entering the traditional web-based component first, learners encounter a set of modules that comprise the core curriculum. These modules, which consist of multimedia materials, learning activities, and self-paced quizzes, were developed by assessing critical aspects of culture for develop cultural competence. Exploring this content prior to entering Second Life, students may acquire target culture knowledge that will help them more deeply experience and practice behaviors in the 3D environment.

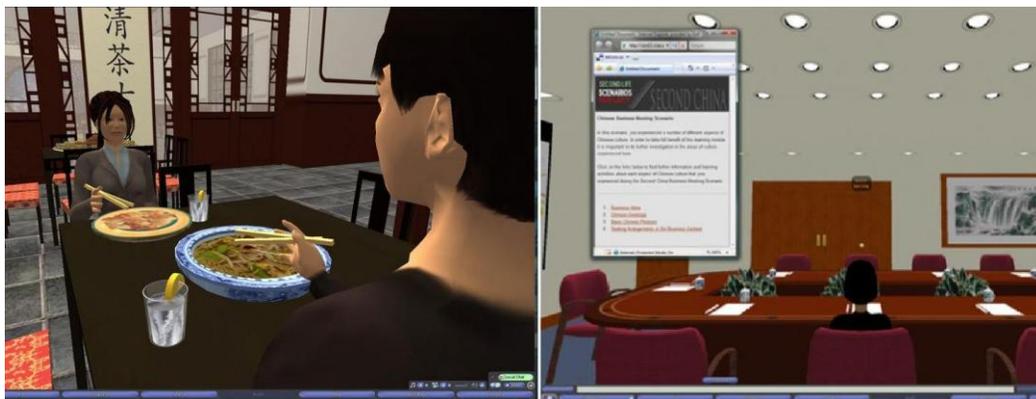


Figure 3.3. Second China for Chinese cultural learning. (© 2009, University of Florida)

Entering the Second China island first, the learner may accept an offer from Jiang, a guide character, or continue to explore the island independently. Exploration may take the form of observing or performing culturally significant activities (such as tai chi in a park), which use scripted animations. Learners

may also examine the architecture and other aspects of the 3D world, or participate in information quests. In addition to unguided exploration, embedded scenarios are facilitated by bots in culturally appropriate roles (such as a receptionist). Upon detecting the presence of a visitor, bots initiate a cultural experience, as seen in a restaurant on the left in Figure 3.3. To support learning, the learner receives just-in-time questions where the content either is necessary for the continuation of the scenario, or reinforces prior knowledge from the web repository. In this way, attention is drawn to critical cultural moments for those who choose a more experiential route.

Evaluation of the environment has begun in the form of a peer review committee comprised of experts in various related areas. Committee members were guided through Second China and participated in one of the learning scenarios described above. Henderson and colleagues report that initial feedback was positive, but highlighted the need for continued research and attention to instructional design, and defining the role of a virtual world in education. The suggested role for Second China was in “understanding how to participate in a culture” rather than simply learning about a culture.

#### 3.1.4. TACTICAL LANGUAGE AND CULTURE TRAINING SYSTEMS



Figure 3.4. Tactical Dari for Arabic language and culture learning. (© 2009, Alelo, Inc.)

The goal of the Tactical Language and Culture Training System (TLCTS) is to teach functional skills in various foreign languages and cultures (Johnson & Valente, 2008a). To date, four versions of TLCTS have been implemented: Iraqi, Dari, Pashto, and French. The emphasis is on spoken communication: learners must speak the foreign language to complete lessons and play games. Two views of the mission game in Tactical Dari are shown in Figure 3.4. Each instance of TLCTS provides three kinds of interactive activities:

1. Skill Builders provide part-task lessons that focus on core communication skills.
2. Arcade games provide opportunities for basic vocabulary practice, specifically with way-finding (both speaking and listening).
3. Mission games embody realistic practice and give the learner a chance to apply knowledge learned in the skill builders and arcade games.

Skill builders intentionally limit the context and allow the learner to focus on basic elements like vocabulary, pronunciation, grammar and nonverbal behaviors. Here, feedback is immediate and learners have the opportunity to retry until skills are mastered. Arcade games also provide limits by giving the learner simple goals, such as directing a character out of a maze with Arabic directional words. These skills are then put to the test in the mission game environment where the learner navigate through a high-fidelity 3D virtual world, interact with characters, and achieve more realistic goals, such as finding out names of important people in the town. In all interactions, the learner must communicate clearly in the target language in culturally appropriate ways (including non-verbal behaviors, selected via a drop-down menu).

Characters in the mission environment are driven by AI models of speech, language, and emotion (Johnson & Valente, 2008a). The speech recognizer is trained using novice data so errors can be better detected. Such errors are dealt with directly in the skill builder and arcade games, while they are integrated naturally into the mission game (e.g., characters will react with confusion when they don't understand you). Further, using game performance

and quiz results as evidence, TLCTS maintains a probabilistic student model that tracks learning throughout use of the system. TLCTS also includes authoring tools that allow non-experts to create game content (Johnson & Valente, 2008b).

Several evaluations of TLCTS have been conducted. A broad study on the effectiveness of Tactical Iraqi on military participants reported significant gains in learning of Iraqi Arabic and cultural knowledge from pre- to post-usage of the system (Surface & Dierdorff, 2007). Tactical Iraqi was also used to investigate the “politeness effect” (Wang & Johnson, 2008); it was found that feedback messages that helped the learner “save face” led to better learning than less polite messages on multiple-choice and matching quizzes. Finally, based on in-game performance data, military participants with high motivation and time in service, as well as those who spent the most time with the Skill Builder tool, demonstrated the largest learning gains (Johnson & Wu, 2008).

### 3.1.5. VIRTUAL ENVIRONMENT CULTURAL TRAINING FOR OPERATIONAL READINESS



Figure 3.5. VECTOR system for intercultural communication and peacekeeping training.  
(© 2008, CHI Systems, Inc.)

The Virtual Environment Cultural Training for Operational Readiness system (VECTOR) shares many similarities with Tactical Iraqi and BiLAT. The goals of the system are to improve learners’ knowledge of Arabic culture (specifically,

the Kurdish sub-culture of West Asian Arabs), although the system has an authoring system specifically designed to extend the system to new cultural contexts. Like Tactical Iraqi, the focus is on face-to-face intercultural communication skills, but with an emphasis on peacekeeping (Deaton, Barba, Santarelli, Rosenzweig, Souders, et al., 2005). A screenshot of a face-to-face encounter is shown on the left side of Figure 3.5. VECTOR situates learners in a virtual foreign town working as a Military Policeman or with Military Intelligence. Learners communicate with characters through contextual, menu-based selections and characters respond using synthesized speech. To succeed in missions, the learner must navigate the country, interact with locals, and take actions to solve problems. An example of a game goal in VECTOR is to find the identity of a bomber and stop him from attacking his next target.

VECTOR scenarios can be created with a set of associated authoring tools that are intended to permit non-experts to create game data (Barba, Santarelli, Glenn, Bogert, & Belanich, 2006). Cultural rules are derived from common superclasses that encode cultural norms, and can be tweaked such that they react differently according to personality differences. Emotional modeling plays an important role in VECTOR – the emotional state of NPCs changes based on what happens to them. Actions taken by the user are one of the primary influences on NPC emotions, as well as interactions that happen “off camera” with other NPCs (Barba, Deaton, Santarelli, Knerr, Singer, & Belanich, 2005). This means that a meeting with one character may negatively or positively influence a different character’s attitude towards the learner. VECTOR includes a “synthetic instructor” character that can provide proactive guidance during the game and conduct an after-action review with the learner. Correct answers are reinforced with explanations for why the actions had a positive effect on the characters, whereas mistakes inform the learner what actions would have been more appropriate. Actions are assessed by the synthetic instructor by inspecting their relative impact on the emotional states of the NPCs.

No empirical evaluation of VECTOR or its authoring system has been published at the time of this writing.

### 3.1.6. BiLAT

BiLAT is a game-based immersive environment that teaches the preparation, execution, and understanding of intercultural bilateral meetings (Kim, Hill, Durlach, Lane, Forbell, Core, et al., 2010), and is the context for the work in this dissertation. The BiLAT architecture is built on Unreal Engine 2.5 and integrates research technologies such as virtual agents and intelligent tutoring support. The focus is on both knowledge and skills in Arab culture, but with emphasis on the culture of business meetings and negotiation skills. Unlike TLCTS, there is no coverage of the Arabic language. BiLAT is supported by a video that depicts good and bad examples of intercultural meetings, and provides a summary of the learning objectives addressed by BiLAT. Most scenarios in the simulation place the learner in the position of a high-ranking U.S. Army officer who must solve a peace-keeping or rebuilding mission in a small Iraqi community, although other scenarios have been authored, such as for training Iraqi recruits and conducting searches in Afghanistan.

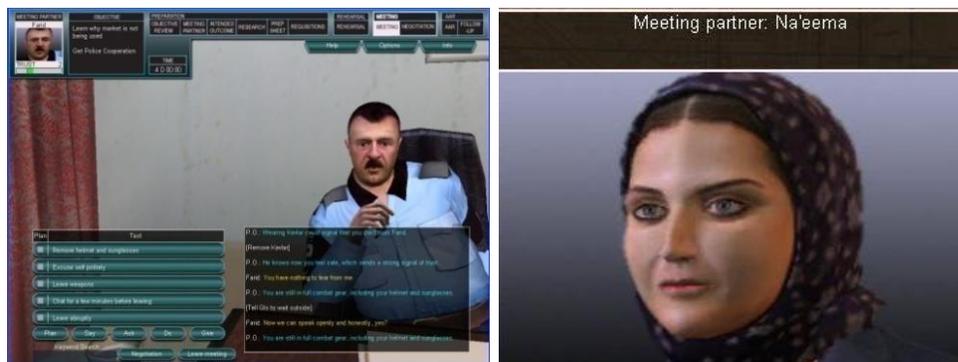


Figure 3.6. At left, a meeting in BiLAT with police captain Farid with the goal of solving a problem with a market in an Iraqi town. At right, meeting partner Na'eema, a doctor. (© 2009, University of Southern California)

## 3.1.6.1. BILAT GAMEPLAY

Scenarios derived from real-world situations encountered in Iraq drive the game experience. For example, the Market Scenario that is used extensively in this dissertation takes place in a town with a market that was recently built by the United States. The market is being avoided by local businessmen for unknown reasons, and the learner must figure out why. First, learners must study the background story for a scenario, review their objectives, and select a character with whom to meet who can provide important information and assist in solving the problem. In the Market scenario, the first character to meet is a local police officer named Farid (seen in Figure 3.6).

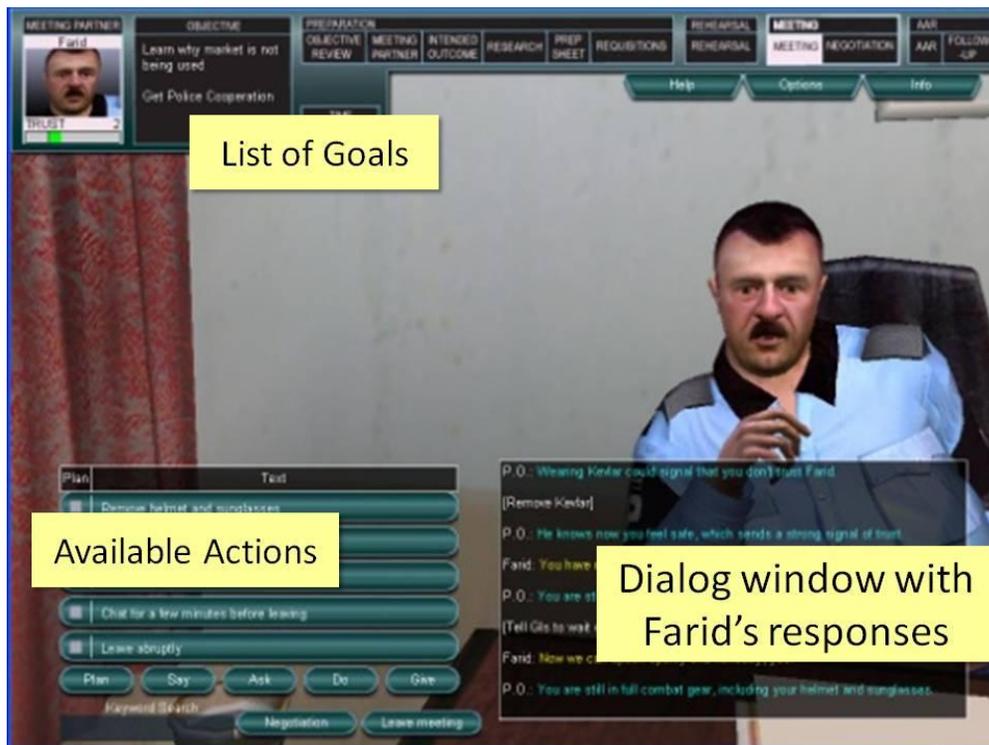


Figure 3.7. Annotations describing the various parts of the interface in BiLAT.

Next, learners then conduct background research in the Prep Room on those characters to learn about their particular likes, dislikes, needs, and desires. For instance, Farid, the police captain, is a family man who is concerned about the

state of the local school, but holds an illegal weekly poker game. Players can store this information in a meeting preparation sheet that identifies these important pieces of information, and includes goals for the meeting, potential impasses, and more. This information is provided by a variety of “sources” of varying degrees of trustworthiness, which simulates gathering such information in a real world situation. During preparation, concrete, negotiation task-related goals for the meeting are presented, which learners must select before moving forward. The set of objectives for the initial meeting in the Market Scenario is, “Learn why the market is not being used” and “Gain police cooperation.”

Once research is complete, the learner is sent to meet with the character. A screenshot of the BiLAT meeting screen appears in Figure 3.7. BiLAT uses menu-based interactions for communication with the virtual human characters. The set of available actions depends in part on the quality of the background research conducted by the learner for each character. In addition to the meeting actions available in the menus, the virtual character can initiate dialogs in a format called a “challenge network”. In these challenge networks, the character initiates a dialog and the player may be required to reply from the standard menu of meeting actions, or a special limited set of responses. For instance, if the player attempts to flatter one character three times, a challenge network might be initiated in which the character complains that the compliments have lost meaning.

The player receives feedback on their actions in three ways. First through the character’s responses, which are played through a speech synthesizer as well as posted in a dialog window. Non-verbal behaviors are animated during the response and correspond to the content of the character utterances. BiLAT also maintains a “trust meter” that is updated after each turn – the amount of trust the character currently has in the player is modeled based on the player’s actions, and is displayed in the interface. This scaffolds the player in both monitoring their success and learning what actions may hurt and help achieve

their objectives. Finally, an intelligent tutoring system, described in the next section, provides guidance that fades over time as students master the content.

Once the meeting objectives have been obtained, the meeting is complete and the player will receive the next set of objectives for the scenario and view new characters who might help them achieve those objectives. It is also possible to have an unsuccessful meeting (e.g. by offending the host), in which case the character asks the learner to leave and return for another meeting to try again. When the scenario is complete and the mission is solved, learners may select another scenario with new objectives.

#### 3.1.6.2. LEARNING OBJECTIVES AND ADAPTIVE GUIDANCE

Success in BiLAT requires the application of a number of different skills. Adherence to Iraqi cultural and social interaction rules is necessary, as well as the application of integrative negotiation tactics. A primary learning objective is to consider their counterpart's interests such that a "win-win" result can be realized, as described in "Getting to Yes" (Fischer, Uri, & Patton, 1981). While negotiation is the context for gameplay in BiLAT, negotiation in a larger sense involves conducting meetings, setting up subsequent meetings, following up on promises, etc. This dissertation focuses on the cross-cultural issues that surround this negotiation. The aim of both BiLAT and the work presented here is not simply to guide students towards obtaining the most lucrative negotiation outcomes, but rather to successfully build lasting relationships through cultural understanding. "The Handbook of Negotiation and Culture" describes how an understanding of the impact of culture is central to an understanding of intercultural negotiation (Gelfand & Brett, 2004).

BiLAT was designed to address the specific cultural knowledge and skills that support more effective negotiations in the Iraqi culture. By rewarding students who take actions that conform to appropriate cultural behaviors, BiLAT

addresses cultural concepts like polychronicity (differences in the meaning of time), the value of social relationships and building trust, and face-saving, along with basic etiquette concepts like greetings, giving gifts, and leave-taking (see Nydell (2006) for a more complete description of Iraqi cultural specifics).

Cultural knowledge is encoded in several forms in BiLAT. All meeting content is linked to a detailed representation of learning objectives which is managed by an intelligent tutoring system (Lane, Core, Gomboc, Karnavat, & Rosenberg, 2007). Table 3.1 shows an example of a learning objective concerning starting the meeting with small talk. The knowledge representation is then broken down into a number of related standards which contribute to achieving this objective. Given the ill-defined, non-procedural nature of the domain, even experts may disagree about when or why some actions are appropriate in different contexts. The knowledge representation deals with this difficulty in part by providing conditions under which the objective should be applied, and also by labeling standards with how imperative they are: “required”, “rule of thumb”, “usual”, or “avoid”. The cultural elements of these user actions and character responses have defined effects on trust, power, and other character attributes.

Additionally, the characters each have meeting phase encodings that define typical Iraqi business meeting etiquette. The phases involved in these meetings are labeled *opening* (involving greetings and the like), *pre-business* (the social, relationship-building phase), *business* (where the negotiation happens), and *closing* (relationship-building continues to set the stage for the next meeting). Most meeting actions are only appropriate in a limited number of phases. For instance, “Greet in Arabic” should only be done in the opening stages of a meeting. These restrictions are recorded in the learning objective representation.

The ITS coaching component uses this representation to provide guidance and feedback during face-to-face meetings in BiLAT (Lane, Hays, Core, Gomboc,

Forbell, et al., 2008) as well as an after-action review for each meeting (Lane et al., 2007). All communicative actions taken in a meeting are assessed based on the meeting context, the related learning objectives, and the trust changes they induce. This process categorizes actions as having a positive, mixed, or negative impact on the character. The coach uses these to decide whether to give prescriptive feedback (i.e., hints) and/or reactive feedback (i.e., about an action just taken). Both hints and feedback are delivered to learners in the dialog window, presented by a character called the P.O. (“process observer”), a role taken from a similar position in the Army.

Training objective	Start a meeting with a social period
Short name	small talk
Condition	while meeting with a partner with whom you are trying to build trust
Standards	<ol style="list-style-type: none"> <li>1. (REQ) expose yourself (do not overly protect)             <ol style="list-style-type: none"> <li>a. (USL) remove protection (e.g., weapons)</li> <li>b. (REQ) remove coverings such as sunglasses and helmet</li> </ol> </li> <li>2. (REQ) begin with greetings and introductions             <ol style="list-style-type: none"> <li>a. (REQ) greet in Arabic</li> <li>b. (REQ) conduct formal introductions</li> </ol> </li> <li>...</li> </ol>

Table 3.1. A sample learning objective from BiLAT. REQ indicates a required action, USL indicates a usual action.

### 3.1.6.3. VIRTUAL AGENTS IN BiLAT

Within the scenarios presented by BiLAT exist a large number of EECAs, with whom the learner interacts to practice intercultural negotiation skills. These agents simulate members of the Iraqi culture with different roles, e.g. doctor or

merchant. In the case of the Market Scenario, the first character learners meet is Farid, the police captain in the town (seen on the left in Figure 3.6).

Learners interact with each BiLAT character by selecting from a menu of communicative actions that includes greetings, small talk, task-related dialog, and more. With each action taken by the learner, the agent updates their model of trust in the learner; in general, culturally appropriate actions increase trust, while inappropriate actions decrease trust. Personality also influences the appropriateness of particular actions taken by the learner; e.g., characters may be more or less likely to respond negatively to suggestions that contradict religious teachings, such as tasting alcohol.

Agents respond to players' actions in both text and synthesized speech as well as culturally-appropriate non-verbal behaviors such as gaze and gesture. Guiding each virtual character's gestures, gaze, and posture is an animation engine called SmartBody (Lee & Marsella, 2006). Verbal responses depend on a number of factors, including the agents' personality, the current meeting phase, his or her current level of trust in the learners, and a virtual dice roll. The dice roll is intended to simulate uncertainty in human behavior. Finally, the Psychsim social simulation (Pynadath & Marsella, 2005) determines negotiation decisions during the sub-phase of the meeting where business terms are being offered and requested. The dialog utterances for each agent were written by a team of game developers after consultation with subject matter experts (see Hill, Belanich, Lane, Core, et al., (2006) for a more complete description of the development of these agents).

#### 3.1.6.4. PREVIOUS EVALUATIONS

BiLAT is one of the very few cultural training systems on which evaluations using assessments external to the system have been conducted. In general, these studies have focused on learner reactions, overall learning from the

system, or on manipulations to the intelligent tutoring system. At the time of writing, no other studies in any of these systems have investigated learners' relationships with the agents or attempted to manipulate learners' social orientation as this dissertation does. Nor have they been compared to typical classroom activities for teaching the same skills. Examples of prior studies with BiLAT include:

Whether BiLAT produced learning gains and for which kinds of learners: Durlach, Wansbury, & Wilkinson (2008) ran a study with U.S. Army officers and found that learners with minimal prior negotiation experience did demonstrate learning gains on a Situational Judgment Test (see Appendix B), while those with prior experience did not. The authors note that it was possible not enough time was devoted to training, suggesting that users with prior experience may have just needed more time to use the system in order to see a benefit. In addition, it was also reported that learners benefited regardless of their self-reported video game-playing habits.

The effect of coaching during meetings in BiLAT: In one experiment, feedback from the ITS during meetings was found to be effective at helping learners understand Arab business meeting expectations as they relate to time and "following the lead of your host" (Lane, Hays, Core, Gomboc, Forbell, et al., 2008). Specifically, feedback that focused on the expectations of their meeting partner (e.g., "You should not be talking about business topics now.") translated into better in-game performance on later meetings with characters with no coaching, as well as on related questions on a written test. A subsequent study compared conceptual feedback with very concrete feedback, and this again produced better in-game transfer (Hays, Lane, Auerbach, Core, Gomboc, et al., 2009).

Most of the environments described in this section are still in developmental stages and research efforts have focused on design and integration of interactive technologies rather than empirical evaluation of learning. Therefore,

one avenue for future research is in the evaluation of such systems for student learning, compared to either typical classroom approaches or the gold standard of one-on-one human tutoring of such skills. Additionally, it is not yet understood how people treat, react to, and learn from the enculturated virtual humans in these environments. Studies of learning from BiLAT in this dissertation fit into the broader movement to understand how people treat, react to, and learn from virtual humans.

## 3.2. EMBODIED ENCULTURATED CONVERSATIONAL AGENTS

Perhaps the most promising feature of these virtual learning environments for intercultural training is the development of Embodied Enculturated Conversational Agents (EECA) (Rehm, 2010). This concept initially stemmed from Embodied Conversational Agents (ECA), agents able to communicate with users through the production of verbal responses, gestures and postures of their virtual body (Cassell, 1999). Since verbal and nonverbal behavior is known to differ greatly from one cultural group to another (Bonvillain, 2008), cultural considerations must be taken into account when designing agents for these environments. Developing effective EECA requires mastery of many issues such as the genesis of realistic 3D behaviors and communication styles, the computationalization and integration of cultural competences, and the consideration of cognitive and affective implications (Rehm, 2010). The next two sections describe current research efforts into cultural frameworks for EECA, and studies investigating user perceptions of agents.

### 3.2.1. FRAMEWORKS FOR CULTURALLY-AWARE VIRTUAL AGENTS

In the past few years, several frameworks have been proposed for infusing culture into virtual agents. Although not necessarily designed for educational

use, computer-assisted educational technology is a natural target of these generic approaches to culture. Some of these initiatives are reported here.

Based on a formal analysis of the domain, Blanchard and colleagues (2010) have used heavyweight ontology-engineering techniques to propose the Upper Ontology of Culture (UOC). The UOC is a theory-driven conceptualization of the cultural domain including cognitive, affective, behavioral, contextual, and physical dimensions (see Figure 3.8 for the ontological conceptualization of “culture” in the UOC). It provides guidelines for the development of culturally-aware applications, the consistent computerization of cultural data, as well as the development of culture-driven automatic reasoning processes. Focusing more specifically on multi-agent systems, Birukou and colleagues have proposed the Implicit Culture Framework (ICF) (2010) to formalize how cultural knowledge transfers between community members or between communities.



Figure 3.8. The ontological conceptualization of the “culture” concept in the UOC; ‘p/o’ refers to part-of links, and ‘a/o’ refers to attribute links. See Blanchard et al. (2010) for an in-depth explanation of this conceptualization and the ecology of concepts into which it is integrated.

Whereas the UOC and ICF frameworks describe relations between culture and other human factors of interest for educational purposes (e.g., affect and cognitive processing), other researchers have developed models that have been implemented into agent systems. In the past few years, a number of projects with very different research foci have emerged in this area. For instance, Nazir and his colleagues have proposed a model that merges cultural, affective, and personality features (2009). This model refers to Hofstede's system of value (2001) for its cultural component, the PSI model (Doerner, 2003) for affect, and the Big Five model (McCrae & Costa, 1996) for its personality component. Based on a qualitative evaluation, Aylett and colleagues (2009) reported that integrating this model into embodied agents populating a role-playing environment could increase children's intercultural empathy. The CUBE-G system (Rehm, André, Bee, Endrass, Wissner, Nakano, Nishida, & Huang, 2007) also suggests the use of Hofstede's dimensions in combination with empirical data to support the development of culturally-appropriate nonverbal agent behaviors. In a separate initiative, Huang and colleagues (2009) have proposed the Generic Embodied Conversational Agent (GECA) framework in order to speed the development of EECA. Using GECA, only a module describing verbal and non-verbal communication specifics of a targeted group has to be developed in order to provide cultural intelligence to an ECA. This concept was showcased in an application where an EECA played the role of a culturally-intelligent tour guide for serving Japanese, Croatian, and general Western users by displaying appropriate verbal and non-verbal behaviors.

In support of cultural models, modeling universalisms in communication are also worth consideration. For example, Brown and Levinson's theory of universal politeness (1987) has been applied to EECAs (Johnson, Mayer, & André, 2005), and Miller and colleagues recently proposed a formalized computational model for agents (Miller, Ott, Wu, & Vakili, 2010) that further facilitates the integration of this politeness theory. Although not currently being

employed in the service of EECAs, Cassell and Bickmore (2003) have developed a user-agent social linguistic model based on the dimensions of the universal ‘interpersonal relations in conversation’ model developed by Svennevig (1999). These dimensions are theorized to have an effect on collaborative activity and trust, which are important considerations in learning.

### 3.2.2. USER PERCEPTIONS OF VIRTUAL AGENTS

With a framework for EECAs in place, researchers have begun to study user perceptions of these agents. For instance, for training purposes, it is important that learners are able to perceive cultural differences in their behaviors. Endrass and her colleagues proposed a system where EECAs were attributed culturally-marked communication styles with varying usages of pauses and overlapping speech (Endrass, Rehm, & André, 2010). They found that, even if the fantasy language EECAs used to communicate with each other was unknown to human observers, they perceived the agents as having a Western or Asian orientation depending on the agent’s communication style. Furthermore, in a preliminary evaluation, observers reported preferring agents with a communication style similar to their own.

Similarly, Mascarenhas and colleagues have developed scenarios in which agents act within the principled parameters of simulated cultures. In one study, two groups of agents acted according to individualistic or collectivistic rules for behavior. Users were able to successfully distinguish between the two groups in what values they held, but believed that the differences were due to personality rather than culture (Mascarenhas, 2009). In a second study, the same agents were designed to follow rituals based on either high power distance or low power distance rules (Hofstede, 2010) for sitting down to eat dinner. In this case, users identified the rituals as belonging to separate cultures, but could not identify different values for the two cultures. These examples illustrate some of the difficulties associated with instruction in cultural domains;

the values associated with cultural dimensions are often unconscious and therefore are harder to interpret as cultural by learners than more overt rituals or symbols.

Imbuing virtual agents with cultural traits can successfully enable learners to perceive cultural differences, but may lead to unexpected negative consequences. Baylor and Kim (2004) reported variation in perceived competence of pedagogical agents according to their ethnicity. This effect may be due to the expression of stereotypes and unconscious mental preconditioning within a cultural group. On a positive note, research suggests that showing more virtual experts with characteristics from minorities could help to reverse negative mental programming (Yee & Bailenson, 2006).

In addition to using cultural specifics or universals as the model for agent behaviors, researchers have explored an individualist approach by mirroring users' own verbal and non-verbal behaviors. Gratch and colleagues (2007) report that users feel similar feelings of rapport with a virtual agent who mimics their head gestures and posture as with a human listener. Bailenson and Yee (2005) also found increases in the persuasiveness of virtual agents based on mirroring head motions. These strategies may be more or less successful based on user characteristics; Burlison and Picard (2007) found that gender affected student perceptions in a similar mirroring intervention. These and other findings are beginning to establish that models that drive agents' behaviors can have significant effects on users' perceptions, affect, and cognitions.

Since Embodied Pedagogical Agents (EPA) (Rickel & Johnson, 1997) are ECAs with additional intelligent tutoring capabilities, it is reasonable to posit that improving their ability to efficiently communicate would improve the quality of their relation with learners and their overall efficiency. EPAs with cultural models have already been implemented into several systems for intercultural competence instruction (Johnson, 2007; Kim et al., 2010).

## 4. STUDY I: EXPERT VALIDATION

During the development of BiLAT, the virtual learning environment utilized in this dissertation, subject matter experts (SMEs) were consulted to produce compelling and realistic game scenarios and content. In particular, game designers and educational researchers collaborated with two expert groups: Iraqi nationals and Army officers who conducted extensive cross-cultural negotiation while deployed in Iraq. These efforts contributed to iterative design of the system by providing feedback on demonstrations of system prototypes. Following development, however, no qualitative data had been gathered from SMEs that documented their experiences while they played the completed game. Prior to studying learning and social engagement with BiLAT, I believed it was important to take steps towards validating the content of the virtual learning environment with experts.

Simulation validation, particularly when dealing with social systems, is not a simple procedure; in fact, a tool that uses knowledge and ontological representation to assist in this endeavor was the subject of a recent doctoral dissertation (Yahja, 2006). The main issue to address in this work, instead, is that the content looks and feels realistic to those who have been in intercultural experiences in Iraq, therefore suggesting its applicability to training. In addition, I was interested in exploring how experts engaged with the virtual learning environment: What are expert goals as they approached the task of intercultural negotiation? What are their preconceptions? Do they interact with training agents in meaningful, social ways? To this end, I conducted

think-aloud protocols with experts from both the military and Iraq as they interacted with the virtual characters.

#### 4.1. ARMY SUBJECT MATTER EXPERT INTERVIEWS

The U.S. Army rank of Captain is a commissioned officer rank that typically corresponds to the command of a company of between seventy-five to two hundred soldiers. While deployed overseas, Captains are required to interact with local leaders and other citizens, negotiating to conduct projects in the area while generating goodwill and trust. Therefore, soldiers at the rank of Captain are expected to have previously encountered exactly the type of situation that is covered in the BiLAT environment. At Ft. Riley, Kansas, I conducted interviews with three Army Captains who had each been deployed several times.

##### 4.1.1. PARTICIPANTS

All three participants had been recommended by their commanding officer as successful negotiators in an Iraqi cultural context. None of them had been involved in development of BiLAT, nor had they had any prior knowledge of the game. Prior to the situated interview, each participant took a demographic survey.

Participant 1 was male, age 31, and an Army captain. He had been deployed to Iraq, Afghanistan, Korea, and Europe and conducted bilateral negotiations while deployed. While he had only had some formal training in negotiation, he considered himself a competent negotiator and more than competent at knowledge of Middle Eastern cultures. He had lived outside of the United States for more than a year while deployed, but spoke only English. He rarely played video games and did not consider himself a gamer, but those he did play were first-person shooters. He did not have any programming experience.

Participant 2 was male, age 34, and an Army captain. He had been deployed to Iraq, Afghanistan, and other theaters, and had conducted bilateral negotiations while deployed. While he had had little formal training in negotiation, he considered himself a slightly more than competent negotiator but slightly less than competent at knowledge of Middle Eastern cultures. He had lived outside of the United States for more than a year and spoke English and Spanish. He rarely played video games and did not consider himself a gamer, but those he did play were first-person shooters or Wii sports games. He did not have any programming experience.

Participant 3 was male, age 29, and an Army captain. He had been deployed to Iraq and had conducted bilateral negotiations while deployed. While he had had no formal training in negotiation, he considered himself a competent negotiator and competent at knowledge of Middle Eastern cultures. He had additionally been outside of the United States on vacation and spoke only English. He rarely played video games and did not consider himself a gamer, but those he did play were sports games. He did not have any programming experience.

#### 4.1.2. METHOD

The interviews were conducted individually on the base, each lasting approximately one hour and all taking place within the span of one day. Participants were told that the purpose of the interview was to evaluate BiLAT for authenticity, as well as to gather their experiences in intercultural negotiation abroad for use in future design. In order to facilitate data collection and elicit situated responses, the participants played BiLAT while they discussed these experiences. I prompted them to speak while they played but did not interfere when they encountered difficulties, instead choosing to observe how they resolved the problem. If they became too frustrated, they were moved to another character so that they experienced some variety within the game. Following the goals of the interview, these participants tended to engage in

meta-talk about their experiences as they related to the current context in BiLAT. The verbal protocol was recorded and the data analyzed for major themes.

#### 4.1.3. OBSERVATIONS

None of the three Captains had received negotiation training before they were deployed. However, each of them felt that they did not have trouble adjusting to the environment when they arrived abroad.

P2: "I was in Afghanistan in 2004, and that's when I was just... here you go! Meet with the tribal leaders every week!"

P3: "I never had any training in any of this, and I never had a problem. I just knew not to put my feet up in front of him, and there were certain things I couldn't talk about and that was it."

Once in Iraq, negotiations often started out with a purely social meeting that allowed their meeting counterparts to begin building a relationship.

P2: "I know in Iraq, they kind of felt me out, because the leaders came over to my COP, and they asked to sit down with me and we just talked about family, family and religion, and that was my first meeting with the guys that I worked with over there in Iraq. No business, just family and religion. And politics of course."

In these social meetings as well as in the subsequent negotiations, participants believed that, in general, the abilities to perform this job are either innate (you are personable or you are not) or require a few general learnable skills such as showing respect:

P1: "... and if you are comfortable interacting with people and you know how to communicate regardless of their culture you just know how to open up the conversation - keeping in mind the cultural norms for Iraqis and Arabs ... you just get it after you practice a few times"

P3: "I always thought this sort of negotiation training was overblown. You're either a jackass or you're not. And certain people just don't do that well, no

matter how hard you try to train somebody on negotiating and talking, some people just aren't any good at it and they never will be. Some people just aren't personable."

P2: "The biggest thing is, when you get in there, you show respect, I think a lot of the times... I think watching the [BiLAT preparatory] videos, with Captain Miller, it was a lack of respect... But my thing was, just treating everyone with respect."

Participants did acknowledge that cultural differences exist and must be accounted for in the negotiation.

P1: "you do want to know as much as you can about the guy before you go talk to him, at least knowing what his motivations are and what his reputation is, but at the end of the day he's still Iraqi, and being Iraqi, the Iraqi factor is always an effect."

They believed that success was also dependent on being able to constantly reevaluate the current situation and take into consideration both the meeting partner and the external circumstances.

P1: "... it's just such a dynamic situation and every situation is different ... If there's a lack of security, everyone in the room acknowledges that small talk or social issues are not that important. But the security situation drives a lot of that."

P3: "Everything depends on your meeting partner, who may be slow or fast, business or social, and is usually trying to get something from you."

However, even though the situation required constant monitoring and adapting, while abroad each captain had developed a particular way of opening meetings and conducting a social meeting phase that he felt was repeatedly successful for him.

P1: "... so what I'm looking for is, what I typically do when I enter into a meeting, is to do the social interaction."

"... since I don't know where everything is at [in BiLAT], I'm having to navigate through it for the first time, so that's the challenge, because I know the order

in which I'd want to start a meeting, but I'm looking for stuff and not finding it right away"

"Always current events are good, if there's any type of national soccer game on that gets everyone's interest, the Iraqi team is very popular, so that's something you can always find common ground on."

Participant 2 walked through the opening stages of BiLAT while verbalizing his typical routine:

P2: "Basically I would start with a brief introduction, always saying Assalamu-alaikum ["Peace be upon you"], shaking the hand... is there a greeting?"

And then I would take off my helmet and depending on the surrounding area I would remove my IBA [Individual Body Armor].

Is it just Farid and I? <leaves weapons>

Usually I would try to get involved in, show him I am interested in the cultural aspect first, but he says he doesn't have time so...

Since he seems like he's in a hurry, I would talk to him about community problems."

At first, each participant had difficulty accomplishing goals in the game. They all claimed this was due to the unavailability in the game of their preferred (routine) methods of developing social relationships.

P1: "... it really is a good tool, and you just have to learn what the game is requiring here for the correct sequence of responses, because it's hard to pick a correct answer when real life interactions are so fluid and change based on the current conditions on the ground."

"... so if I sat down and went through this and figured out what the approved solution is, cause I do know what the right answers are I just don't know what the sequence of the game and all that"

P3: "See I would have asked him why people are still using the old market. But that question isn't there.

Punch him in the face!"

This preceding quote provides evidence that some learners (even those with prior experience) hold a “gamer mentality”, or a non-social approach to gameplay that ignores the “feelings” and “desires” of the virtual agents involved. While Participants 1 and 2 treated every interaction with the characters as if they were in a serious negotiation, Participant 3 continued,

P3: “There’s no punch the guy in the face! I wouldn’t really add that, because somebody like me would actually just do it 5 times just to see the reaction.”

An additional finding was that while they had successfully accomplished negotiation tasks in Iraq, they did not tend to have favorable opinions about Iraqis:

P1: “They’re extremely hypocritical. A lot of the Iraqi men are. But they don’t want to lose face. So they want to look good and they want to say the right things and do the right things around those they think are watching or care but the reality of the situation might be something very different.”

“once you kind of crawl into the Iraqi mind and you understand where they’re coming from and that they want one of two things, money or power, and it may or may not be in that order. Even the patriots still want money and power - the guys you think are uncorruptable [sic], that have great integrity, that’s their culture. That’s the way they’ve been.”

When asked about the training potential of a tool like BiLAT, all participants thought that the game would be more useful for complete novices, but possibly helpful as a refresher before being deployed again.

P1: “This is probably a better tool for guys that haven’t done it before, to teach them some of the basics, in that it would have been a valuable tool to have seen this before I went to Iraq the first time.”

P2: “I think it would be more useful for someone who hasn’t gone over before, you know like good training, and maybe throw in some language, actually Arabic language, and I know over in Iraq they love when you try to speak the language.”

In general, participants were impressed with the scope and detail of the environment, and would recommend it for use with novices. Conclusions and implications of this work are discussed in Section 4.3.

## 4.2. IRAQI SUBJECT MATTER EXPERT INTERVIEWS

The second type of subject matter expert who had input into the development of the game was Iraqi nationals. In addition to determining that the game felt congruent to the experiences of Army negotiators, I wished to evaluate whether the content of the Iraqi cultural model incorporated into the game felt accurate to these experts. To this end, I traveled to El Cajon in San Diego County, California, to conduct interviews with four Iraqi women. Including the second generation, there are an estimated 35,000 Iraqis living in San Diego County, making it the second-largest Iraqi community in the United States (the largest being Chicago). El Cajon is a town outside of San Diego with a thriving community of recent Iraqi immigrants, who have established a Little Baghdad in the center of town.

### 4.2.1. PARTICIPANTS

All four participants were women. Selecting a specific participant gender, while potentially generating a less complete understanding of the simulations' realism, facilitated later agent development and is more fully described in Chapter 7. Three of the women were immigrants to the United States and in their 50s. They had lived in Iraq for the majority of their lives, and now lived in a community of Iraqi ex-patriots residing in California. One of their stated roles was as part of the "welcoming committee" who brought newly arrived Iraqi citizens into the community. One of these three participants worked as a professional role-player in a simulated Iraqi village in the U.S., where Army soldiers train in negotiation with Iraqis prior to deployment (see Figure 4.1). As

part of this employment, she spent weeks at a time living in the Army-built town, interacting with soldiers in a manner similar to that which BiLAT intends to simulate.

The fourth participant was 24, and was the daughter of one of the older participants. She had never lived in Iraq, but was part of the ex-patriot community with her mother. As the other three participants spoke varying levels of English, she translated when necessary.



Figure 4.1: Iraqi national playing BiLAT.

#### 4.2.2. METHOD

A group informal interview was conducted with the four participants outside at a coffee shop in the Little Baghdad area of El Cajon, in front of a laptop computer. Participants were asked to read about the game scenario, meet with Farid, and respond to a survey regarding the cultural accuracy and realism of the agent and his behaviors. This survey used a Likert scale of 1 (“Not true at all”) to 5 (“Very true”). Due to their relative unfamiliarity with computers, I modified the intended procedure so that I controlled the game interface as

they stated what they would like to do next. The main goals for the interviews were to understand whether the verbal and non-verbal behaviors displayed by the agents in BiLAT felt authentic and appropriate, and determine whether there was anything in the game that they believed was offensive to natives of the Iraqi culture.

#### 4.2.3. OBSERVATIONS

Participants confirmed that Farid displayed behavior authentic to Iraqi culture, in particular noting that the dialog felt appropriate and realistic. According to one participant, the smalltalk and social conversation incorporated into Farid's character were "very accurate". Accuracy in dialog was rated a 5 (very true). Participants also thought non-verbal behaviors were realistic – rated a 5 - including arm movements and even Farid's seated posture. They also readily ascribed a disposition to the character, noting, "[Farid] seems mad." Overall, they concluded, "he's exactly the same as an Iraqi policeman" and rated the authenticity of his behavior as a 4 out of 5.

Indeed, Farid appeared so realistic that he evoked an instinctive emotional response. One of the first things Farid does after greetings are exchanged is offer tea, as a social custom, and I had expected that the group of women would accept his offer. Instead, as soon as one participant saw Farid, she appeared visibly upset, and refused the tea, thinking he was trying to trick her into giving something up. When asked why, she said that "he seem[ed] exactly like an Iraqi policeman". Later, the others expanded: "Usually police don't want to chat. They want to finish the job first," and "Police are always upset. They are very rude."

Iraqi nationals were solicited as an expert group not only to ensure the system accurately portrayed the Iraqi people, country and culture, but also that there were no offensive elements present (as the software was developed by non-

Iraqis, who may not fully appreciate the nuances of a foreign culture). This was an important secondary goal, as the system is used by members of the US armed forces for training, which places strong emphasis on positive US-Iraqi relations. When asked if there was anything offensive to the Iraqi culture in the game, they said that there was not, rating offense as a 1 out of 5 and adding “No, [Farid] just seems mad.”

### 4.3. DISCUSSION AND CONTRIBUTIONS

In two naturalistic settings, I conducted interviews with participants with expertise highly relevant to the social simulation contained in BiLAT; Army Captains and Iraqi nationals. A main purpose of these interviews was to get expert assessment of the systems’ readiness and accuracy. Participants indeed found the experience to be realistic and engaging. They believed that the cultural behaviors and appearance of the agents were authentic, and the phases and social interactions in the negotiation were accurate. Further, they believed it would be a useful training environment for novices in intercultural negotiation. Therefore, I deemed the simulation to be appropriate for use in future learning studies.

Beyond evaluating the utility and usability of the experience, a chief goal was to identify examples of successful approaches to cross-cultural skills in the real world and how they could be used to enhance BiLAT. What were the experts’ objectives and preconceptions as they approached the task? The interviews made clear that a suite of interpersonal skills were important for successful negotiations, and highlighted skills such systems can focus on developing. In general, the captains indicated that a focus on the social interaction, especially small talk and appropriately opening a meeting, is critical to later success. It is also important to get to know your partner prior to the meeting and listen closely and be respectful through the course of the interaction. At the same

time, there is a need for flexible or adaptable thinking, because new information about your partner might come to light or external circumstances might change at any moment. The structure of the BiLAT environment currently provides opportunities to research meeting partners, conduct meeting openings, and opportunities to engage in small talk. Chapter 5 investigates whether novices take advantage of these opportunities to applying the social interaction strategies recommended by the experts.

It was also important to identify any misconceptions or potential barriers to learning that still remained at an expert level. One area for improvement that was identified was that even experts can approach the environment with a “gamer mentality” that prevents them from engaging with the agents. The captains attempted to explore actions that would deliberately provoke the agents, and often blamed the interface for their failures in communication. The captains had difficulty engaging with the BiLAT agents in meaningful, social ways, even though they professed the importance of such tactics. Also, it appeared that negotiation skill and positive cultural opinions or attitudes do not always coincide; changing attitudes may be a potential target for BiLAT. Chapter 5 will also investigate the presence of these barriers in a novice population.

The contributions of Study I include a significant step towards validation of the content of the simulation, both from the perspective of the Iraqi culture as well as Army cross-cultural negotiation techniques. It also elaborated an understanding of how Captains in the Army approach this task in the field. Finally, it leads towards understanding lasting misconceptions that novices might have in this domain, in particular the attitudinal and motivational challenges that may be targeted in an instructional system.

## 5. STUDY 2: NOVICE EXPERIENCE OF GAMEPLAY

Although BiLAT has undergone iterative cycles of development and has been evaluated for learning results in other contexts, there has not yet been any study that looks qualitatively at how learners interact with the simulation, how they experience the mechanics of the game, and their relationships with the characters. Given the results from the Army experts interviewed in Chapter 4, I was particularly interested in examining the nature of the explanations students generate for events in the game. In training intercultural competence, it is important that learners be able to generate cultural explanations for others' actions, but also acknowledge the individual differences that characterize particular responses. It is also important that they acknowledge their own role in interpreting and responding to others' actions. In BiLAT, students get many opportunities to practice these skills. However, as was evident with the Army experts, learners may also attribute undesirable events to the game mechanics rather than the virtual characters, impeding the learning that should occur from receiving feedback on errors. These are the aspects of the learner experience that I wished to explore.

A related motivational construct to how students make cognitive attributions for game events is how they make attributions for their own performance within the game, called locus of control. Locus of control (Rotter, 1966) is a construct in social learning theory that could give some insight into whether players viewed the virtual characters as having any agency in their learning. In

this chapter, I explore the following questions: how do students interact with the game and the virtual characters, and how do students' beliefs about locus of control relate to success in this game?

## 5.1. SOCIAL CONSTRUCT: LOCUS OF CONTROL

The locus of control construct describes whether the cause to which events are attributed is internal to the student (e.g., high ability) or external (e.g., a “mean” teacher). Students who attribute their performance to internal causes like amount of effort put forth tend to have better learning outcomes than those who attribute performance to external causes (Phares, 1976). These locus of control attributions also seem to be causal; a number of studies have shown that we can retrain students to make adaptive attributions that lead to more learning (e.g., de Charms, 1976; Perry, Hechter, Menec, & Weinberg, 1993). Therefore, understanding people's locus of control attributions in games may ultimately lead to better educational game design.

In contrast to other domains, with skills that involve social interaction (such as intercultural negotiation), it may be the case that social factors play a role in making appropriate attributions. When negotiating with a counterpart in the real world, the outcome is most likely dependent on both parties. Realizing that one does not have full control, and that outcomes are influenced by the other participants in the interaction, may actually be an adaptive rather than a maladaptive pattern. On the other hand, as seen in the expert interviews in Chapter 5, learners who are struggling may place the blame on the interface or interactions in the game. While research on the locus of control typically focuses only on whether a cause is internal or external, I expanded the investigation to differentiate between various external actors as possible loci: the virtual agents, the culture to which the agents belonged, and the game itself. I hypothesized that key differences from standard locus of control findings

would be that successful students give more credit for their performance to the virtual game characters, and less successful students blame the game.

## 5.2. PARTICIPANTS

I recruited participants from the University of Southern California who varied in their frequency of game play on a 4-point scale (never, rarely, monthly, weekly;  $M = 2.8$ ,  $SD = 1.15$ ), as well as in their self-rating of negotiation skill on a 7-point scale where 1=poor negotiation skill and 7=expert negotiation skill ( $M = 3.15$ ,  $SD = 1.4$ ). The 13 participants ranged in age from 19 to 54 ( $M = 34$ ). They were compensated \$40 for a two and a half hour session.

## 5.3. MEASURES

Following each negotiation with an agent, students took a survey based on the Revised Causal Dimension Scale (McAuley, Duncan, & Russell, 1992), in which they rated their own performance at the task and wrote down the most salient cause for their performance. They then rated this cause on a set of 9-point Likert scales (see Table 5.1 for examples). Because I modified the locus of control factor to address individual external actors, students rated whether the cause of their performance was the game itself, the agent with whom they had interacted, or the culture to which the agent belonged. Each possible locus of control included 3 items, which were averaged to create a single score.

**Is the reason for your performance something:**

Over which Farid has no control.	1 2 3 4 5 6 7 8 9	Over which Farid has control
Not under the power of the BiLAT game.	1 2 3 4 5 6 7 8 9	Under the power of the BiLAT game
You cannot regulate.	1 2 3 4 5 6 7 8 9	You can regulate

Table 5.1. Sample items for which learners rated the extent of control over their performance.

## 5.4. METHOD

Students were seated in the lab and took a demographics questionnaire on paper. As described in Chapter 3, they first watched an introductory video about intercultural concepts and skills related to the learning objectives of the game. Students then entered BiLAT and met with the first character, Farid, until an agreement was reached in the negotiation or the allotted time ran out. At this point students were given the locus of control survey described below, and then moved along to a negotiation with a second character, Hassan. Once students reached an agreement or ran out of time with Hassan, they completed a second locus of control survey. Throughout this process the students were prompted to think out loud using the protocol developed by Ericsson & Simon (1993). After the student met with both characters, he or she left the game and entered an open-ended interview with the experimenter.

## 5.5. VERBAL DATA RESULTS

### 5.5.1. GAME-RELATED GOALS AND ATTRIBUTIONS

From the think-aloud data, I found that most students' verbal commentary during the game was highly focused on achieving the proximal task-related goals of the game, or on other game or interface elements. As a consequence, they repeatedly referenced game elements, and attributed events in the game to game mechanics.

For example, the game has a meter that shows the current level of trust that an agent has in the player, which was the spotlight of much of their dialog:

84: "Get that little trust bar up.. yeah..."

910: "... and I'm still sucking wind on the trust meter...okay."

In fact, comments about the trust meter were often made with little understanding of what may have contributed to the change in trust on the part of the agent:

92: "I don't know why he can't tell me, my trust is high enough!"

925: "... and the trust meter is going up so ... that's good I suppose."

Another game mechanic that was mentioned frequently was the menu system for selecting dialog options. Similarly to the experts in Chapter 4, some participants found the options offered by the menus to be restraining, and therefore to blame for their lack of success:

84: "When I'm going through the things that I can say and do, I feel that what I personally want to do is not an option and so I think I try to merge what my options are and what I want to do and it's often failing."

88: "I'm trying to figure out how best to bring this up and there isn't like one that best fits into ..."

While perusing these options, one participant remarked on the lack of time constraints, a feature designed to allow novices to adequately reflect on their choices:

85: "It doesn't matter how long it takes me to look though this stuff, right? Okay, I mean in terms of... I mean [Farid]'s not waiting?"

This difference from human-human social norms, along with others that are designed to support learning, may contribute to feeling a lack of agency on the part of the virtual characters:

3: "Fortunately [Farid] will talk to you as many times as you like, because he is AI and doesn't know any better."

Given these game-related attributions, many learners approached the virtual learning environment as a game environment, displaying exploratory behavior towards the agents rather than behavior that would mirror real-life reactions. While this is an adaptive game-playing skill (Gee, 2003), it may not be

appropriate for contexts where game behaviors are supposed to transfer to social interactions in the real world:

925: “Well, were this an actual meeting... Were this not a game I probably would not have tried that but... like I said, having read that I was curious about whether it worked.”

925: “Honestly, like especially with the bottle of alcohol... I also find sometimes I'm just driven by wanting to see what happens and play with things instead of like actually pretending I'm in negotiations... so perhaps that's not best..”

910: “Do you care if I deliberately do some things wrong? Just because I'm curious as to how it reacts back. That's all um.. exploratory I guess. Things that I know will probably cause me to come to a screeching halt.”

Learners who identified themselves as frequent gamers were especially focused on these issues:

9183: “This just reminds me I just got addicted to this other game on facebook... so it's like every time I fail I've got to go back on and re-click it and try it again.”

Again, this may have led some participants to discount the reactions they received from the agents as accurate feedback on their intercultural interactions, and to rely instead on the training materials they viewed prior to entering the game:

910: “Yeah, my first thought is the game acted incorrectly. Because I took a correct action and it gave me incorrect feedback.”

While these training materials contain appropriate general information on the Iraqi culture, a main goal of the practice opportunities in the game is to push learners to go beyond these materials and adapt to changing circumstances.

### 5.5.2. CULTURAL ATTRIBUTIONS

On the other hand, many participants were able to attribute characters' actions to values or perspectives from their home culture:

84: "Well it said he dabbles in alcohol right? I dunno, that's really risky with these Arabs... I mean they could like.. cause he's religious too... They might just freak out.. I just.. you can't go wrong."

2: "I think I need to be careful now not to be too complementary because of fears of bad luck is kind of a cultural thing."

In other cases, they recognized that each character has a distinct personality, although they at times emphasized these differences over the cultural similarities the two agents shared:

2: "So they were obviously, they were very different people culturally, maybe it was partially an appearance thing and a background thing. So Hassan was strongly devout, religious, and it was in his own personal home, so kind of family issues, perhaps religious issues thing like that were more important and appropriate there, whereas a police officer - kind of with my understanding generally of police officers, kind of by the book."

1: "Hassan was more into ... uh.. religion. Farid was a good family man, but he cared about community a lot more whereas Hassan... not so much."

Unlike the Army experts in Chapter 4, none of the novice participants voiced any negative or stereotyped views about Iraqis. Rather, they relied on the information they received in their training materials.

### 5.5.3. METACOGNITIVE MONITORING AND PERCEPTIONS OF CONTROL

There was a subset of participants who mentioned concepts like *empathy*, *perspective*, or *rapport* with the characters when interviewed after they played, e.g., "I was trying to take his perspective, and put myself in Farid's shoes." These participants also exhibited metacognitive monitoring that suggested that they were acquiring these skills, and recognizing their own influence on game events:

2: "I wonder if I did those two things in the wrong order or if he's just in a certain mood right now that's making him respond differently than before."

918: "I tried to understand his point of view ... I'm not sure that I properly processed uh... um... the meaning of what he was trying... the indirect meanings of what he was trying to say."

2: "Maybe he realized that I'm making an effort to engage in smalltalk, or... I guess he recognized I'm making an effort, that I am sensitive to some issues, I'm not diving into negotiation or anything like that so maybe he wants to trust me and it's kind of naturally increasing by me not diving into the negotiation itself."

918: "I was trying to... try to think from his perspective and couldn't ... on several occasions thought I could predict his response to my response but did so incorrectly... so I think that I don't understand Iraqi culture enough to negotiate with an Iraqi."

They also referenced the desire to interact with the characters in social ways, and an awareness that their actions would have a social response:

9291: "I don't want to ignore or dismiss what he's complaining about and um I don't want dismiss him and keep talking about what I want."

9291: "The meeting with Farid was more, the character that I was playing... the objective was to be more um... not investigative more like um try to gain trust."

4: "I want to do the socially appropriate thing."

917: "He's probably heard this promise before from other Americans."

They exhibited planning to attempt to determine the social consequences of their actions:

929: "Some of the information said that you know, it seemed to indicate that giving a bottle of alcohol would be appropriate but you know I think I'm a little hesitant but I think I'm going to try it anyway."

These participants were less focused on task goals, tended to give the virtual characters more credit for the outcome, and be more successful at the game.

## 5.6. LOCUS OF CONTROL RESULTS

These qualitative results invited further examination of not only how participants attributed game events, but how they attributed their success, using the locus of control construct.

I measured success in the game by determining the average number of negotiation objectives each student met per meeting they attempted in the negotiation. In the first of the two negotiations (with the character Farid), I found that students' ratings of locus of control were significantly correlated to their success. Specifically, success was highly correlated to having an internal locus of control ( $r=.659$ ,  $p=.014$ ); e.g., P87 listed the main cause of his performance as “decent planning,” and rated his own control over the situation at an average of 6.33 out of 9. Success was negatively correlated to blaming BiLAT for their performance in the game ( $r=-.581$ ,  $p=.037$ ); e.g., P93 listed “I was confused on when to bring up the topic of why I came” as the reason for her poor performance, yet rated her own control over the situation as a 4 and BiLAT's control over her performance at an 8.

An independent-samples t-test showed that there were significant differences in ratings between genders in this negotiation. Women were more likely to attribute their performance to the game than men, and marginally less successful than men at achieving objectives in the game (see Table 5.2).

	t(11)	p	Women		Men	
			M	SD	M	SD
Attributions to game	2.31	<b>.042*</b>	7.0	.86	4.56	2.0
Objectives achieved	2.02	<b>.068</b>	.13	.25	.67	.5

Table 5.2. Attribution and performance success between genders; \* indicates a statistically significant difference ( $\alpha = .05$ ).

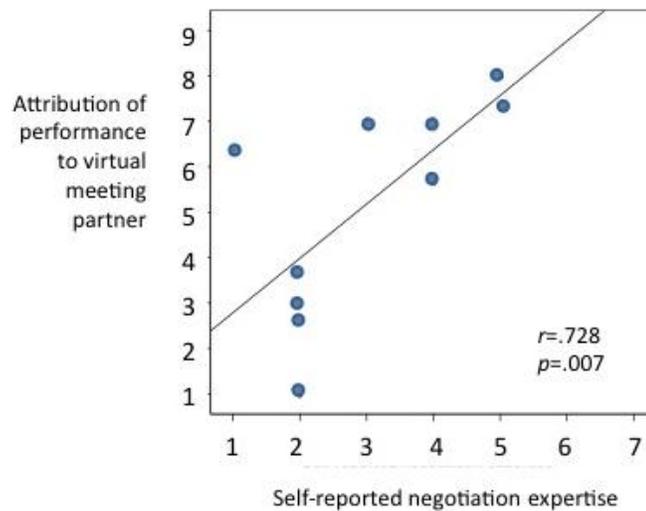


Figure 5.1. Attributions towards virtual agent by negotiation expertise.

In the second negotiation (with character Hassan), performance in the environment dropped significantly (Farid  $M=.54$ ,  $SD=.49$ ; Hassan  $M=.25$ ,  $SD=.45$ ;  $F(1,11)=5.04$ ,  $p=.046$ ). In this negotiation, gender no longer had any effect on attributions along any dimension (all  $r$  values  $< .2$ ). However, significant correlations between locus of control ratings and prior knowledge did appear. Learners with greater self-rated negotiation expertise were more likely to attribute the success of their performance to the character in the game ( $r=.728$ ,  $p=.007$ , see Figure 5.1) and the culture ( $r=.596$ ,  $p=.041$ ), and less likely to attribute their performance to internal factors ( $r=-.775$ ,  $p=.003$ ). For instance, P917, who rated himself a 5 in negotiation expertise, listed the cause for his successful performance as, “building trust with him in the beginning.” He rated Hassan’s control over the situation as a 7 out of 9, and also rated his own control over the situation as a 7 out of 9. Additionally, giving more credit for performance in the game to Hassan, their meeting partner, was correlated with more success in the game ( $r=.587$ ,  $p=.01$ , see Figure 5.2). In comparison, participants’ self-rating of ‘knowledge of Arab cultures’ was not significantly related to their locus of control.

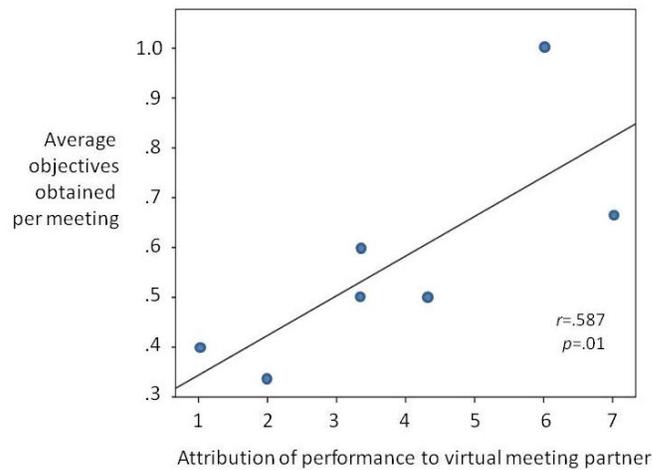


Figure 5.2. Number of objectives obtained per meeting by strength of attributions towards virtual agent.

## 5.7. DISCUSSION AND CONTRIBUTIONS

This chapter explored both qualitative, verbal protocols of learner experiences within BiLAT, as well as quantitative measures of the attributions they made to explain their performance. The key insights gained from the verbal data were that the task goals of the game seemed to be the most salient for the participants, perhaps distracting them from engaging with the virtual characters as social agents. When participants remarked on increasing or decreasing trust with an agent and referenced the game mechanic of the trust meter, they often did so with little understanding of what they had done to drive the change. Some participants engaged in exploratory behaviors that were intended to either find the boundaries of appropriate actions in this context, or simply to disrupt the experience for humor. Like the experts, these participants often blamed their failure on external factors such as lack of appropriate menu options. Participants that engaged with characters as social agents were more successful, and exhibited more of the skills that transfer to real-life intercultural negotiation (such as taking the perspective of another culture). Unlike the

experts, novice participants did not display overt negative attitudes towards Iraqis; this may be because they had no direct experience in context as the Army Captains did.

In the first negotiation, prototypical relations emerged between performance and participants' locus of control; success was linked with feeling like one is personally responsible. As I hypothesized, of all possible external loci, success was specifically negatively correlated to feeling like the environment is responsible for one's performance. I also found that females may have had difficulty in adjusting to the environment – they were slightly less successful, and blamed the game instead of themselves. This is congruent to findings by some researchers that females tend to have maladaptive patterns of attributions (see Dweck, Davidson, Nelson, & Enna, 1978, or Eccles, 1983, for a discussion). This pattern may be detrimental to females' self-efficacy and to their persistence in continuing with the game (Schunk, Pintrich, & Meece, 2007).

By the second negotiation, however, the task became more difficult for everyone. The character Hassan was more difficult to negotiate with, and performance dropped. When this happened, it appeared that some prior knowledge factors became predictive of locus of control patterns. Students with more negotiation expertise were more likely to share responsibility with the character they were engaging with, as well as the culture. As I hypothesized, these participants were also more successful at the game. This may indicate that these students were engaging in a schema of social interaction from their prior experience. In fact, prior knowledge of a domain has been thought to influence attributions (Fiske & Taylor, 1991). However, self-rating of knowledge of Arab cultures did not correlate to attributions in the same way. Among possible explanations is that *knowledge of a culture does not always translate to skill in interacting with that culture.*

The work in this chapter makes a contribution through the first exploration of verbal protocols focused on the intercultural agents in this environment. It uncovers misconceptions and errors made by novice players, and reinforces the successful strategies employed by the experts. Additionally, although the number of participants is small, it suggests quantitatively that investigating learners' feelings of social agency, and understanding the role of prior knowledge, would be fruitful future directions for promoting learning in this environment.

# 6. STUDY 3: MANIPULATING LEARNERS' SOCIAL GOALS THROUGH EXPLICIT PROMPTS

The insights gained from the exploratory Studies 1 and 2 suggest that to enhance benefits from the use of BiLAT, learners may need a greater focus on social interactions with the agents. This chapter describes an empirical study that investigates the effects of manipulating learners' social goals by providing them with an explicit social objective. The study examined whether practice in an intercultural simulation indeed leads to learning, and whether students who hold more social goals both interact more socially with the virtual agents in the simulation, and benefit more from the interaction.

## 6.1. SOCIAL CONSTRUCT: SOCIAL GOALS

### 6.1.1. GOAL ORIENTATION

Motivation is a *goal-directed* process that instigates or sustains behavior (Schunk, Pintrich, & Meece, 2007). In order to motivate students, virtual environments such as BiLAT are often given game elements such as proximal, task-related goals that provide explicit reward structures (Reiber, 1996). Thus, goal theory (Dweck & Leggett, 1988) is a relevant motivational theory for studying games. In educational psychology, two goal orientations have emerged as the most studied: performance orientation, the desire to demonstrate successful outcomes, and mastery orientation, the desire to learn and understand the material. However, my focus is instead on social vs. task goal orientation in a

learning context. In a domain like culture that focuses on social interactions, social motivation can have a great influence on learning (Urduan & Maehr, 1995).

One reason I focus on social goals is because they have been shown to be relevant to student learning. According to Ford and Nichol's taxonomy (1987), there are two main categories of social goals: self-assertive and integrative social goals. The concept of identity is foremost in understanding these goal categories. *Self-assertive* goals relate to asserting one's identity as an individual. This category subsumes goals such as the desire for superiority, self-determination, and individuality. *Integrative goals*, on the other hand, relate to a social identity, or one that is part of a larger community. This category subsumes goals such as affiliation, or the desire to build and maintain attachments, and social responsibility, or the desire to conform to social rules and avoid social transgressions. In the learning sciences, these goals have mostly been studied in the context of a classroom to determine how they correlate with broad measures of success (e.g., Wentzel, 1989; Patrick, 1997). For example, students who have a high GPA tend to hold more social goals, such as being responsible and seeking approval, than those with a low GPA, even while controlling for academic goals like mastery (Wentzel, 1991). Prior research tends not to focus on how social goals may differentially influence learning in specific domains.

Additionally, even outside learning contexts, social goals have been shown to lead to successful interactions. Negotiation researchers have shown that having a sense of a shared group identity, as indicated by holding social goals such as the desire for affiliation, can increase positive attitudes, lead towards a win-win perspective, and even increase negotiation outcomes for both parties (Weingart, Bennett, & Brett, 1993; De Dreu, Weingart, & Kwon, 2000). This sense of shared identity has been implicitly manipulated by researchers prior to the negotiation through discussion of similarities between the participants. Social goals have also been manipulated to reduce outgroup bias in cross-

cultural contact with successful results (Gaertner, Dovidio, Anastasio, Bachman, & Rust, 1993). When cultural differences arise, people with different cultural identities are often categorized as members of an “outgroup” (Prentice & Miller, 1999), which can exacerbate biases and lead to social goals like the desire to be seen as superior to the outgroup (Taylor & Moghaddam, 1994). This result may be particularly applicable to the intercultural skill of *savoir-être*, the ability to express positive attitudes and empathy towards another culture. While improvements in performance outcomes have been demonstrated in various related domains, it is an open question whether such goals can be explicitly manipulated in a way that improves learning in environments in which learners interact with virtual characters from a different culture. Additionally, researchers tend to emphasize integrative goals, and specifically the minimization of differences between groups. Because understanding cultural differences is critical to the intercultural skill of *savoir-faire*, and because social goals may interact with existing academic goals, there is a need to study whether these results apply to a learning context.

#### 6.1.2. INDUCING SOCIAL GOALS

If learners do not produce social responses naturally, it may be possible to introduce social orientations into human-agent interactions such as the ones in BiLAT. A significant area of research involves understanding and influencing learners' motivational orientations towards learning environments (Dweck & Leggett, 1988). To study goal orientations, researchers generally give brief instructions that encourage specific goals for the task, which has been shown to have a significant impact on students' goals (Dweck & Leggett, 1988). For example, Dweck and colleagues told students to focus on either learning as much as they could about the task, or on performing as well as they could. They found that students who were told to focus on learning did learn significantly more than the performance condition. In the same way, giving

learners explicit social goals could influence their social orientation towards virtual humans.

In the randomized, controlled study presented in this chapter, learners played BiLAT with or without the addition of an explicit social goal. This manipulation was implemented with a slight adjustment to previous goal modification work. The goal was presented in an interface element in the virtual learning environment that learners were required to click on, rather than simply constituting a sentence in the task instructions.

## 6.2. EXPERIMENTAL DESIGN

In this study, I gave students pre and posttests to determine the extent of the benefit of the system, both the effects of social goals within BiLAT and the (yet unproven) effects of BiLAT more generally. Further, a major advantage of the use of virtual learning environments is the amount of data generated as students move through and take actions in the environment. This data has not been mined in previous studies with BiLAT, with the exception of looking at the number of errors students made. Investigating learner interactions could provide a more complete picture of the differences between learners, and could support the further development of an understanding of what it means to be social with agents in a virtual learning environment. The hypotheses investigated in this study were:

*H1*: Through practice with a social simulation, learners acquire intercultural skills.

*H2*: Learners presented with a social goal have more social goals for interaction with virtual agents.

*H3*: Learners with social goals acquire more intercultural skills from a social simulation than those without.

*H4*: Learners with social goals have more social patterns of interaction with virtual agents.

### 6.2.1. PARTICIPANTS

Participants were 59 students (32 males, 27 females,  $M$  age = 20.8,  $SD$  = 2.71) recruited by email and flyer from two universities. All participants were U.S. citizens. In a demographics questionnaire participants reported judgments of their own formal negotiation training ( $M=1.20$ ,  $SD=0.58$ ), knowledge of Arab cultures ( $M=1.98$ ,  $SD=1.22$ ), and frequency of gaming, each using a Likert scale of 1 (none) to 7 (extensive). They were compensated \$40 for a two and a half hour session.

### 6.2.2. MEASURES

Intercultural negotiation is an ill-defined domain and assessment of learner expertise is therefore challenging. As noted in Chapter 3, there is no singular measure that gives a complete picture of a student's abilities across such a complex set of skills. However, a number of measures have been developed and validated by researchers for various contexts and purposes that address specific learning objectives associated with intercultural competence. This study employed three measures that assess different skills in this domain.

#### 6.2.2.1. INFORMATION INTEGRATION

The *Information Integration* assessment was created specifically for this task, and was linked to the content in the two scenarios played by the participants. It assesses a learner's ability to develop an accurate model of the virtual characters and tasks in a scenario, by asking participants to rate the truth of various items relating to the task and to the character (e.g., "Farid could be described as a family-oriented man"). These items were taken from the

information participants received in the Prep Room, information which they knew could be accurate or misleading. Participants evaluated each item as true, false, or unknown. Throughout the course of a meeting with a character, successful participants should be able to revise their understanding of the character and the task through their interactions, and make more appropriate judgments about the accuracy of the information from the sources in the Prep Room. I deployed this assessment both before and after participants met a new character. Participants were given a point for each item that matched a subject matter expert's rating of the information.

#### 6.2.2.2. CULTURE ASSIMILATOR

The second assessment used a selection of questions from a validated instrument called the Culture Assimilator (CA) (Cushner & Brislin, 1995), which had never been used in conjunction with BiLAT previously. This instrument is an assessment of the skills of *savoir-faire*, and can be used as a training tool when paired with an instructor and classroom context. It was validated using a known-groups method. In each item, participants read a short scenario in which a person (or group of people) experience confusion in a foreign culture. Participants then select the most likely of four possible cultural explanations for the events that occurred in that scenario. The Culture Assimilator deals with scenarios from both Iraqi and other cultures that relate to the learning objectives in BiLAT, therefore testing whether students can transfer their knowledge to novel situations. Item examples are shown in Appendix A. Because of the high likelihood of scenario recall from pre to post, this assessment was counter-balanced using two versions which each contained six isomorphic items. To score the measure, participants were given a point for each participant answer that matched the validated response.

#### 6.2.2.3. SITUATIONAL JUDGMENT TEST

The third measure, a Situational Judgment Test (SJT), addresses both cultural and negotiation learning objectives. Each item asks learners to read a scenario and then rate the appropriateness of a set of potential follow-up actions that one of the actors in the scenario could take (Legree & Psotka, 2006). During development, several subject matter experts (SME) are recruited to take the SJT. In order to score the test, participants' scores are standardized and correlated with the average of the standardized SME scores. This type of assessment, where correctness is determined by the responses of multiple experts, was designed to be used for measuring recall and understanding in ill-defined domains where there is some disagreement about correct answers. While the SJT is a general technique for assessment, the U.S. Army Research Institute and the Institute for Creative Technologies collaborated on the creation of an SJT for testing knowledge of bilateral negotiation in an Iraqi cultural context. This assessment was validated by verifying a high correlation between the responses of several subject matter experts. The final measure used in this research had nine scenarios with either three or four evaluation items per scenario, chosen because they relate strongly to the learning objectives in BiLAT. Examples are shown in Appendix B. This assessment has been used to collect data in several previous studies with BiLAT (e.g., Lane et al., 2008, Hays et al., 2009).

#### 6.2.2.4. GOALS MANIPULATION CHECK

Finally, to evaluate whether goals were effectively manipulated in the game, participants listed their goals in free text after meeting with each character. They were prompted with the heading, "My goals for this meeting were:". To analyze these responses, two independent raters coded participants' answers to their free-response goals as either social or task goals.

Because the main focus of the experiment was to investigate the effects of playing the game while holding social goals, participants were categorized as “no reported social goals” and “reported social goals”, regardless of their condition. The two coders rated anything that focused on social interaction with the virtual character in the “reported social goals” category.

### 6.2.3. PROCEDURE

Step	Task Goals	Social Goals
1	Demographics	
2	Video Intro	
3	CA + SJT Pretest	
4	Farid Prep	
	+ Task Goals	+ Task & Social Goals
5	II Pretest	
6	Farid Meeting	
7	Goals Check + II Posttest	
8	Hassan Prep	
	+ Task Goals	+ Task & Social Goals
9	II Pretest	
10	Hassan Meeting	
11	Goals Check + II Posttest	
12	CA + SJT Posttest	

Table 6.1. Study procedure for the two conditions.

For all activities, participants were in the user studies lab, seated at a desktop computer and wearing headphones. They began by taking a demographics questionnaire. Next, they watched a twenty minute preparatory video that explained concepts and skills related to the learning objectives, as described in Section 3.1.6. Learning in BiLAT is intended to come through experience and practice, which is often best done in conjunction with some form of direct instruction. Therefore, the video gave participants a necessary introduction to the material that is covered by BiLAT, as well as an introduction to using the system. All participants were told in the video that interpersonal aspects of the

interaction are an important consideration for successful negotiations. Following the video, participants took the CA and SJT pretest assessments. The procedure followed this sequence (video first, and then pre-assessments) in order to test the effects of BiLAT rather than the effects of the video.

Next, participants entered the game (see Section 3.7 for a description of the BiLAT environment). With minimal guidance from the experimenter, they explored the Prep Room to learn about the scenario and the character they were going to meet. At this point, the goals for the meeting were introduced, including the social goal for those participants in the social goal condition. The social goal displayed on the screen was, "Come to understand Farid's point of view" (where Farid was replaced with the name of the current meeting partner). BiLAT displayed this goal on the screen as a single sentence which learners had to select each time they started a meeting, and which was also available in the list of goals throughout the meeting (see Figure 3.7). This social goal was chosen by reflecting on the main focus of both the negotiation and cultural learning objectives: perspective taking.

After leaving the Prep Room, the Information Integration items for that character were administered. Participants then saw the meeting goals again and met with the first character until an agreement on the negotiation was reached. The experimenter only provided assistance in the case of computer error, or minimal guidance if confusion occurred on the separate negotiation screen. If the participant did not come to an agreement within forty-five minutes, the experimenter moved the participant along to the next step. At this point, participants were asked to list their goals for the meeting in free text and then took the Information Integration items for that character again as a posttest. Following this posttest, participants repeated this procedure with a second character who was part of the same scenario. Finally, participants took the CA and SJT posttests. Table 6.1 outlines the steps of the procedure for each condition.

## 6.3. RESULTS

While 59 participants completed the study, 5 participants were dropped due to computer error or complete lack of engagement. In the final analyses, I compared 25 participants in the social goal group to 29 participants playing with the standard task-related game objectives.

### 6.3.1. LEARNING GAINS

*H1* posited that playing BILAT would improve the cultural learning of the participants. To analyze the data, I conducted a paired t-test to compare pretest scores to posttest scores across all students. On the Information Integration items, students showed significant learning from pre to post ( $t(49)=9.213, p=0.004$ ). On the Cultural Assimilator, there was also significant learning from pre to post ( $t(47)=4.582, p=.038$ ). On the Situational Judgment test, however, there was no significant change in participants' scores. Given that their pretests were almost at ceiling (on average greater than 70% correlation with experts), I hypothesized that previously seen learning outcomes with the SJT (e.g., Durlach, Wansbury, & Wilkinson, 2008; Lane et al., 2008) may be attributed to instructional order effects. In prior studies, the SJT was administered before participants viewed the AIDE video. In fact, when researchers at the Institute for Creative Technologies then conducted a small study designed to test learning from the video only, AIDE increased learning as measured by the SJT, but not as measured by the CA. The “remembering and understanding” skills measured by the SJT may be influenced by watching an instructional video, whereas the deeper analytical skills measured by the CA are best achieved through a practice environment (a more detailed explanation is given in Hays, Ogan, & Lane, 2010).

## 6.3.2. REPORTING OF SOCIAL GOALS

H2 stated that the social goals group would have more self-reported social goals for interacting with the virtual characters. Confirming H2, the number of participants with reported social goals was significantly influenced by condition according to a chi square test ( $\chi^2(1, N = 54) = 5.868, p=.015$ ; see Table 6.2).

<b>Condition</b>	Not social	Social	Total
Task	20	9	29
Social	9	16	25
Total	29	25	54

Table 6.2. Number of participants by condition and reported goals.

Even though students in the task condition saw only task goals, a number of them reported social goals anyway. Additionally, the reported social goals from both conditions appear very similar in nature (see Table 6.3). Students in both conditions saw task goals, and reported task-related goals that very closely resembled the goals listed in the interface (see Table 6.4).

<b>Task goals condition</b>	<b>Social goal condition</b>
To help both sides get what they wanted and maintain a peaceful negotiation.	To establish a positive personal relationship with Farid
Establish a good relationship	Maintaining a strong personal rapport with Farid.
Build trust	Building the foundations for a long term relationship with Farid.
Establish a trustworthy relationship between Farid in myself. I wanted to get support and unity with the Iraqi police but I also wanted to attend to Farid's needs in order to keep the relationship open and favorable.	To understand Farid's point of view and to get the backing of the Iraqi police.

Table 6.3. A comparison of goals labeled "social" that were reported by participants in each condition.

**No Reported Social Goals**

<b>Task goals condition</b>	<b>Social goal condition</b>
To address the market issue and come up with ways to solve the problem	Learn info about the market get cooperation
To get police cooperation and solve problems in the market.	Police cooperation, find out why people aren't using the market
Fix the market problem, get more help from the IP	Resolving market problems

Table 6.4. A comparison of goals labeled “not social” that were reported by participants in each condition.

It was difficult to determine why some students in the task condition reported certain social goals, and why some students in the social condition did not. Reporting a social goal was not significantly related to the demographic characteristics I measured: age, negotiating experience, knowledge of Arab cultures, or frequency of playing games. Neither did the reported social goal participants have higher pretest scores on the Information Integration items or the Culture Assimilator (all  $p > .2$ ).

### 6.3.3. LEARNING GAINS BY REPORTED GOALS

*H3* stated that students with social goals would perform better on measures of learning. I first examined all of the learning results with respect to the learner's *given goals* (social vs. task), using a repeated measures analysis of variance (ANOVAs) with condition as between-subjects factor and test time (pretest and posttest) as within-subject factor. Here, my hypothesis was not confirmed. However, because almost a third of all participants did not report their goals as expected based on the condition they were given, I examined all the learning results again with respect to the learner's *reported goals* from the manipulation check (“reported social goals” vs. “no reported social goals”). “Reported social goals” was added as a between-subjects variable to repeated measures ANOVAs with test time (pretest and posttest) as within-subject factor. On the

Information Integration items, the ANOVA showed that reported social goals significantly influenced learning ( $F(1,49) = 3.979, p=.052$ ). An ANCOVA model of the Culture Assimilator test also showed that reported social goals significantly influenced learning ( $F(1,47) = 8.314, p=.006$ ). On both assessments, learners who reported social goals outperformed learners who did not report social goals.

#### 6.3.4. GAMEPLAY RESULTS

	t(52)	p	Reported social		No reported social	
			M	SD	M	SD
Total actions	2.45	<b>.018*</b>	79.0	27.0	107.0	49.0
Business actions	2.60	<b>.012*</b>	29.6	12.6	40.0	16.1
Social actions	1.82	.073	50.1	21.3	67.1	42.2
Unique actions	2.24	<b>.029*</b>	37.0	7.1	43.0	8.6
Meetings	2.04	<b>.046*</b>	2.8	1.6	3.9	2.3
Time played	0.23	.800	46.2	10.3	47.0	13.6

Table 6.5. Gameplay differences between “reported goals” groups; \* indicates a statistically significant difference ( $\alpha = .05$ ).

*H4* stated that students with social goals would have more social patterns of play in the game. To examine this hypothesis, I investigated participants' actions taken in the game and compared them across reported goals. Game actions were classified by the developers into opening (e.g., greetings), social, business, and closing (e.g., leave-taking) categories. Participants who *reported social goals* took significantly fewer total actions in the game than participants who did not report social goals (see Table 6.5 for all statistics). Broken down by action type, participants with self-reported social goals took significantly fewer business actions, while taking a statistically similar number of social actions. They also held fewer meetings before they achieved their objectives. Because actions could be repeated in the game, I additionally assessed the number of *unique*

actions each student took. Participants with reported social goals also took significantly fewer unique communicative actions. There were no significant differences between groups in the amount of time the game was played.

#### 6.3.5. GENDER GAP

Although it was not one of the initial hypotheses, in exploring the data, a trend emerged indicating that gender might be related to learning outcomes, as was also seen in Study 2. Related research indicates that there may in fact be gender differences in social behaviors, social roles, or social focus (e.g., Hoffman, 1977; Hall, 1984). In fact, in most learning measures in this study, males emerged as significantly more successful. The next chapter presents several hypotheses for why this overall trend of gender differences exists, and potential solutions that may be implemented within the system to correct them.

### 6.4. DISCUSSION AND CONTRIBUTIONS

The four hypotheses of this study were confirmed, although not entirely as expected. First, overall, participants did learn from interacting with the virtual humans in BiLAT. Showing that a social simulation for intercultural training can produce increases on a validated assessment is a contribution given the current state of evaluation in the field (Ogan & Lane, 2010), which relies mainly on qualitative measures such as interviews (or has not yet progressed to evaluation at all). The results of the explicit goals manipulation in the study require closer consideration. Relevant to *H2*, a significantly higher percentage of participants who were explicitly given a social goal in the game interface did report having a social goal as a meeting objective than those who were not given such a goal. Therefore, some students in the social goal condition did appear to be influenced to consider such goals as important. However, a third

of the students in the social goals condition did not seem to heed the manipulation. These students may not have understood how to achieve the social goal they were given, or may not have wanted to achieve it. Additionally, unlike other similar learning orientation manipulations (e.g., Dweck & Leggett, 1988), simply attempting to manipulate students' social goals was not beneficial to learning.

What the data showed instead, confirming *H3*, was that learners with *self-reported social goals* for the interaction had increased learning over learners without such goals. The results of this study extend two sets of seemingly conflicting evidence. Okita et al. suggest that students who believe that virtual characters are real benefit more from interacting with them (Okita, Baileson, & Schwartz, 2008; Rose & Torrey, 2005). However, the Media Equation (Reeves & Nass, 2006) suggests that social responses to virtual characters are automatic. In the current study, participants with self-reported social goals, regardless of condition, learned more about the scenario and characters and were better able to transfer their knowledge to novel situations. It may be that while social responses to pedagogical agents are not automatic, learners who hold social goals for interacting with them treat the interaction as more authentic and thus benefit more.

Additionally, students who reported a social goal played the game in a qualitatively different way (confirming *H4*). In an identical amount of time in the game, they took fewer actions, which may indicate that they spent more time reflecting on each action to consider their partners' perspective. They took fewer total actions relating to business and a higher percentage of social actions than students who did not report social goals. Additionally, they took fewer unique actions, signifying less exploration of the conversation space (seemingly avoiding dialog actions that could potentially be seen as offensive). Together, these patterns seem to present a social orientation towards gameplay, where participants hold some theory of mind about their virtual partner's thoughts,

feelings, and motivations. This outlook is in contrast to a prominent view of learning from gameplay, which involves exploring a risk-free, task-oriented environment for discovery, as suggested by the PsychoSocial Moratorium principle and other theories of game-based learning (Gee, 2003; Johnson, 2006). This principle states that games are a place where learners benefit from taking risks they would not normally be comfortable taking in the real world. In BiLAT, learners might manifest this principle by intentionally offending the virtual character, or experimenting with all available actions in an attempt to understand the boundaries of acceptable behavior. In fact, participants did engage in these behaviors, as can be seen in the verbal protocol data presented in Chapter 4. This type of gameplay is at odds with a social perspective, in which learners would carefully consider their partner's perspective, attempt to avoid giving offense, and avoid exploration that would take them into unknown territory of culturally acceptable behaviors. A contribution of this study is that "mindful" social interaction in a game rather than possibly "mindless" trial and error behaviors is associated with greater learning.

These results show that a social orientation towards interaction with virtual humans in learning intercultural negotiation and perhaps similarly complex social skills may not always happen, but when it does, it is associated with increased learning. It appears that learning social skills in virtual environments can be more effective if we can promote a social orientation. However, presenting explicit social goals was not the most effective way of doing so. Instead, in the next chapters this dissertation presents ways to *implicitly* promote a social orientation. A second avenue of research suggested by these results is to investigate those students who arrive at a learning environment already holding social goals. Reporting a social goal was not tied to any of the demographics measured in this study (e.g., prior negotiation training, knowledge of Arab cultures), nor did these participants appear to be of higher

ability based on their pretest scores. It may be that other measures such as social intelligence or personality traits can provide a better characterization of these learners. The next chapters discuss how social goals are influenced by and interact with these learner characteristics.

## 7. SOCIAL INFORMATIONAL DIALOG DEVELOPMENT

The results from Study 3 suggest that learners who hold social goals for interacting with the agents in BiLAT learn more than those who do not. However, explicit telling was not a successful way to change learners' goals. Instead, in continuing work, I investigated ways to implicitly affect the interpersonal relationship between the learner and the agents. In particular, I selected dialog as the means for influencing learners' social orientation. Traditionally, the instructional dialog of ECAs has a task orientation, in that it focuses on the instructional task. For example, in a physics tutor, the main focus of agent dialog might be to assist students in solving the next step in a momentum problem, or in better understanding a concept like force or velocity (e.g., Litman & Silliman, 2004; Jordan, Makatchey, & Pappuswamy, VanLehn, & Albacete, 2006). Increasingly, the motivational and affective components of student-agent interactions are receiving greater attention (e.g., D'Mello, Picard, & Graesser, 2007; Kim & Baylor, 2006). Especially when considering instruction in interpersonal domains, such as negotiation, the development of an interpersonal relationship with one's pedagogical agent may play a significant role in learning.

As discussed in earlier motivation, belief that a human was generating the instructional dialog in a system was enough to increase learning outcomes - a sign perhaps that pedagogical agents must be socially "believable" in order to be most effective (Rose & Torrey, 2005; Okita, Bailenson, & Schwartz, 2008). Thus, it is important to develop dialog models and agents that generate desirable interpersonal effects with the learner. I targeted three relational

outcomes that may affect learning, especially in an intercultural domain: trust, entitativity (the perception of a group as one entity), and shared perspective.

How to create agent dialog that leads to these desirable outcomes is an open question. In this chapter, I describe the development of an instructional dialog model aimed at achieving these outcomes, which I call *social informational dialog* (SID). In its development, I created a process for generating agent dialog with the input of participants from different backgrounds. While agent dialog has been developed in other systems, this work provides a formalization of a procedure for doing so and offers insights into lessons learned during the process. At the same time, I turned to communications literature for conversational strategies that are hypothesized to have effects on entitativity, trust, and perspective, integrating these findings with the insights gained from iteration with users. This chapter also describes a corresponding agent with (more typical) task instructional dialog, which was used as a comparison to measure the success of these strategies in Studies 4 and 5, described below in Chapters 8 and 9.

## 7.1. SOCIAL CONSTRUCTS: TRUST, ENTITATIVITY, AND SHARED PERSPECTIVE

Although there are many potential interpersonal outcomes that may affect learning, in this work I have targeted three: trust, entitativity, and shared perspective. These concepts and their relation to learning are described in the subsequent sections.

### 7.1.1. TRUST

In instructional tasks, *trust* increases peoples' ability to make reliability judgments about the accuracy of the information they are receiving. In general, researchers have found relationships to be important for acquiring information

(Burt, 1992). Early work by Pelz and Andrews (1966) and Mintzberg (1973) indicates that people are preferred over documents as a resource for information, which would suggest that virtual agents are a useful delivery mechanism for instruction. Trust literature has found considerable evidence that higher levels of trust lead to an increased willingness to listen to useful knowledge and absorb it (see e.g., Levin, 1999; Mayer, Davis, & Schoorman, 1995). Thus, agents should engender a feeling of trust through their actions to address social and instructional objectives.

### 7.1.2. ENTITATIVITY

A second desirable relational outcome is *entitativity* - the feeling of working together as a team (Lickel, Hamilton, Wierzchowska, Lewis, Sherman, et al., 2000). Although the concept of entitativity specifically has not yet been linked to learning, numerous studies have reported the positive effects of collaboration and groupwork on learning (e.g., Johnson & Johnson, 1990; Lou, Abrami, & d'Apollonia, 2001). In particular in negotiation literature, this feeling of working as a team leads to more positive affect towards negotiation partners, and also significantly better negotiation outcomes (Gelfand & Brett, 1983). While one of the two proposed antecedents for entitativity is similarity of goals (Ip, Chiu, & Wan, 2006), Lickel et al. (2000) have found that intimacy groups such as family and small groups of friends have higher entitativity than task-oriented groups such as committees or work groups. Pedagogical agents should be able to stimulate a feeling of entitativity in the learner by developing a sense of intimacy that goes beyond simply having shared task goals.

### 7.1.3. PERSPECTIVE

Finally, agents should be able to influence the *perspective* that learners take in the interaction. The agents in BiLAT are enculturated, meaning their behaviors (gestures, utterances, etc) are driven by a model that is, in part, culture-

specific. They are intended to help learners understand behaviors and values from the agent's cultural perspective, as well as assist the learner in taking actions that conform to that cultural perspective. In both intercultural competence and negotiation, perspective-taking has been shown to be important for success (Galinsky & Mussweiler, 2001; Gelfand & Brett, 1983; Jones & Nisbett, 1987; Neale & Bazerman, 1983). Even beyond intercultural education, the ability to take on the shared perspective of an organization is of great value. For example, in STEM (science, technology, engineering and math) education, one central objective is for students to be able to see themselves taking on the persona of scientists or mathematicians. Whether the culture is national or organizational, interaction with an enculturated agent should persuade learners of the importance of sharing that perspective.

## 7.2. DIALOG DEVELOPMENT PROCESS

Within the framework of the BiLAT game, I developed a new agent named Zahora (see Figure 7.1) to evaluate the two instructional dialog models described in the next section. Utilizing the same non-verbal behavior models that underlie the rest of the agents in BiLAT (see Hill, Belanich, Lane, Core, et al., 2006 for more detail), she moves with culturally-appropriate gestures, gaze, and posture. Zahora speaks in English with a computer-generated voice that has a slight accent. She is introduced as an Iraqi interpreter with knowledge of the culture and negotiation experience.



Figure 7.1. Zahora, the interpreter. The agent is fully animated: she moves with culturally-appropriate gestures gaze, and posture.

Unlike the rest of the meeting partners in BiLAT, Zahora is designed to take a more explicitly instructional role. Her instructional dialog model is based on a set of branching mini-dialogs that typically last two to six turns (depending on the learner's choices), each of which are organized around a particular learning objective about Iraqi culture or negotiation. Learning objectives vary from casual subjects, such as food, to more complex social dynamics, such as the differences in the Iraqi and American concept of privacy.

This section progresses chronologically through the steps of how I developed the dialog for the agent. First was *Preparation*, in which I evaluated and selected learning objectives for the content of the dialog, and developed a backstory for the character. Second was *Dialog Generation*, accomplished through role-play techniques using American nationals with knowledge of Iraqi culture. This generation created a large corpus of natural dialog. Then, *Iterative Development*

revised the dialog using multiple lightweight prototypes with students. Finally, in *Dialog Validation*, I presented the completed dialog to Iraqi nationals for one final revision that aligned the dialog to Iraqi conversational norms.

### 7.2.1. PREPARATION

To prepare for creating the agent dialog, I first generated a list of potential conversation topics for the content. These learning objectives were taken from material in the widely cited book, *Understanding Arabs* (Nydell, 2006), that describes Arab culture with specific details on Iraq. With a thorough review of the material, I chose issues that were important in the Iraqi culture which also have a strongly contrasting perspective in a prototypical American cultural setting; for instance, the concepts of privacy, fate, and family relationships. A second criterion for the objectives I chose was that they were not already covered by BiLAT, so that the content presented by Zahora would not overlap with the content of the rest of the learning environment.

A complementary preparatory step, also used by the developers of other BiLAT agents, was the creation of a backstory for Zahora. This backstory, known as a *character bible*, contained more extensive demographic information about the character than I expected to be used during dialog generation, as well as details about her family, their occupations, her relationship with them, and so on (see Table 7.1). There has been a wealth of material on the creation of narrative, including for video game environments (e.g. Bateman, 2007; Despain, 2008; Mirrielees, 2008); I followed the same general writing processes, which typically include generating this type of information prior to writing for a character. When questions arise about whether the character would be likely e.g. to say a particular utterance or feel a particular emotion, a development team can return to the character bible and have a better understanding of how she might act in that situation. The character bible became an important go-to resource throughout all stages of dialog development.

---

**Zahora**

---

Female, 28 years old, 150 pounds, 5'3"

Nationality: Iraqi, Ethnicity: Arab, Religion: Shia Muslim

Wears a headscarf

Has worked as a translator for the American forces for 3 years

Used to be a teacher in a secondary school that was shut down when a bomb struck the sports fields

---

Table 7.1. A sample of content from the character bible for Zahora.

Zahora's backstory began with her role as interpreter. In creating the character bible, I continuously consulted the learning objectives to ensure that her story gave her reason to discuss some of them in the dialog. For example, one objective related to family relationships, and so Zahora's bible had a reference to the family-run business that employed the rest of her siblings. Development of the character bible also influenced the learning objectives list. As I brainstormed interesting features for Zahora's life, I reciprocally consulted the Iraqi cultural materials to determine whether there existed a potential new learning objective that could take advantage of her backstory. It was also important, to allow for an engaging narrative, to introduce a little drama into her life. Therefore, Zahora became an interpreter after an errant bomb destroyed the school where she had been teaching. This gave the potential to create a dialog with buildup and climax, with a sympathetic character.

An early and critical decision in the creation of the agent was deciding on a gender. In general, great care should be taken in understanding the role that an agent will play in a system. Agents may take on many different roles, from coach, to peer, to tutor (Chou, Chan, & Lin, 2003). It is also known that the

characteristics of an agent, whether it is skin tone, age, or speech style, can affect user interactions (Cassell, 2009; Baylor & Kim, 2004). In this case, as part of a learning environment that delivers training in intercultural interactions within the Iraqi culture, gender was an important concern. For cultural reasons, the majority of people that soldiers (the population for which BiLAT has been designed) interact with in Iraq tend to be male; most of the soldiers performing these interactions will be male as well. For this reason, BiLAT is largely populated with male agents.

In this case, there were several reasons to consider changing the gender of the agent. In Studies 2 and 3, I found significant differences in learning by gender. Female learners were at a significant disadvantage in learning gains, despite starting out with similar pretest scores to their male counterparts. This finding may have come about for several reasons. In general, women tend to be intimidated by the concept of negotiation (Babcock & Laschever, 2003), one of the key instructional components in BiLAT.

The all-male cast of characters in the environment may be another source of intimidation. That source, however, is one that can change. In further research into the role of women in the Iraqi culture, it became evident that the role of interpreter was one that increasing numbers of Iraqi women are taking on. Therefore, it seemed appropriate to both prepare male learners for undertaking interactions with women in this context, as well as giving female learners a gentler introduction into the system.

### 7.2.2. DIALOG GENERATION

Following preparation of the character and the learning objectives, I began the development of agent dialog. To aid in formulating the procedure for the dialog generation, I worked closely with a former executive who managed content creation for a network television station. I began by selecting a set of

participants to play the roles of Zahora and the game player. Rather than an instructional designer writing all of the dialog by hand, the use of role players allows for a more natural and dynamic basis for discourse, which can then be revised to ensure that it contains the appropriate instructional content. It also greatly accelerated the process of generating the large amount of content necessary for inclusion in the system.

The role-players used were American citizens, who were employees of the Institute for Creative Technologies at the University of Southern California. While the participants I recruited were not part of the BiLAT project, they had worked on other intercultural projects for the Army and were familiar with the aspects of Iraqi culture that were intended to be presented in the game. Therefore, while they did not consider themselves to be members of that culture, they were able to produce utterances that may be in line with that cultural perspective. The cultural perspective was later validated as described in Section 7.2.4.

The role-play was set up in a conference room, with recording equipment to capture both audio and video for future analysis. A good microphone was essential to record audio from multiple parties. While capturing video was not strictly necessary, it allowed for a dual channel review of emotion, attentional focus, and other nonverbal cues that might influence the dialog direction. During the session, two women played the role of Zahora and one man took the role of the Player (see Figure 7.2). The decision to include two people in the role of Zahora was to elicit varying perspectives on the conversation, which would generate multiple dialog utterances to choose from.



Figure 7.2. Role-players involved in the development of the new agent dialog.

Each of the participants taking the role of Zahora was given a short packet that presented the character bible, a brief description of the circumstances leading up to the current meeting with the Player, and a listing of Zahora's goals and values as related to the particular learning objectives that were expected to be present in the learning environment. To elicit a variety of responses, the two participants were each given different tones to take in the conversation. One was given a positive role, as an interested person who wanted to help the soldier in gaining comfort in his new environment. The other was instructed to be slightly more confrontational and not always in agreement with the soldier. This role was intended to produce responses with potentially greater tension or drama, to engage players in the conversation.

The participant taking the role of the Player was given a sheet with a description of his role as an Army officer newly arrived in Iraq to assist in reconstruction efforts. The sheet also described the goals for the interaction, which were to get acquainted with an Iraqi and begin to understand a little bit about their culture. The Player was instructed to begin the conversation by entering the room where Zahora was seated and offering a greeting.

In the conversation, participants were seated facing each other, and instructed to interact with each other as if it was their first conversation. The Player started the conversation with a greeting, and the two Zahoras took turns responding to it. Over the next hour, participants continued conversing based on the goals and values described on their character sheets. At first, I asked each Zahora to respond to each statement or question from the Player, and the Player would choose to address only one of their utterances (e.g., the one he found most interesting). This was done in order to obtain more (varied) options for later selection into the dialog. The role-players themselves modified this strategy as the conversation got more in-depth; as the discussion intensified, one Zahora would often continue responding for several turns before passing the conversation off to the other Zahora. Although this meant that there were fewer options to select from to compose the dialog, I felt that when the participants became so invested that they did not want to give up the floor, they were producing a compelling narrative, and so I did not stop them.

After an hour, the role-players took a fifteen-minute break, during which they were separately given further instructions to take the conversation in a new direction. These directions came from observation by the research team as to what learning objectives needed more content, and to interesting points from the previous conversation that could use follow-up. Following the break, participants continued for approximately another hour of conversation, until they felt they had exhausted their understanding of the conversation topics. Overall, this process generated two hours of conversation and over a thousand utterances from different perspectives to use as a basis for the agent dialog.

### 7.2.3. ITERATIVE DEVELOPMENT

As I was told by entertainment industry employees working at the University of Southern California, “dialog is about rewriting.” The next stage in dialog development was to manipulate the conversation generated in the previous

step into the appropriate format for BiLAT, and then iterate on the wording with input from users. The virtual meeting environment within BiLAT offers a set of actions and dialog choices to the learner, which are always available through a button-click interface (refer to Figure 3.7). In general, the interaction in BiLAT is turn-based, with each choice by the learner being followed by an utterance from the agent. The learner can then select another action or utterance from the same menu and receive another response.

In addition to this set of actions, which is specific to the current meeting, a second type of interaction is allowed through “challenge networks”. In a challenge network, the agent initiates an utterance, the standard set of actions disappears, and the player is given a choice of several new responses. This type of interaction can be used when agent-initiated dialog is desired (e.g., the agent asks if the player wants tea, and then the player is given several potential replies), or when the scenario designer wants to offer choices that are only available at a specific time in the meeting. In previously existing scenarios, these actions and dialog choices were authored by the development team in collaboration with Iraqi nationals and military officials.

The initial interaction with the new agent Zahora looks much the same as those in the rest of BiLAT. The learner sees a set of actions and dialog choices that I selected from interesting “conversation-starter” utterances made by the Player in the role-play, such as a question, or a statement that opened up further discussion. When the learner selects an utterance from this initial set, the agent responds with an utterance that I selected from the role-play dialog. This response provides entry into a mini-dialog that covers a particular learning objective.

I chose to implement these mini-dialogs through a branched series of challenge networks. After each agent turn, the learner is given three choices of how to respond, which were developed to represent different perspectives: American, Iraqi, or a compromise of the two. The American reply was taken from the

Player's verbal data in the role-play. The Iraqi perspective was generally developed by writing a statement of agreement with the agent utterance, and adding on additional information taken from an Iraqi perspective in the role-play. Finally, the compromise perspective was generally developed by writing a new utterance that took a middle road; offering a contrasting opinion, but with some concession to the Iraqi perspective. Table 7.2 shows an example of three learner choices offered within one dialog turn. Presenting the learner with a choice of perspective at every step allowed the system to log that perspective for future data analysis.

Perspective	Description	Sample Utterance
Iraqi	A prototypical Iraqi perspective, generally in agreement with the utterance spoken by the agent	"I agree, my family is my rock."
American	A prototypical American perspective, developed through role-play with American nationals	"Well, I wouldn't say that. I love my family, but I wouldn't go into business with them!"
Compromise	A compromise perspective in which neither strong cultural perspective was emphasized	"Actually, I'm close with just a few."

Table 7.2. A sample of three learner choices within a dialog about the role of family in Iraq. Zahora has just replied, "Yes, family is very important to us as well. My family are the only people that I can really trust. This is why brothers and sisters all stay in the family business. My family sells pottery."

After the learner makes a choice, the next challenge network begins with an agent utterance, until a reasonable ending point to the mini-dialog is reached. Transitions between the mini-dialogs are completed at these points through discourse markers in the agents' utterance, e.g. "Anyway...", indicating that the agent is ready to move on to a new topic. At these points, the learner is returned to the initial menu of choices.

A benefit I found to implementing the dialog mainly through challenge networks, compared to the way most dialogs are implemented in BiLAT, is that learners could delve further into particular cultural topics. This is something that novice participants in Study 2 had stated that they would like to do. It also reduced the number of initial actions to a manageable set for beginners (from approximately 50 actions down to 10), placing more emphasis on depth of conversation than breadth.

The tradeoffs to making extensive use of challenge networks, in particular in this branching fashion, is the complexity of organization and programming. To assist in organization as I was developing and refining the dialogs, I used PowerPoint slides with internal links based on user choices. The initial set of learner choices resided on the first slide. Each learner utterance was then linked to the appropriate agent reply, on a different slide. Each agent slide had a “Continue” link to move to the next set of learner choices, and so on. When a branched dialog was ending, it linked back to the first slide to allow learners to enter into a new conversation. PowerPoint, although a comparatively simple technology, allowed for rapid iterative development including rewriting, rearranging, and deletion or addition of content, often on the fly as users suggested revisions.

Challenge networks also require greater development of content, because learners are offered multiple response choices, each of which must be followed by an agent response. One way in which I mitigated this tradeoff was to collapse some of the branches of the networks. If several of the learner choices in a dialog could be reasonably answered with the same agent response, I linked them to the same subsequent challenge network. This reduced the amount of content development, but made the network structure more complex.

Once the translation from recorded role-play conversation to digitized challenge-network format was complete, I used the linked PowerPoint file to iterate on the dialog with approximately fifteen participants. I showed each

participant the file on a laptop, allowing them to click through the dialog making their desired choices. While progressing through the dialog, they used a think-aloud protocol focused on whether they understood Zahora's comments, if they believed each challenge network contained a dialog choice that they felt comfortable selecting, and their affective reactions to the differences in perspective presented in the conversation. Viewing user reactions was a critical step in the writing process, especially given that the dialog was intended to support interpersonal relationships.

I continued this procedure, making modifications after each participant, until participants seemed to converge on finding appropriate choices and feeling a positive affective reaction to Zahora. I observed that most participants' initial reaction was to look for the dialog choice that I had labeled as "American", giving some confirmation that these choices indeed captured some resemblance to an (university-educated) American perspective.

Following these iterations, I moved to a second modality of presentation. Given that the wording and the branching structure had been tested and were fairly stable, I developed linked HTML pages to put the dialog online. This allowed for both remote testing, as well as the tracking of the responses that learners clicked. For example, I was able to confirm that American participants tended to select the "American" perspective response.

#### 7.2.4. VALIDATION WITH IRAQI NATIONALS

The final step of development was to validate the completed dialog with Iraqi nationals (using the Iraqi participants from Chapter 4). This was important in order to account for the American participants' cultural biases. I chose the methodology of dialog generation with Americans followed by validation with Iraqi women for several reasons. One was the conservation of resources. Recruiting Iraqi participants who are appropriate for this type of activity is

difficult and costly. Having a draft of the dialog for them to evaluate and revise allowed them to skip the generation stage, which is the most time-consuming. Although it would have been beneficial to consult an Iraqi perspective at various points in the process, this method allowed for the best use of available resources. Additionally, while they are acting as our cultural experts, they are not experts in pedagogy. Generating a first draft of dialog with members of research teams who have experience in developing such content can also assist in introducing the appropriate pedagogical content. The Iraqi women consulted during this stage generally approved of the dialog, but provided a number of minor tweaks and one or two major modifications (saying, e.g., “I would *never* say that! That is rude.”). Following this evaluation, the dialog was considered ready to be implemented into BiLAT.

### 7.3. SOCIAL INSTRUCTIONAL DIALOG MODEL

Imparting instructional objectives is not the main goal of Zahora’s dialog. Throughout the process of dialog development, my investigation focused on how to design Zahora so she implements the notion of *social* instructional dialog, introduced above. Surrounding the branching conversational framework described above, I developed a model of social informational dialog (SID) based on communication theory. The model (shown in Figure 7.3) uses three main conversational strategies to communicate learning objectives: self-disclosure, narrative, and affirmation. I chose these strategies both because they are posited in the communications literature to affect the relational variables described in Section 7.1, but also through the insights gleaned from working with the various participants described in Section 7.2.3.

While the choice of the strategies used in the model is heavily influenced by previous work done by Bickmore and Cassell (2001) and other work on individual features of social relationships in pedagogical agents, such as

politeness (Wang, Johnson, Rizzo, Shaw, & Mayer, 2005), this work differs in two important ways. First, this dialog is intended to be instructional rather than transactional. The agent is intending to convey learning content with each dialog move, and therefore the strategies are employed to deliver that content. Secondly, the utterances created through the implementation of our model are transformed through conversational strategies that holistically change the dialog, rather than tacked on to task dialog as additional utterances (see Table 7.3 below as an example). This both allowed the length of the dialogs to be relatively constant between task and social implementations, as well as to keep the social flow of the conversation.

The following sections describe the conversational strategies in the SID model, their relation to the positive interpersonal outcomes, and how they were influenced by the dialog development process.

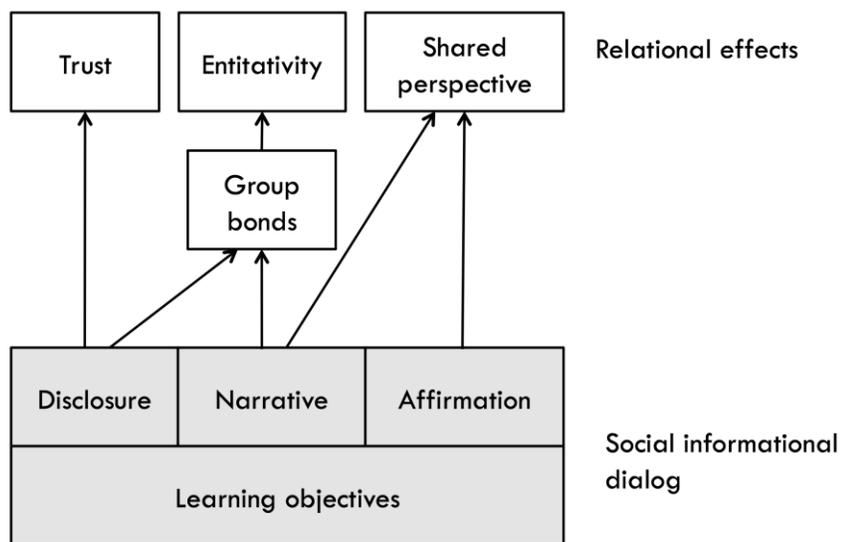


Figure 7.3. Model of social informational dialog and effects on relational variables.

### 7.3.1. NARRATIVE

Narrative is a fundamental form of human communication, which is increasingly popular in the Artificial Intelligence in Education (AIED) community as a method of conveying information, (e.g., McQuiggan, Rowe, Lee, & Lester, 2008; Riedl, Arriaga, Boujarwah, Hong, Isbell, & Heflin, 2009; Schank, 1990). This conversational modality has many hypothesized effects on communication. For example, stories embedded in narrative allow the management of social bonds among group members (Bochner, Ellis, Tillmann-Healy, 2000). Secondly, the use of narrative as a communication tool provides context and allows for more personalized and persuasive presentation. Bruner (1986) describes two ways of knowing: paradigmatic, which utilizes argument to describe truth, and narrative, which conveys descriptions of experience through personal perspectives. Communications literature suggests that this narrative way of knowing is as good for knowledge transmission and retention as the paradigmatic. The key benefit, it seems, stems from our aptitude for memorizing and processing stories – a tradition that has existed for millennia. Specifically, stories allow learners to leverage pre-existing schemas to acquire new information. Additionally, heightened affect produced by narrative leads to arousal, which, in turn, affords greater and more focused attention from the learner (Lang, Bolls, Potter, & Kawahara, 1999).

In my implementation, SID presents learning objectives within the context of a story. As described in the Preparation stage, I first introduced narrative with the development of the character bible for Zahora which included details about her life and her family. In Dialog Generation, players were given a brief verbal prompt, “Incorporate any information that you find useful from the character bible to enhance your replies in this conversation.” As the role-players conversed, they very naturally turned to this document which provided them with a strong basis to weave a narrative about Zahora’s observations of Iraq. For instance, Zahora uses the story of her family’s pottery business to describe

her perspective on familial duty and closeness. The Player, who was given less of a backstory to study, often drew from his own personal life while replying to questions. For instance, he mentioned his new baby daughter and his relationship with his in-laws, when Zahora asked him about family matters.

As described in the Iterative Development stage, I then used the narrative generated by the role-players to divide the conversation into mini-dialogs that each encompassed a learning objective. The SID model as implemented in BiLAT also has means of tracking topics discussed, so that if for example Zahora's cousins are introduced in one mini-dialog, they may be used as a referent in a subsequently chosen mini-dialog. This enables the narrative to flow more smoothly, and conforms to assumptions about common ground. Finally, in the Validation stage, the narrative was confirmed to be appropriate in the Iraqi culture.

### 7.3.2. AFFIRMATION

The second strategy the SID model incorporates is affirmation, the acknowledgement that the receiving party in a communication has been heard and understood. In my implementation, affirmation is accomplished through affirmative statements. These markers (e.g., "I see", "Yes") are presented in the agent dialog after the learner chooses to reply with an American perspective. These markers indicate that although the agent holds a different perspective, she acknowledges the learner's perspective as valid. When the learner chooses to mirror Zahora's perspective, affirmation is not necessary, as the two already agree.

Inclusion of this strategy was driven by the Iterative Development stage in which participants first reported feeling like Zahora "didn't care about their perspective and only wanted to share her own." A likely cause for this feeling was an initial strong focus on imparting the cultural learning objective; Zahora

(as played by the role-players) was insistent on getting her point across. When affirmation was subsequently increased in the dialog, she was described as too “zen-like” and accepting, as if “she really had no opinions of her own” and therefore was not able to convincingly agree with the learner’s opinion either. I pursued an iterative development approach for this particular strategy, coupled with repeated pilot testing, to find an appropriate balance in the dialog. Affirmative statements were included in approximately one fourth of the agents’ utterances.

### 7.3.3. SELF-DISCLOSURE

The third strategy the SID model employs is self-disclosure – revealing information of a personal nature. Although self-disclosure may occur within the context of a narrative, they are not identical. Narrative may involve secondhand stories, invented stories, or official stories that tell an innocuous version of events (Schank & Berman, 2002). Self-disclosure, instead, reveals real information about oneself, family, or similarly private items (Cozby, 1973). Disclosing information to another is a behavior that communicates that we trust that person to respond appropriately. This should have the effect of generating reciprocal trust from the learner. Additionally, this strategy is hypothesized to lead to greater social bonds (Cozby, 1973). Moon showed that a computer with reciprocal, deepening self-disclosure in its conversation will cause the user to rate it as more attractive, divulge more intimate information, and become more likely to buy a product from the computer (Moon, 1998).

In the SID model, self-disclosure needed no preparation and was first introduced in the Dialog Generation stage. This strategy was followed instinctively by the role-players, who needed no prompting. Over the natural course of their dialog, they delved more deeply into each topic with each conversational turn, revealing more information from the character bible or their imagination. In the Iterative Development stage, I verified that the amount

of self-disclosure contributed by an utterance increased over the course of each dialog turn within a particular topic. In order to do this, I assigned each utterance a self-disclosure score from zero to five, and ensured that the self-disclosure score was monotonically increasing over time. The greater the number of turns, should the learner choose to continue discussing that particular topic, the greater the amount of disclosure from the agent. For example, as the learner continues with the topic of marriage, Zahora will disclose first that she has been married for three years, and later that she has no children, but hopes to someday.

#### 7.4. TASK INFORMATIONAL DIALOG MODEL

The task informational dialog (TID) model was developed as a comparison condition for the SID model, using the same structure of mini-dialogs each covering a learning objective. The informational content of each utterance was kept as close as possible to the content of the corresponding utterance from the SID model. However, it was delivered in an impersonal manner. None of the utterances contained any self-disclosure from the agent. Instead, the agent referred to the content as coming from the perspective of Iraqis in general. To this end, the narrative components were removed so that the content was decontextualized. The agent also did not make any affirmation of the learner's perspective, whether they choose to express an American perspective or any other perspective in their responses. Additionally, all personal pronouns were removed so that the agent never referred to itself, but rather to the Iraqi population.

An important note here is that the design of the dialog was careful not to advocate stereotypes by claiming that all Iraqis feel a certain way, or always perform any particular behavior. Table 7.3 shows an example of the contrast between a TID model utterance and the corresponding SID model utterance.

Instructional content	Model	Utterance
Reciprocal provision for contacts	TID	“Your contacts will give you information if you can offer them something in return. For example, translators might need protection for their families.”
	SID	“Of course, that is what friends are for. I can give you some information, and maybe you will offer my family your protection in return.”
Role of family in Iraq	TID	“Family life is very important to Iraqis. Some people feel like family are the only people that they can really trust. This is why often brothers and sisters all stay in the family business.”
	SID	“Yes, family is very important to us as well. My family are the only people that I can really trust. This is why brothers and sisters all stay in the family business. My family sells pottery.”

Table 7.3. Sample utterances for two learning objectives, contrasting the task and social dialog models.

## 7.5. CONTRIBUTIONS

The work presented in this chapter contains two main contributions. First was the formalization of a process for the generation of agent dialog. While procedures for machine-generating dialog for narrative-based agents have been described (Riedl & Young, 2003; Cassell & Bickmore, 2003), formal procedures for writing such dialog with user input have not been developed. In the process I describe, first a set of learning objectives and a backstory for the character are developed in a reciprocal manner. Second, role-players with these tools in hand are used to generate a large quantity of raw content. In the third step, this raw content is converted into an appropriate format, and rapid prototyping tools are used to present the dialog to users for iterative revision. Finally, the completed dialog is taken to subject matter experts for validation of the (in this

case cultural) content. This process can be used for the development of dialog for agents in a narrative-based learning environment, in particular in a context where building interpersonal relationships is an important consideration in the system. This chapter contains descriptions of the pitfalls and successes encountered along the way.

The second contribution of this work is the creation of a model of social informational dialog that links conversational strategies to desirable interpersonal outcomes. I identified three interpersonal outcomes that could lead to stronger social ties between an agent and a learner, potentially increasing learning results in the process. In addition, I identified three conversational strategies that are hypothesized in communications literature to lead to these interpersonal outcomes, and investigated their application in the iterative development procedure described above. This model differs from other dialog models in its instructional nature, as well as the incorporation of conversational strategies directly into the dialog rather than as separate utterances added to a task-based model. The next two chapters investigate the effects of this dialog on participants' interpersonal relationships with an agent.

## 8. STUDY 4: LEARNERS' INTERPERSONAL RELATIONS WITH AN INSTRUCTIONAL AGENT

Following development of the agent dialog, I ran a small empirical study to understand the varying effects of SID and TID on learner relations with the instructional agent, Zahora. In this study, I compared two versions of Zahora that were identical except for their dialog model. In the first, the dialog was task-focused (TID), and in the second the instructional dialog was interwoven with social conversational strategies (SID). In order to understand whether the social dialog produced the desired effects, I studied these dialog models in isolation from the rest of the instructional scenarios in BiLAT. This chapter describes the experimental measures and procedure used in the preliminary study, as well as the results on the three positive interpersonal outcomes introduced in the previous chapter (trust, entitativity, and shared perspective). Chapter 9 continues to study these outcomes in the context of the full virtual learning environment. Given the theoretical grounding behind the conversational strategies in the SID model, I hypothesized that trust, entitativity, and shared perspective with the agent would be greater in the social condition.

### 8.1. PARTICIPANTS

Thirty-nine people participated in the study (64% female). Participants were students recruited using an online subject pool from two university campuses.

Requirements were that they be U.S. citizens, over the age of 18 ( $M = 21.7$ ). Participants were paid \$10 for the half-hour session.

## 8.2. METHOD

Participants were randomly assigned to either the TID or the SID condition. They were seated at a desktop computer running the BiLAT software, wearing headphones to hear the agent responses. First, participants were given a briefing describing the role they were about to assume and introducing the character of Zahora, the interpreter. Participants were given the role of a government official preparing to do sensitive negotiations in Iraq. In both conditions Zahora was introduced as an authority on local Iraqi culture. Participants then interacted with the agent for as long as they wanted (around ten minutes on average). Finally, they took a post-interaction survey that asked them to rate various qualities of their interaction with the agent and collected demographic information.

## 8.3. MEASURES

*Shared Perspective* was measured using two seven-point Likert items (1 = strongly disagree, 7 = strongly agree) asking whether 1) participants felt they had attempted to express an American perspective with their dialog choices and 2) whether they tried to conform to an Iraqi perspective with their dialog choices. Due to computer error, only twenty-five participants' responses were recorded for these measures.

*Entitativity* was assessed with four items on a seven-point Likert scale (1 = strongly disagree, 7 = strongly agree): "This team is a unit", "The members of this team can act as one", "There is great togetherness in this team", and "We are as one" (Leach, van Zomeren, Zebel, Vliek, Pennekamp, Doosje, Ouwerkerk, & Spears, 2008) (Cronbach's (1951)  $\alpha = 0.88$ ).

*Trust* was measured through a standardized scale of trust on a seven-point scale (1 = strongly disagree, 7 = strongly agree) (Wheless, Grotz, 1977).

*Learner Characteristics* were also captured. I assessed self-reported social intelligence, a quality hypothesized to influence intercultural interactions, using a validated Social Intelligence scale (Silvera, Martinussen, & Dahl, 2001). This scale was divided into three parts, each with seven items on a 7-point Likert scale: social awareness, social information processing, and social skills.

Given that Bickmore and Cassell in their work on transactional agents (2001) found an interaction between personality and condition, where introverts were more trusting in a task condition and extroverts more trusting in a social condition, I included measures of introversion and extroversion. Extroversion and introversion were each measured by an index composed of six extrovert or introvert adjective items (Wiggins, 1979) (extrovert  $\alpha = 0.85$ ; introvert  $\alpha = 0.85$ ).

*Learner Reactions* were captured through an open-ended, post-intervention interview that asked them to express their thoughts about the character and the experience.

## 8.4. RESULTS

To investigate the models' effects on perspective, I ran a one-way ANOVA with condition as the independent variable and perspective as the dependent variable. I found a significant main effect of condition on perspective, which is shown in Table 8.1. Learners in the TID condition were more likely to state that they were espousing an American perspective in their responses to the agent than the SID condition. In the social condition, on the other hand, participants made comments like "I altered my responses to be in agreement with what she was saying" and "I mostly answered to be aligned with the Iraqi culture, to build rapport."

Scale	ANOVA	TID M(SD)	SID M(SD)
American perspective	$F(1,24)=8.87, p<.01$	4.42 (1.30)*	3.21 (0.66)
Entitativity	$F(1,24)=0.66, p=.42$	4.86 (1.02)	5.14 (1.11)
Trust	$F(1,24)=4.49, p<.05$	5.63 (1.12)*	5.00 (1.12)

Scale	ANOVA	Extro M(SD)	Intro M(SD)
Trust	$F(1,24)=4.49, p<.05$	5.62 (1.07)*	4.92 (1.16)

Table 8.1. ANOVA results comparing SID and TID on feelings of trust, entitativity, and perspective, and extroverts and introverts on trust. Items with a \* have a higher mean.

Next, I explored how participants differed in feeling entitativity with Zahora. Although the SID mean rating was higher, a one-way ANOVA showed that it was not significantly different from the TID mean. Given this trend, I also investigated correlations that may influence or be influenced by entitativity in this context. I found that participants' feelings of entitativity were strongly associated with having a shared perspective ( $r=.62, p<.001$ ); participants who were attempting to conform to the Iraqi perspective were more likely to feel like they were working together with Zahora (see Figure 8.1). Entitativity was also strongly correlated with trust ( $r=.446, p=.004$ ). Additionally, participants with a higher self-reported Social IQ, specifically on the Social Information Processing subscale, also felt significantly more like they were working on a team ( $r = .5, p < .001$ ).

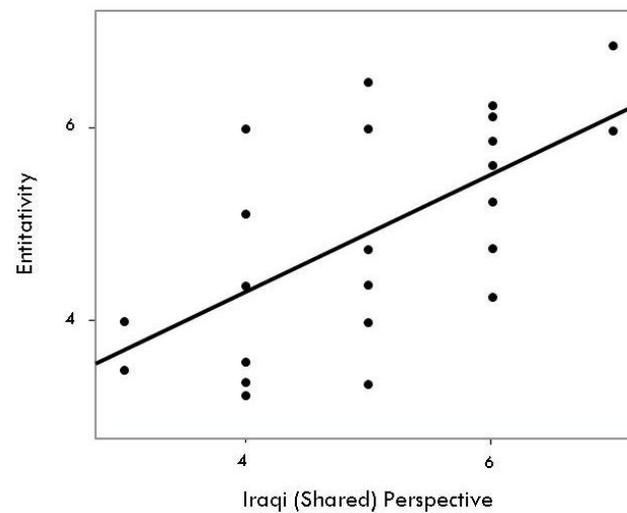


Figure 8.1. Correlation of entitativity judgments and shared perspective.

Finally, I investigated Zahora's effects on trust. Given Bickmore and Cassell's (2001) finding of an interaction between personality and condition on trust, I conducted an ANOVA with extroversion and condition as the independent variables, and trust as the dependent variable. Unlike prior results, I did not find an interaction. Instead, I found main effects of both condition and extroversion (see Table 8.1). Learners in the TID condition were more likely to trust Zahora compared to learners in the SID condition. Extroverts were also more likely to say that they trusted the agent than introverts, regardless of condition. These two effects were additive (as seen in Figure 8.2). In addition, trust was strongly correlated to the social information processing dimension of reported Social IQ ( $r=.519, p<.001$ ).

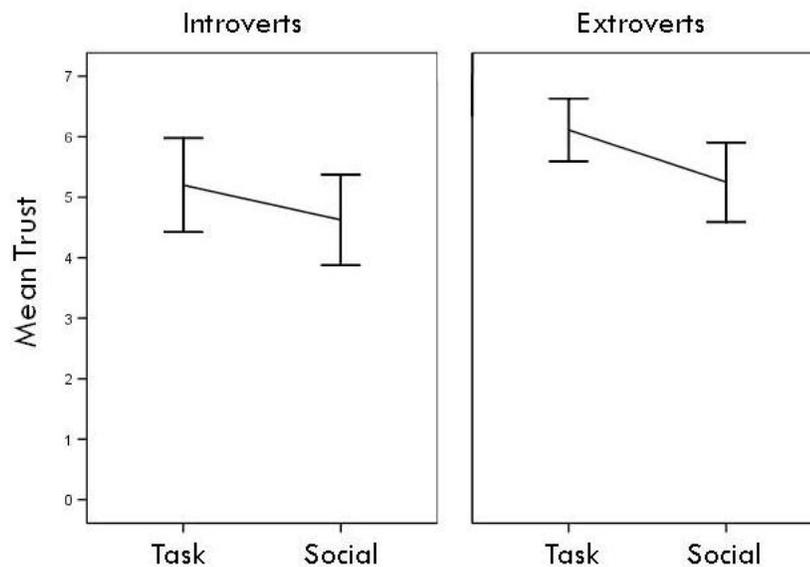


Figure 8.2. Trust judgments by condition and personality.

Qualitatively, I was concerned that the interaction with the dialog menus would feel forced, but participants generally found it to be fairly natural. One participant commented that “It wasn’t always what I would have typed in myself, but I usually felt like there was one answer more applicable to what I wanted to say”. Additionally, I wanted to explore whether the context and content felt realistic. One participant of Arab background remarked that it “felt true to Arab culture”, and a participant who had previously been deployed to Iraq said that the “options seemed realistic.” Participants in the task condition had comments like, “This seemed like more of a work meeting,” and even in one case, “I wanted more social interaction.”

## 8.5. DISCUSSION AND CONTRIBUTIONS

In Study 3, I looked at how introducing social conversational strategies into instructional dialog affects interpersonal relations with virtual agents.

A main contribution of this work is the finding that incorporating social conversation strategies (namely, narrative, self-disclosure and affirmation) into a

pedagogical agent's instructional dialog (SID) has a significant effect on the perspective that learners desire to present to an agent. With respect to intercultural education, this is a significant win. In general, language and culture instructors want learners to both understand what the target cultural perspective is and be willing to (at least) temporarily consider and assume that perspective. Even beyond intercultural education, the ability to take on a shared perspective may be highly valued; e.g., in STEM education, one central objective is for students to be able to see themselves as scientists or mathematicians. I also found that those who demonstrated a shared perspective with Zahora had significantly higher ratings of entitativity - they felt like they were working together on a collaborative team. The increased entitativity not only reflects positive interpersonal relations with the agent, but may also have an effect on learning as well. As noted previously, there are many studies that report positive effects of groups on learning.

On the other hand, learners in the task instructional dialog (TID) condition expressed more trust in the agent. While this finding was not in line with my predictions, it may possibly be explained by the authoritative feel of the task dialog. In TID, the information given was not contextualized or taken from personal experience, but presented as a general representation of the Iraqi culture. Therefore learners may have given more credence to the authoritative presentation of this information. The finding that social conversational strategies may lessen trust is an important contribution because many systems are beginning to employ some of these strategies, such as narrative, for educational purposes. What the agents in these systems gain in terms of positive interpersonal relations with the learner, they may lose if the learner does not trust the learning content they are given. In the end, a balance between TID and SID might yield the strongest way forward. The conditions that lead to these results are discussed further in Chapter 9, in the context of Study 5.

Additionally, the study shows a difference from related prior work in that extroversion had a positive effect on trust in both social and task dialog, rather than interacting with condition. This result may be due to our model's infusion of social conversational strategies into all of the agent utterances, rather than the addition of social utterances to already existing task dialog. It is possible that in prior work looking at adding social elements to transactional dialog, introverts reacted negatively to the length of the dialog – they may have felt the social agent was simply talking too much. Or these effects may be related to the inclusion in SID of conversational strategies not found in this prior work.

Extroversion was not the only effect of personality on learner-agent relations. In fact, in this study I also found that participants who rated themselves highly on social intelligence reported a higher entitativity rating – they felt more strongly that they were working together with the agent – and a higher rating of trust. These findings were specific to the social information processing scale, which contains items like, “I can predict how others will react to my behavior,” and “I understand other peoples' feelings.” With respect to intercultural communication, this result is not unexpected. In previous work, I have seen that predicting behavior is a good indicator of intercultural ability (Ogan, Alevan & Jones, 2010). These findings contribute to evidence that personality may have a significant impact on learners' interpersonal relations with virtual agents, and that systems should be able to take this into account.

A critical avenue of research moving forward is to investigate the mediated relationship between conversational strategies, learner-agent interpersonal relations, and learning results. While in Chapter 7 I hypothesized, based on related literature, that positive interpersonal relations will lead to greater learning, the study discussed in this chapter does not address this question. Moving forward, it is important to understand whether these interpersonal findings hold true in the full context of a virtual learning environment, and subsequently how they affect learning. Study 5, presented in the next chapter,

seeks to answers these questions. Additionally, to further explore several of the hypothesized results that had positive trends but not significance in this study, an immediate need filled in Study 5 is to increase the number of participants. Understanding how these strategies and in turn, interpersonal relationships, relate to learning will enable the creation of agent dialog with real benefits for education.

## 9. STUDY 5: COGNITIVE AND INTERPERSONAL EFFECTS OF SOCIAL INFORMATIONAL DIALOG

Earlier results indicated that participants with social goals for interacting with the agents in BiLAT were more efficacious learners (Study 3). However, explicitly presenting social goals was found to be unsuccessful for modifying students' goals. Thus, in Study 5, I aimed to positively influence students' social orientation through properties of the agents themselves. However, the agents that existed in BiLAT were not good candidates for modification; they were designed to simulate authentic Iraqi meeting partners. Modifying these agents to be more social would remove the authenticity of the experience and leave students underprepared for transfer to a real-world negotiation context.

On the other hand, a new agent, encountered early in the experience and imbued with social qualities, may be able to change learners' orientation towards the rest of the environment and the agents in it. The results from Study 4 showed that social dialog with an agent in BiLAT has significant effects on learners' perceptions of the agent compared to a task-based control. Learners changed their perspective to share the agent Zahora's, in the process gaining a greater sense of entitativity. What is not known yet is whether the social orientation that Zahora helps to induce carries over in the students' subsequent encounters with the other agents in the simulation. If this agent can change students' orientation towards the other agents in the game to be more social, this may also increase their learning.

Using the newly developed agent Zahora, I conducted a study to investigate whether the social effects seen in Study 4 persisted within the full context of

BiLAT, and whether they were sustained across the other agents encountered in the scenario. The study was also intended to uncover a link between these interpersonal outcomes and greater learning. In addition, given the current lack of understanding of the specific educational benefits of virtual environments for intercultural training, I compared the use of BiLAT, which is focused on practice of social skills, to an ecological control condition in which no practice occurs. In this comparison group, learners worked on a reading activity such as they might typically be given in a classroom that teaches intercultural negotiation. This chapter describes the results of this study and the implications for social design of pedagogical agents.

## 9.1. SOCIAL CONSTRUCTS

This chapter draws together many of the social constructs that have been explored elsewhere in this thesis. As in Study 3, I elicited social goals by asking students to write out their goals for the interaction with the agents. In this study, I additionally use a validated social goals survey to uncover learners' goals. Most surveys of this type have been developed for use with children. Instead, using a questionnaire developed by McCollum (2006) that is specifically aimed at college students, this study is the first to take a quantitative look at the social goals learners have for interacting with agents. As in Study 4, this study investigates the influence of agent dialog on trust, entitativity, and shared perspective. Additionally, I look at the effects of agent dialog on presence, the subjective experience of being in one place or environment (Witmer & Singer, 1998).

Moreover, rather than looking at how agent dialog affects these social constructs individually, I take a more integrative approach. In the data analysis, I use structural equation modeling to find connections between the different social constructs. I also look at how individual characteristics and attitudes

affect the constructs, and investigate how the dialog influences these constructs as it relates to other agents in the environment. Using this approach, it is possible to take into account the complexity of the learner's context and look at how the many variables affecting the learner's interactions and performance interact with one other.

## 9.2. EXPERIMENTAL DESIGN

This study compared three conditions to understand effects on interpersonal relations with agents and learning. In the two conditions that employed the BiLAT environment, participants used the standard software described in Chapter 3, with the addition of the character Zahora described in Chapter 7. The task condition interacted with a version of Zahora configured with task-based instructional dialog (*TID*). In the social condition, Zahora used the social instructional dialog (*SID*) model as a basis for her dialog. Unlike Study 4, participants then continued on to negotiations with other agents in BiLAT.

The third condition was an ecological control group (*control*) who read a document pertaining to the same information that learners practice in BiLAT, taken from "Through the lens of cultural awareness" (Wunderle, 2007). This document was written for a military audience by one of the experts who was most influential in the development of BiLAT, and is similar to material students might study in an intercultural negotiation course in the Army.

I had three main hypotheses:

- H1. Social and Interpersonal Perceptions:* The use of *SID* in a virtual agent influences participants' perceptions of the agent to be more social compared to *TID*-based dialog.
- H2. Social Goals:* The use of *SID* increases learners' social goals for interacting with an agent compared to *TID*-based dialog.

*H3. Learning:* Use of a social simulation increases learning gains over reading a high-quality text on the subject of intercultural negotiations, with *SID*-based agents increasing learning over *TID*-based agents.

### 9.3. PARTICIPANTS

96 people participated in the study (53% female). Participants were recruited using an online subject pool from two university campuses. The stipulated requirements were that participants be between 18-25 years of age, and U.S. citizens. Participants were paid \$40 for participating, which was expected to last around 2.5 hours. Four participants were dropped due to technical errors, and three others due to missing data. Therefore 89 participants are included in the analyses presented in this chapter.

### 9.4. MEASURES

The measures used in this study were taken from assessments used in previous studies in this dissertation, as well as additional measures that investigated learners' social orientations towards the virtual agents, feelings of presence in the environment, and liking of the activity. Furthermore, I measured learner characteristics that have been hypothesized as antecedents to intercultural competence (Abbe, Gulick, & Herman, 2008).

#### 9.4.1. LEARNER CHARACTERISTICS

The review paper by Abbe, Gulick, and Herman (2008) has identified three main characteristics of practitioners that affect their effectiveness in cross-cultural settings: dispositional, biographical, and those relating to self and identity. They report, however, that research is needed to address the extent to which these dimensions and their components are responsive to education and training. Therefore, in this study I investigated antecedents described by

Abbe et al. (2007) and additional characteristics which I believed would have the most potential to affect learners' progress in a social simulation.

With respect to dispositional variables, I used the extroversion and introversion scales (Wiggins, 1979), as described in Section 8.3. Although not described by Abbe et al. (2007), I also measured the other commonly studied goal orientations that may impact how learners approach a learning environment and interact with social goals: performance and mastery goals. These I assessed using the Goal Orientation Scale (Elliot & Murayama, 2008). In this scale, performance and mastery orientations are each divided into approach and avoid goals. For instance, students may take a positive approach towards mastery, trying to learn everything they can about a topic. On the other hand, students may take an avoidance stance towards performance, attempting to avoid receiving a low grade (Schunk, Pintrich, & Meece, 2007).

Although Abbe et al. (2007) list characteristics like "cultural intelligence" as skills, I believe that these self-report measures may in fact be assessing dispositions towards engaging in social or intercultural relations. Thus I ask participants to complete the social information processing and the social awareness sections of the Tromso Social Intelligence scale (Silvera, Martinussen, & Dahl, 2001), and a shortened form of the Cultural Intelligence scale (Ang et al., 2007).

With respect to biographical variables, I asked participants to report on their age, gender, GPA, knowledge of Arab cultures, and number of languages spoken. The context of the virtual learning environment is what makes this work unique from the studies compiled by Abbe et al. (2007). Therefore, I also investigated learners' prior experience with games (and in particular, first-person shooters, which use the same interface elements as BiLAT), programming, and negotiation, the secondary content domain of the learning environment.

With respect to identity, I asked participants to report on their citizenship and prior experience with foreign cultures as described in Chapter 6.

#### 9.4.2. ATTITUDES

Given the immersiveness of the military context of BiLAT, learners' attitudes towards the military might affect their interactions with agents or their learning. Reciprocally, interactions with the environment, whether positive or negative, may affect learners' perceptions of the military. Therefore, an attitude assessment was given both pre- and post-interaction. I measured attitudes towards the military with a three-item scale taken from Newcity (1997). Each item was given on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree). Items are shown in Table 9.3. These questions are *projective*, in that they ask about other peoples' feelings; this technique of assessing attitudes has been developed by clinical researchers to allow participants to protect and defend their ego by externalizing sensitive thoughts and feelings, ascribing them instead to other individuals or inanimate objects (Donoghue, 2000; Kline, 1983).

The cultural context of BiLAT is also clearly very salient. Similarly, interactions with the agents or learning might be affected by learners' a priori attitudes towards people of Arab descent. In turn, interactions in the environment might affect learners' attitudes towards the Arab culture. I measured pre- and post-attitudes towards Arabs with a three-item scale taken from the Subtle Prejudice Scale (Pettigrew & Meertens, 1995) using the same 7-point Likert scale. These items are also shown in Table 9.3.

#### 9.4.3. MANIPULATION CHECK: PERCEPTIONS OF ACTIVITY AND DIALOG

To gain a broader understanding of the learner perceptions of social simulations compared to typical classroom activities, participants were asked

whether the activity was fun, engaging, natural, and tedious. These items were given on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree).

In order to determine whether learners perceived the TID and SID dialog differently, they answered four manipulation check items (“Zahora was focused on the task”, “Zahora did not make much small talk”, “Zahora shared personal stories”, “Zahora was very social”). These items were given on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree).

#### 9.4.4. SOCIAL AND INTERPERSONAL MEASURES

As in Study 3, I measured learners’ social goals for the interactions with the agents. Participants wrote out their goals for meeting with the characters in free text, as described in this previous study. In the current study, I also introduced a quantitative measure of learners’ social goals for meeting with Zahora by using a short form of the Social Goals Questionnaire (McCollum, 2006). This scale was developed and validated specifically for university-age students and focuses on five subscales that were found to be particularly salient in that age group: Social Responsibility, Social Attractiveness, Power, Receiving Assistance, and Belongingness. Items are measured on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree).

As described in Study 4, I measured learners’ trust, entitativity, and shared perspective with the agents.

Presence, the “subjective experience of being in one place or environment”, is a frequently-measured construct in virtual environments that investigates how immersed learners feel (Witmer & Singer, 1998). Because feeling greater social presence is believed to positively influence learners’ perceptions of embodied agents (Lee, Jung, Kim & Kim, 2006), I included a social presence scale with five items: My onscreen workspace felt like a real place, I felt like I could see my meeting partners, I felt the presence of my meeting partners, When we were

talking, it felt like we were in a real meeting, I felt like my meeting partners and I were in the same place. (1 = strongly disagree, 7 = strongly agree).

Finally, other work on agents (e.g., Cassell & Bickmore, 2003; Kumar, Ai, Beuth, & Rosé, 2009) has included a set of individual social relationship measures to give a more complete picture of how participants perceive the agent. I used the same measures on 7-point Likert scales: I like [agent], I feel like [agent] knows me, I feel like I know [agent], where agent was replaced with the name of each virtual agent the learners encountered in the environment. (1 = does not describe me, 7 = completely describes me).

#### 9.4.5. LEARNING

Learning of the cultural knowledge components was assessed through two measures used in the previous experiments: the Culture Assimilator (Cushner & Brislin, 1995), and the Situational Judgment Test (Legree & Psozka, 2006), as described in Section 6.3.2.

Change in attitudes towards Arab cultures, using the survey described in Section 9.2.2, was used as an assessment of *savoir-être*, or openness towards new cultures. *Savoir-être* is a learning objective of interest for intercultural competence as described in Chapter 2.

In this study I introduced an in-game, behavioral assessment of learning by having all participants negotiate with a character named Aziz, a hospital administrator who is part of a different scenario. To investigate learning gains in this assessment, I measured the negotiation objectives achieved and the number of errors committed in this negotiation. Errors are defined by the intelligent tutoring system which is described in Section 3.6, and may include taking a culturally inappropriate action, taking an appropriate action at an inappropriate time (known as a phase error), or other missteps.

## 9.5. PROCEDURE

Step	BiLAT – TID	BiLAT – SID	Control
1	Demographics + Attitudes Survey	Demographics + Attitudes Survey	Demographics + Attitudes Survey
2	PowerPoint Intro	PowerPoint Intro	PowerPoint Intro
3	SJT + CA Pretest	SJT + CA Pretest	SJT + CA Pretest
4	<b>Zahora – TID</b>	<b>Zahora - SID</b>	<b>Reading</b>
5	Zahora Survey	Zahora Survey	“
6	Farid Prep <b>+ Zahora message</b>	Farid Prep <b>+ Zahora message</b>	“
7	Farid Meeting	Farid Meeting	“
8	Hassan Prep <b>+ Zahora message</b>	Hassan Prep <b>+ Zahora message</b>	“
9	Hassan Meeting	Hassan Meeting	“
11	SJT + CA Posttest	SJT + CA Posttest	SJT + CA Posttest
10	Attitudes + Social Survey	Attitudes + Social Survey	Attitudes Survey
12	In-game posttest	In-game posttest	In-game posttest

Table 9.1. Study procedure for the three conditions: TID, SID, and the reading control. Bold indicates where the conditions differ.

All participants took the demographics questionnaire online prior to arriving at the study. The rest of the study took place in a user studies laboratory. Once they arrived at the lab, participants were randomly assigned to condition and seated at a desktop computer, such that they could not see each other's screens. The experimenter sat in a control room with a one-way mirror from which the participants' computer screens could be seen. For 10 minutes, participants in all three conditions reviewed a set of PowerPoint slides that contained material relevant to the learning objectives of BiLAT. Afterwards, they took the Culture Assimilator and Situational Judgment Test online. At this

point, the control group was instructed to begin reading the document “Through the Lens of Cultural Awareness” on the computer for an hour.

As in Study 4, participants in the two BiLAT conditions were given a briefing describing the role they were about to assume and introducing the character of Zahora, the interpreter. In both conditions the character was introduced as an authority on local Iraqi culture. The difference between the two conditions was the model that was the basis of her dialog (TID or SID). Participants interacted with the agent for as long as they wanted (around ten minutes on average). They then took a survey that asked about their interaction with Zahora. In particular, this survey asked about the perspective taken in their responses, their social goals for interacting with Zahora, and the manipulation check questions.

Following this survey, participants read background information online about the scenario and the character (Farid) they were about to encounter. While they were engaged in reading, the experimenter sent them a chat message from an AOL Instant Messenger account named Zahora, which included an image of the agent (see Figure 9.1 for a screenshot). Given the results of Study 3 (simple prompting does not induce a stronger social orientation), this message was sent to reinforce the manipulation. In the SID condition, the chat message was developed using the SID model: “Farid may be difficult to impress, but he is a friend. He wants to work on our team.” In the TID condition, the chat reinforced the informational nature of the relationship: “Farid is a police officer. Iraqi police often have too much to do and too few resources.”

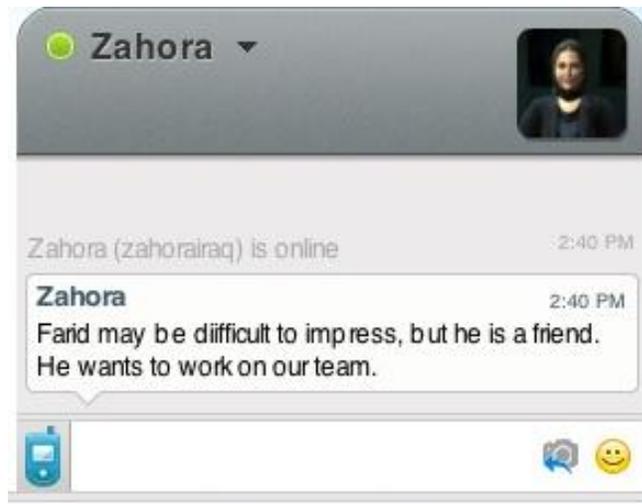


Figure 9.1. Screenshot of the message participants received from Zahora.

Various mechanisms were explored for delivering this message from Zahora, from the experimenter bringing in a personal note on paper, to reconfiguring the BiLAT infrastructure to allow popup messages from previously encountered characters. In the end, Instant Messaging was determined to be the best compromise, without requiring extensive software modifications while still feeling reasonably integrated with the technology. It was also a format for interaction that most college-aged students are comfortable using and needs no instruction. However, the use of instant messages had an unintended consequence; many participants chatted back to Zahora. When this happened, the experimenter replied with a standard message that was intended to close the conversation: “Your officers should be able to help you further. Refer to the background material.”

After they finished reading, participants met with Farid as in Study 4. When they accomplished their goals with Farid, they repeated the process with the following character Hassan, including background reading, a message from Zahora, and meeting. Immediately after a successful negotiation with Hassan, or at the end of one hour, they were finished with the instructional exercise and

took the assessments that asked about their social experiences with all three of the characters.

At this point, participants from all conditions took the post-assessments of learning, and then went to the BiLAT environment to do an in-game assessment, which involved negotiating with a new character named Aziz from a different scenario. In order to complete this in-game assessment, the students in the control condition were given the same instruction on how to navigate the environment that the BiLAT conditions received when they first entered the game. Following the in-game assessment, participants received their payment and an explanation of the goals of the study. Prior to leaving, they were asked if they had any comments or questions regarding the study or the interaction with Zahora in particular. The chronological study procedure is shown in Table 9.1.

## 9.6. RESULTS

### 9.6.1. SCALE RELIABILITY AND RANDOMIZATION CHECKS

Before investigating my core hypotheses, there were several preliminary data analysis steps that needed to be taken. First, for both learner characteristics and attitudes, I checked the reliability of the scales, and whether there were any prior differences between conditions. In Section 9.6.5, I look at how these learner characteristics and attitudes, in combination with condition, affect the social and learning constructs of interest. Second, I performed a basic check on the effectiveness of the experimental manipulation, looking at how the perceptions of the activity varied between the BiLAT and control conditions, and how the perceptions of agent dialog varied between the *SID* and *TID* conditions. These analyses are described in this section.

## 9.6.1.1. A PRIORI DIFFERENCES IN LEARNER CHARACTERISTICS

For each of the learner characteristic scales, the reliability of the combined scale was checked with a test of Cronbach's alpha (1951). In general, a scale with a Cronbach's alpha score above .7 is considered to be reasonably reliable, which allows the individual items to be collapsed into one measure. In Table 9.2 I report the alpha values for each scale, as well the results of a between-condition ANOVA to determine whether there were any a priori differences in characteristics.

Each scale had reliability over .7, and therefore, in the rest of the results section, items in Cultural Intelligence, Social Intelligence, extroversion, and each of the learning goals scales have been collapsed into one measure per scale. The ANOVAs indicate that there were no a priori differences among conditions for any of these measures.

Additional ANOVAs uncovered no between-condition differences in age, gender, knowledge of Arab cultures, number of languages spoken, GPA, or prior experience with games, programming, negotiation, and foreign cultures (all  $p > .2$ ).

In Section 9.6.5, I examine whether these learner characteristics had an impact on social or learning outcomes.

<b>IQ Scale</b>	<b># items</b>	<b><math>\alpha</math></b>	<b>ANOVA</b>	<b>M control</b>	<b>M TID</b>	<b>M SID</b>
Cultural Intelligence	8	.858	$F(2,88)=.62, p=.53$	4.82(.98)	4.85(.76)	5.05(.79)
Social Intelligence	14	.824	$F(2,88)=.45, p=.64$	4.91(.59)	4.88(.73)	4.76(.62)

<b>Extroversion Scale</b>	<b># items</b>	<b><math>\alpha</math></b>	<b>ANOVA</b>	<b>M control</b>	<b>M TID</b>	<b>M SID</b>
Extrovert	6	.862	$F(2,88)=.18, p=.83$	4.58(1.02)	4.72(1.04)	4.72(1.09)
Introvert	6	.815	$F(2,88)=.01, p=.99$	3.74(1.09)	3.72(1.11)	3.77(.98)

Learning Goals Scale	# items	$\alpha$	ANOVA	M control	M TID	M SID
Mastery approach	3	.853	$F(2,88)=.39, p=.68$	5.67(1.07)	5.48(1.01)	5.42(1.24)
Mastery avoid	3	.792	$F(2,88)=.46, p=.63$	4.54(1.42)	4.52(1.40)	4.21(1.63)
Performance approach	3	.878	$F(2,88)=.32, p=.73$	5.30(1.14)	5.31(1.07)	5.10(1.23)
Performance avoid	3	.891	$F(2,88)=.76, p=.47$	5.02(1.51)	4.61(1.42)	4.66(1.29)

Table 9.2. ANOVA results comparing all three conditions on a priori learner characteristics (1 = strongly disagree, 7 = strongly agree). Cultural Intelligence includes Motivational, Knowledge, Behavioral, and Metacognitive factors with 2 items each. Social Intelligence includes Social Awareness and Social Information Processing factors with 7 items each. All scales were at a sufficient reliability, with no significant differences between conditions.

9.6.1.2. A PRIORI DIFFERENCES IN ATTITUDES

For each of the learner attitude scales, the reliability of the scale was checked with a test of Cronbach’s alpha (1951). However, neither of the scales achieved a reasonable reliability (Arab attitudes  $\alpha =.346$ ; Military attitudes  $\alpha =.441$ ). Therefore, it appears that these items are measuring different concepts in this population, and will be used separately in subsequent analyses. None of the individual items showed any a priori differences on a between-condition ANOVA, as shown in Table 9.3.

Arab Attitudes Item	ANOVA	M control	M TID	M SID
1. Arabic people are similar to myself in the values they teach their children.	$F(2,88)=.03, p=.97$	4.14(1.13)	4.13(1.33)	4.20(1.25)
2. Many other groups have come to America and overcome prejudice and worked their way up. Arabic people should do the same without special favor.	$F(2,88)=1.1, p=.33$	3.69(1.17)	4.23(1.55)	4.00(1.46)
3. Arabic people living here teach their children values and skills different from those required to be successful in America.	$F(2,88)=.01, p=.99$	4.66(1.11)	4.70(1.26)	4.70(1.29)

Military Attitudes Item	ANOVA	M control	M TID	M SID
1. Most members of civilian society have a great deal of respect for the military.	$F(2,88)=.39, p=.68$	4.79(1.15)	4.63(1.30)	4.50(1.36)
2. I would be disappointed if a child of mine joined the military.	$F(2,88)=.96, p=.39$	4.14(2.00)	3.90(1.95)	3.47(1.72)
3. Most members of the military have a great deal of respect for civilian society.	$F(2,88)=.15, p=.86$	3.90(1.37)	3.83(1.49)	3.70(1.32)

Table 9.3. ANOVA results comparing all three conditions on attitudes towards Arab attitudes and the U.S. military (1 = strongly disagree, 7 = strongly agree). For the reliability check only, Questions 2 and 3 of the Arab scale and Question 2 of the military scale were scored in reverse for consistency with the other items. The scales did not achieve sufficient reliability.

### 9.6.1.3. PERCEPTIONS OF THE ACTIVITY

To investigate general impressions of the activity, Table 9.4 compares participants' responses using an ANOVA with condition as the between-subjects variable. According to Tukey's post-hoc pairwise comparison, both the social and task BiLAT conditions found the activity significantly more fun and more engaging than the control group who read the document. However, there were no significant differences between the social and task group. Fun and engagement were highly correlated with one another ( $r(84)=.751, p<.001$ ).

Although many participants in the control group commented on how boring they thought the reading activity was, there were no significant differences between groups in how tedious or how natural the activity felt. Finding the activity tedious was negatively correlated to both finding it engaging ( $r(84)=-.340, p=.001$ ) and fun ( $r(84)=-.316, p=.003$ ), while controlling for condition.

Item	ANOVA	M control	M TID	M SID
Activity was fun	$F(2,88)=17.74, p<.001$	3.53 (1.43)	5.43(.96)*	4.97(1.35)*
Activity was engaging	$F(2,88)=13.16, p<.001$	4.03(1.75)	5.71(.80)*	5.55(1.41)*
Activity was natural	$F(2,88)=1.5, p=.229$	3.93(1.53)	4.54(1.14)	4.45(1.61)
Activity was tedious	$F(2,88)=.52, p=.599$	4.70(1.60)	4.32(1.42)	4.35(1.72)

Table 9.4. ANOVA results comparing all three conditions on perceptions of the activity (1 = strongly disagree, 7 = strongly agree). Items with a \* indicate a statistically higher mean.

#### 9.6.1.4. MANIPULATION CHECK

To verify that learners indeed found the dialog with Zahora to have different qualities in the two conditions, I asked four manipulation check questions listed in Table 9.5; two related to the task-based nature of the dialog, and two confirming the social nature of the dialog. Table 9.5 also contains the results of a between-subjects ANOVA for each of these questions, with TID vs. SID as the independent variable. There was a significant difference between conditions on ratings for all four questions. The SID condition found Zahora to be more social, while the TID condition found her to be more task-focused.

Item	ANOVA	M TID	M SID
1. Zahora shared personal stories	$F(1,59)=50.37, p<.001$	3.40 (1.57)	5.90 (1.16)*
2. Zahora was very social	$F(1,59)=26.34, p<.001$	5.10 (1.30)	6.45 (0.68)*
3. Zahora focused on the task	$F(1,59)=13.85, p<.001$	4.93 (1.28)*	3.61 (1.48)
4. Zahora did not make much small talk	$F(1,59)=14.92, p<.001$	3.63 (1.43)*	2.13 (1.61)

Table 9.5. Between-condition differences on manipulation check questions (1 = strongly disagree, 7 = strongly agree). Items with a \* indicate a statistically higher mean.

#### 9.6.2. HI: SOCIAL AND INTERPERSONAL MEASURES

Next, investigating *HI*, I looked at whether social dialog increased the positive social and interpersonal perceptions of the learners towards Zahora. I examined four measures: shared perspective, entitativity, trust, and presence. I tested the influence of condition on these measures individually, but also the relationships between the measures with a Structural Equation Model.

##### 9.6.2.1. PERSPECTIVE IN RESPONSES TO ZAHORA

Participants were asked to report on the perspective they were trying to demonstrate in their conversation with Zahora. Table 9.6 contains the results of a between-subjects ANOVA with TID vs. SID as the independent variable. Participants in the social condition were significantly more likely to state that

they were attempting to show an Iraqi perspective. Participants in the task condition were significantly more likely to claim they were attempting to show an American perspective. In order to understand why this might be, I investigated learner characteristics that were associated with taking an American perspective. An American perspective was negatively correlated with foreign experience ( $r(59)=-.345$ ,  $p=.007$ ) and also with believing that Arabs hold similar values to you ( $r(59)=-.285$ ,  $p=.029$ ).

Scale	ANOVA	M TID	M SID
American perspective	$F(1,60)=14.78$ , $p<.001$	4.53 (1.50)*	3.06 (1.48)
Iraqi perspective	$F(1,60)=12.30$ , $p=.001$	4.47 (1.70)	5.84 (1.34)*
Zahora: Entitativity	$F(1,59)=11.79$ , $p=.001$	4.53 (1.10)	5.53 (1.14)*
Farid: Entitativity	$F(1,59)=1.35$ , $p=.25$	4.32 (1.36)	4.76 (1.53)
Hassan: Entitativity	$F(1,59)=4.42$ , $p=.04$	2.66 (1.04)	3.38 (1.54)*
Zahora: Trust	$F(1,59)=6.18$ , $p=.016$	5.03 (1.30)	5.81 (1.11)*
Farid: Trust	$F(1,59)=0.35$ , $p=.792$	4.72 (1.62)	4.62 (1.63)
Hassan: Trust	$F(1,59)=4.68$ , $p=.035$	2.10 (1.14)	2.94 (1.75)*

Table 9.6. Between-condition differences on perspective in response to Zahora, and entitativity and trust with each character. Items with an asterisk indicate a higher mean.

### 9.6.2.2. ENTITATIVITY

After meeting with the three agents, participants were asked to report on how strongly they felt entitativity with each character. Table 9.6 above contains the results of a between-subjects ANOVA comparing TID vs. SID. Confirming H1, participants in the social condition were significantly more likely to claim they identified with Zahora as a team. Continuing the investigation to examine how the differences in Zahora's dialog affected learner perceptions of the other agents in the environment, participants in the social condition were also significantly more likely to state that they identified with Hassan as a team.

However, although means in the social condition were higher, the difference in their entitativity ratings with Farid did not reach statistical significance.

#### 9.6.2.3. TRUST

Participants were also asked to report how much they trusted each character. Table 9.6 above contains the results of a between-subjects ANOVA comparing TID vs. SID. Participants in the social condition were significantly more likely to claim they trusted Zahora. Learners in the SID condition also rated their level of trust in Hassan significantly higher than those in the TID condition. However, their ratings of trust in Farid did not differ.

#### 9.6.2.4. PRESENCE

Participants rated their feelings of presence in the environment. The results of a between-condition ANOVA are shown in Table 9.7, with no significant differences between conditions on feelings of presence.

Scale	# items	$\alpha$	ANOVA	M TID	M SID
Meeting Presence	5	.887	$F(1,59)=.43, p=.515$	4.83(.98)	4.62(1.40)

Table 9.7. Reliability and between-condition comparison for Meeting Presence.

#### 9.6.2.5. MODELING INTERPERSONAL OUTCOMES

The ANOVA comparisons described in the previous sections investigated the effects of the dialog on individual interpersonal outcomes. Because these outcomes have theoretical similarities and may have implicitly influenced one other, I used Structural Equation Models (SEM) to elucidate the underlying structure of the dialog effects (Bollen, 1989). SEM models each selected variable as a linear function of its immediate causes, and independent Gaussian

noise (Bollen, 1989; Spirtes, Glymour, & Scheines, 2000). I used the SEM algorithm implemented in the software package called Tetrad to search for models consistent with our background knowledge. In this case, the background knowledge provided to the algorithm was that condition preceded all of the other effects; i.e., none of the interpersonal effects could have influenced what condition participants were in. The structural model generated by Tetrad ( $\chi^2(26)=35.39, p=.1$ )<sup>1</sup> and shown in Figure 9.2 revealed several significant relationships. Foremost, the model shows that a participant's strength of entitativity from their initial interactions with Zahora carried forward to subsequent interactions with Farid and Hassan, regardless of condition. Thus, if a participant had a high level of entitativity with Zahora, they were likely to have a higher level of entitativity with Farid and Hassan. Further, the model shows that entitativity had a positive effect on trust. In other words, the more a participant felt like they are working on a team with a virtual agent, the more trust they had in that character. This relationship held true for all three agents. Finally, the perspective that participants took in the interaction, while influenced by condition, did not seem to affect either their level of entitativity or trust.

---

<sup>1</sup> In SEM, the  $p$ -value reflects the probability that the deviance between the implied covariance matrix (at the maximum likelihood estimate) and the observed covariance is as big or bigger than observed. Thus, higher  $p$ -values indicate that the model fits the data well (Bollen, 1989).

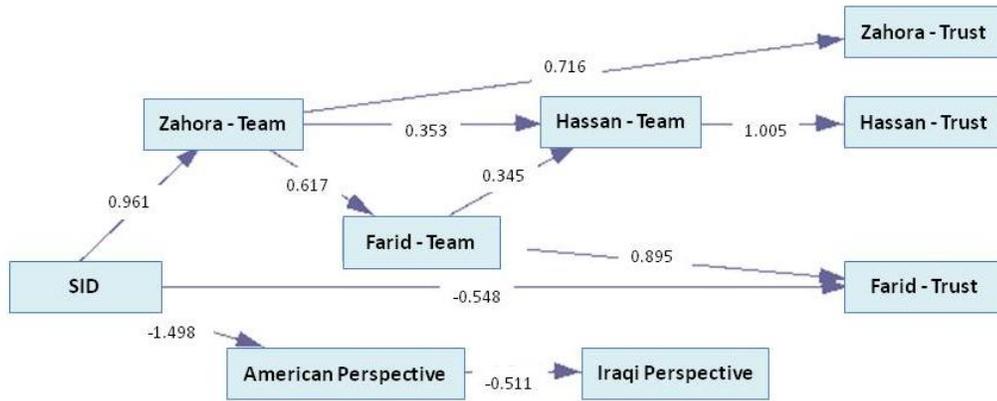


Figure 9.2. Tetrad model showing relationships between condition (labeled “Social”) and interpersonal outcomes. Black numbers indicate the values of the model parameters (the strength of the relationship), while green numbers indicate the overall means for the variable.

9.6.3. H2: SOCIAL GOALS

Next, I investigated *H2*, which hypothesized that the SID condition would increase learners’ social goals. In this section I look at survey responses to social goals, self-reported social goals, and spontaneous social behaviors in the system.

9.6.3.1. SOCIAL GOALS SCALE

Each of the five social goals subscales used in this study was composed of three items. Table 9.8 shows the results of a reliability analysis for each subscale.

Social Goals Scale	# items	$\alpha$
Social Responsibility	3	.554
Social Attractiveness	3	.430
Power	3	.634
Receiving Assistance	3	.599
Belongingness	3	.777

Table 9.8. Cronbach’s alpha values of reliability for each social goals subscale.

The low alpha value for many of the subscales indicates that these particular subscales may not be reliable in this population. Therefore, I performed a Principle Component Analysis with Varimax rotation using all of the items to see whether a different division of the items into subscales had more reliability for this population.

As expected, the Principle Component Analysis (shown in full in Appendix C) showed a different division of items into subscales. Component 1, with items pertaining to Zahora's feelings of trust, acceptance, and liking of the participant, was renamed Acceptance. Component 2 was composed of the items originally in the Power subscale, as well as an item referring to high social status, and therefore kept the title Power. Component 3, renamed Belongingness, was composed of all items that contained a reference to being a member of a team. Component 4 contained only the items originally in the Assistance subscale. Table 9.9 shows the reliability analysis of the new social goals factors. Additionally, it shows the results of a between-subjects ANOVA comparing the two BiLAT conditions, using the new scores for each scale that were generated by the PCA. The social condition had significantly higher ratings desiring Acceptance from Zahora. On the other hand, the task condition had significantly higher ratings of goals for Assistance from Zahora. With Power and Belongingness, no significant differences appeared.

Scale	# items	$\alpha$	ANOVA	<i>M</i> task	<i>M</i> social
1: Acceptance	5	.821	$F(1,59)=32.74, p<.001$	-.60 (.96)	.58(.62)*
2: Power	4	.688	$F(1,59)=.501, p=.482$	.09(1.04)	-.09(.97)
3: Belongingness	3	.765	$F(1,59)=.001, p=.976$	.00(.58)	-.00(1.30)
4: Assistance	3	.599	$F(1,59)=7.48, p=.008$	.34(.66)	-.33(1.16)*

Table 9.9. Reliability and between-condition comparison for each new Social Goals scale. New scores on each scale are centered around zero and range from -1 to 1. Items with a \* indicate a higher mean.

## 9.6.3.2. WRITTEN SOCIAL GOALS

As in Study 3, participants wrote out their goals for interacting with the two agents with whom they negotiated. Two independent raters coded each goal written by a participant, and gave a point to each goal deemed to be social (i.e., not purely task-related). The results of a between-condition ANOVA is shown in Table 9.10. In this experiment, there were no significant differences between conditions on reported goals for either of the characters, Farid or Hassan. In general, reporting a social goal for meeting with these characters was quite rare, with all means at .26 or below. This is in contrast to findings from Study 3, in which the social condition manipulation involved a social goal explicitly written on the screen.

Reported Social Goals	ANOVA	M task	M social
Farid	$F(1,59)=.02, p=.884$	.24(.44)	.26(.44)
Hassan	$F(1,59)=.01, p=.921$	.14(.35)	.13(.34)

Table 9.10. Between-condition comparison for Reported Social Goals for each character.

Learners also indicated social goals in a more organic manner. As described in the procedure in Section 9.5, participants received a note over Instant Messenger from Zahora while they were preparing to meet the other agents in the game. Unexpectedly, half of the participants in the social condition, and one third of the participants in the task condition, wrote back to the agent. Qualitatively, the responses received from participants in the two conditions were very different. Most participants in the social condition replied with effusive, verbose comments including expressions of gratitude and future social goals:

P18: "Thank you for letting me know. I appreciate your help."

P29: "I will be happy to get to know him and hopefully gain his trust."

P40: "Thanks Zahora, perhaps we should all get tea so that I can meet Farid and understand his problems with being understaffed and underfunded."

P87: "Excellent i look forward to him being on the team. We will work together well."

P22: "Thanks for the encouragement. I will try my best."

One participant (P71) who had a family member in Iraq replied with "Shukran" ("Thank you" in Arabic), perhaps indicating a desire to show a shared perspective with the agent.

On the other hand, most replies from the task condition were the simple "Thanks" (e.g., P15, P23, P62) or curt "why?" (P7) or "That is understandable" (P72). The other common response was a question requesting more information, e.g.

P6: "Can we give them more resources?"

P34: "Could Farid be appreciative if the Iraqi police recieved [sic] resources?"

P47: "Would the new market place be something that he would be concerned about?"

Only one participant included a social closing: "thank you for the information, it is nice to know that." (P41).

#### 9.6.4. H3: LEARNING RESULTS

Finally, I investigated H3 to determine whether SID increased student learning, using learning questionnaires, change in student attitudes, and behavioral performance on a transfer negotiation scenario.

##### 9.6.4.1. CULTURE ASSIMILATOR

The Culture Assimilator test contained fourteen total items and was counterbalanced over two forms with seven items each. These forms were

randomly assigned within each condition. Items were divided between forms based on the learning objective they contained.

I first ran a repeated-measures ANOVA with condition as a between-subjects variable and pretest form as a covariate. While time  $\times$  condition was not significant ( $F(2,86)=2.24$ ,  $p=.113$ ), time  $\times$  pretest form was ( $F(1,86)=6.33$ ,  $p=.014$ ), indicating that analysis could not continue until the disparity between forms was corrected. In order to understand which items may have been contributing to the inequality between forms, I conducted an Item Response Theory (IRT) analysis (Lord, 1980). IRT models the response of an participant of given ability to each item in the test, and can help to uncover items that do not fit the underlying construct of the test.

Using posttest scores on the two forms (A and B), I looked at three common IRT models that allow for different numbers of fixed parameters: unconstrained Rasch, 1-parameter, and 2-parameter models. The unconstrained Rasch model had the best fit based on the BIC score. However, it is easiest to see which items are the most problematic in the visualization of the 2-parameter model (shown at the bottom of Figure 9.3), which does not constrain the discrimination parameter. Form A contained one item that was on average 80% correct (seen in red on the bottom left of Figure 9.3) regardless of the ability of the participant, and another item that was on average 40% accurate (in dark blue) and varied little with the ability of the participant. Form B presented one item (in dark blue on the bottom right of Figure 9.3) with an accuracy pattern that is the reverse of the desirable model; participants with an overall low ability were more likely to answer correctly than those with a high ability, possibly indicating that this was an easy question that good students “over-thought”. These three items were seen as good candidates for removal.

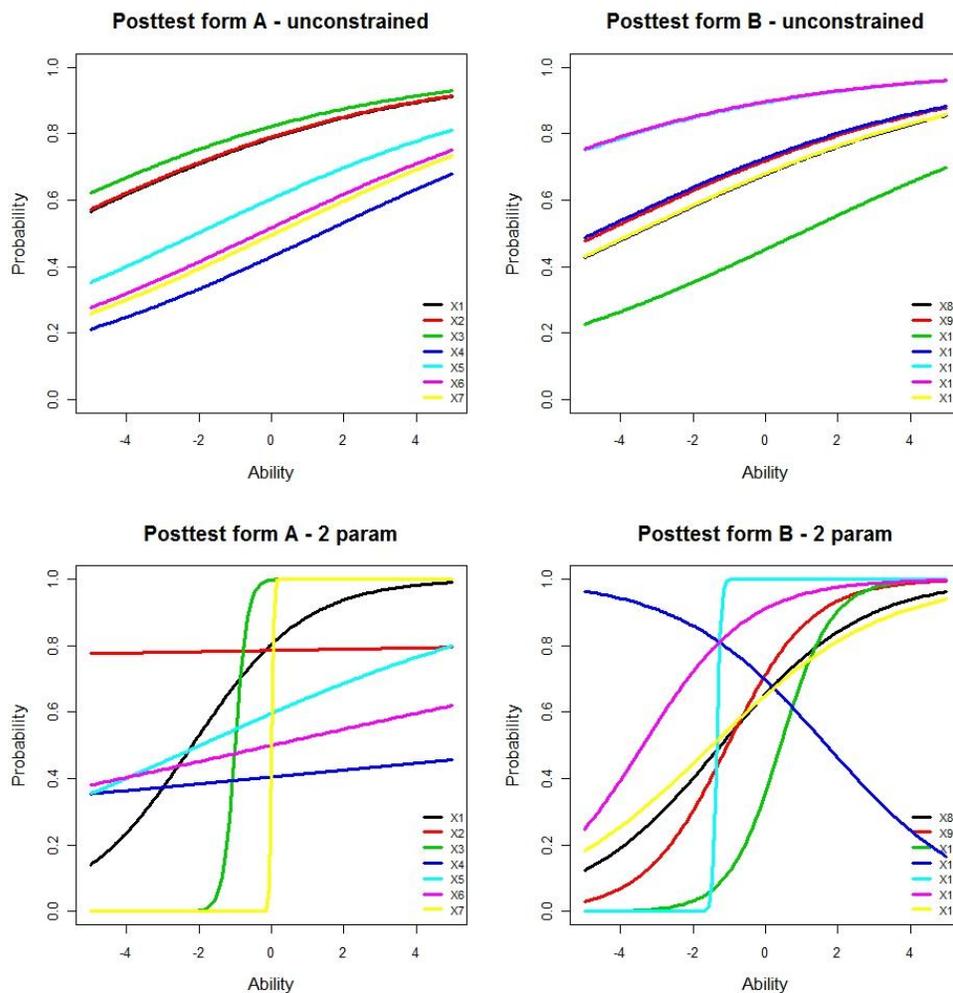


Figure 9.3. IRT models showing the probability of getting each item correct, based on the ability level of the learner.

Using the new scores, with the three items removed, I re-ran the repeated measures ANOVA. Time x condition was still not significant ( $F(2,86)=.632$ ,  $p=.534$ ); however, the difference in pretest forms was reduced, but not gone ( $F(1,86)=3.63$ ,  $p=.06$ ). Therefore, I treated the data as two separate experiments, running separate repeated measures tests based on pretest form. Time x condition was not significant for form A ( $F(2,45)=1.64$ ,  $p=.205$ ), or for form B ( $F(2,39)=.293$ ,  $p=.748$ ), indicating that in neither form were learning gains significantly different between conditions.

#### 9.6.4.2. SITUATIONAL JUDGMENT TEST

The second assessment, the Situational Judgment Test, was scored as in Study 3: participants rated each item on the test as more or less appropriate, and these scores were correlated with experts' scores on the same items.

A between-subjects ANOVA showed no pretest differences between conditions ( $F(2,87)=.13, p=.88$ ). A repeated measures ANOVA with condition as the independent variable also showed that there was no difference between conditions in learning gains across time ( $F(2,87)=.01, p=.99$ ). Similarly to Study 3, both pretest and posttest means were high, about .7 on average, indicating a 70% correlation with experts even prior to training.

#### 9.6.4.3. CHANGE IN ATTITUDES

The items that I used as measures of participant attitudes towards the military and Arab culture showed very little correlation from pre- to post-experiment. A series of repeated measures ANOVAs with condition as a between-subject variable and test time as a within-subject variable showed that condition was not a significant factor in the change in any of these attitudes (all  $p$  values greater than .3). However, given that participants' attitudes did change over the course of the experience, I investigated whether their perceptions of the agents may have influenced their attitudes.

A repeated-measures ANOVA with time as a within-subject variable investigating the item, "Most members of civilian society have a great deal of respect for the military," showed a significant effect of trusting Zahora ( $F(1,55)=7.66, p=.008$ ), and feeling that Zahora was very social ( $F(1,55)=5.93, p=.018$ ) in change of attitudes from pre- to post-experiment. For instance, as seen in Figure 9.4, more trust in Zahora was associated with an increase in feeling like civilians respect the military, while less trust in Zahora showed a decrease in this feeling. Given the theory behind projective attitudes surveys as

described above (Donoghue, 2000; Kline, 1983), this indicates that learners themselves became more positive about the military when they trusted Zahora. No effect was found for trusting or identifying with the other agents, possibly because these characters (unlike Zahora) were not directly associated with the military.

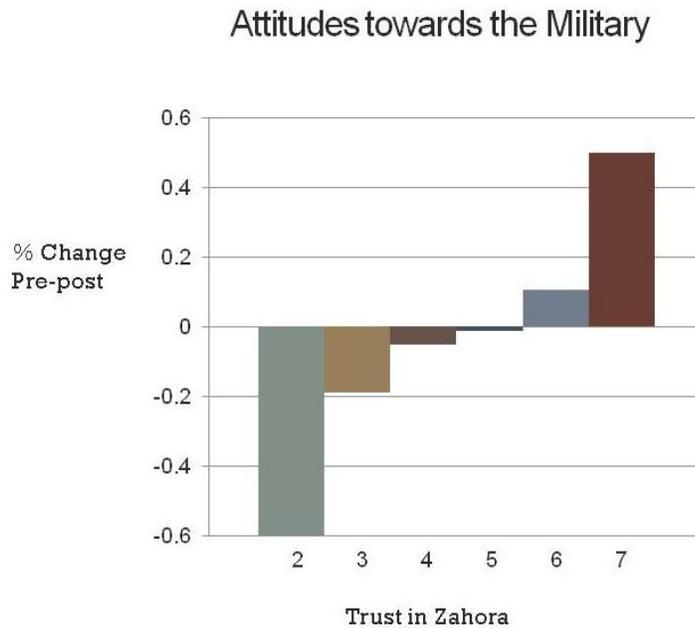


Figure 9.4. Change in positive attitudes towards the military from pre- to post-experience, based on learners' trust in Zahora.

Addressing change in attitudes towards the Arab culture, a repeated-measures ANOVA with time as a within-subject variable on the item “Arabic people living here teach their children values and skills different from those required to be successful in America” showed a significant effect of trusting Zahora ( $F(1,55)=6.03$ ,  $p=.017$ ), identifying with Zahora ( $F(1,55)=6.92$ ,  $p=.011$ ), trusting Farid ( $F(1,55)=4.19$ ,  $p=.045$ ), identifying with Farid ( $F(1,55)=10.53$ ,  $p=.002$ ), trusting Hassan ( $F(1,55)=7.10$ ,  $p=.01$ ), and identifying with Hassan ( $F(1,55)=11.22$ ,  $p=.001$ ). For instance, as seen in Figure 9.5, more trust in Hassan was associated with a decrease in feeling like Arabic values are not

successful in the United States, while less trust in Hassan showed an increase in this feeling. In other words, learners became more positive about the Arab culture when they trusted and identified with the agents, all of whom were associated with the Arab culture.

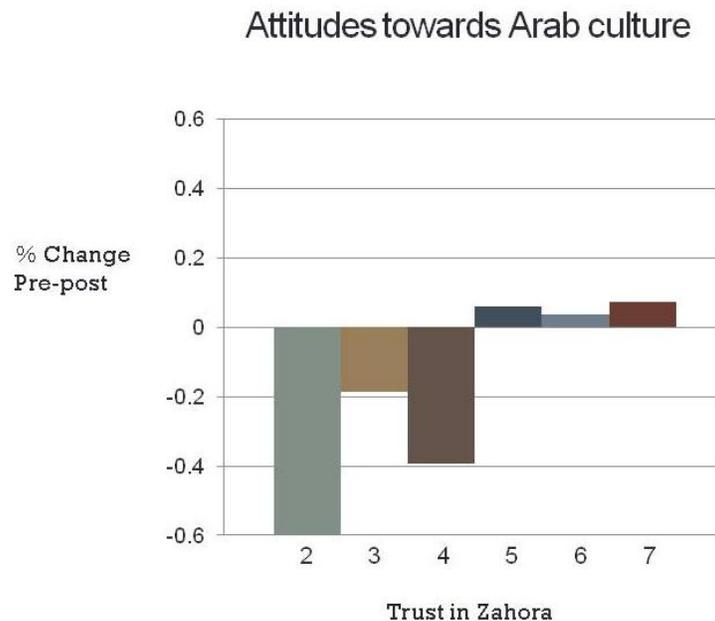


Figure 9.5. Change in positive attitudes towards the Arab culture from pre- to post-experience, based on learners' trust in Hassan.

#### 9.6.4.4. BEHAVIORAL MEASURE: TRANSFER TO A NEW SCENARIO

The third type of posttest I administered was a practice-based, in-game assessment. Participants negotiated with a third character, Aziz, who was part of a different scenario in BiLAT. In this assessment, I measured both the negotiation objectives achieved (in this negotiation, there was only one possible objective) and the number of intercultural errors committed. A between-subjects ANOVA comparing conditions, shown in Table 9.11, reveals a significant difference on both the number of errors made in the negotiation, as well as the objectives met. According to Tukey's post-hoc pairwise comparison, only the social condition committed significantly fewer errors than

the control group, while the task group was statistically equivalent to both the social group and the control group. The control group met significantly fewer objectives than both the task and social conditions, which were statistically equivalent.

Item	ANOVA	M control	M task	M social
Aziz: Errors	$F(2,87)=4.28, p=.017$	10.77(8.22) <sup>a</sup>	7.37 (5.48) <sup>a,b</sup>	5.87 (5.91) <sup>b</sup>
Aziz: Phase errors	$F(2,87)=3.95, p=.023$	7.73(6.84) <sup>a</sup>	5.27 (4.16) <sup>a,b</sup>	4.13 (3.61) <sup>b</sup>
Aziz: Objectives met	$F(2,87)=4.96, p=.009$	.53(.51) <sup>a</sup>	.83 (.38) <sup>b</sup>	.83 (.38) <sup>b</sup>

Table 9.11. Between-condition comparisons for the in-game posttest. Superscript letters indicate statistically equivalent groups.

#### 9.6.5. INFLUENCE OF LEARNER CHARACTERISTICS ON INTERPERSONAL AND LEARNING OUTCOMES

Next, I examined the relationships between learner characteristics and social and learning outcomes. Surprisingly, there appeared to be little connection between many of the measured learner characteristics and outcomes. However, there were a few groups of learner characteristics that repeatedly appeared to relate to outcomes: extroversion/introversion, goal orientation, cross-cultural experience, and gaming experience.

##### 9.6.5.1. EXTROVERSION

Extroversion was related to several aspects of gameplay. The participants' level of extrovertedness correlated with finding the activity engaging, controlling for condition ( $r(84)=.284, p=.008$ ). Extroverts also felt a greater sense of presence in the game ( $r(59)=.292, p=.026$ ).

Extroversion was also related to perceptions of game characters. Controlling for condition, extroversion was correlated to both thinking that "Zahora

shared personal stories” ( $r(59)=.341, p=.008$ ) and “Zahora was very social” ( $r(59)=.286, p=.028$ ). Belongingness goals for interacting with Zahora were also associated with extroversion ( $r(59)=.334, p=.009$ )

Finally, success in the game was negatively related to introversion. In fact, controlling for condition, the only factor that correlated to meeting objectives with Aziz was introversion ( $r(59)=-.286, p=.033$ ). The more introverted a participant was, the less likely they were to achieve meeting goals.

#### 9.6.5.2. MASTERY AND PERFORMANCE ORIENTATION

Academic goal orientation also related to several social outcomes (in particular, social goals). Mastery orientation seemed to be associated with aspects of the environment that were less task-focused and more social. Controlling for condition, belief that Zahora is very social was correlated to having a mastery approach orientation ( $r(59)=.307, p=.019$ ), as was feeling entitativity with her ( $r(59)=-.322, p=.014$ ), and having belongingness goals for interacting with her ( $r(59)=.305, p=.019$ ).

On the other hand, having a greater performance orientation had greater association with aspects of the task. Feeling entitativity with Farid, the first negotiation partner, was correlated with having both a performance approach ( $r(59)=.350, p=.007$ ) and performance avoid ( $r(59)=.308, p=.018$ ) orientation. Trust in Farid was also correlated to a performance avoid orientation ( $r(59)=.284, p=.031$ ). Power goals were correlated with both having a performance approach ( $r(59)=.394, p=.002$ ) and performance avoid ( $r(59)=.272, p=.037$ ) orientation, and assistance goals were correlated to having a performance avoid ( $r(59)=.308, p=.018$ ) orientation.

## 9.6.5.3. CULTURAL EXPERIENCE

A learner's cross-cultural experience was related to several social and game outcomes, including perceptions of the game characters. People who felt more similar to Arabs felt like Zahora was less task-focused and that she made more small talk; i.e., believing that Arabs have similar values to you was negatively correlated to both "Zahora focused on the task" ( $r(59)=-.398, p=.002$ ) and "Zahora did not make much small talk" ( $r(59)=-.390, p=.002$ ). People with greater cultural knowledge also felt like Zahora made more small talk; having greater knowledge of Arab cultures was negatively correlated to "Zahora did not make much small talk" ( $r(59)=-.349, p=.007$ ). Presence in the environment was associated with Cultural Intelligence ( $r(59)=.315, p=.016$ ), as was feeling entitativity with Farid ( $r(59)=.301, p=.022$ ).

Gameplay and success was also associated with cross-cultural experience. Making errors was negatively correlated to Cultural Intelligence ( $r(59)=-.288, p=.032$ ), and believing Arabs have similar values ( $r(59)=-.343, p=.01$ ), controlling for condition. In addition, although there were no evident learning gains on this measure, I found that post scores on the SJT were correlated with Cultural Intelligence ( $r(87)=.273, p=.04$ ).

## 9.6.5.4. GAMING EXPERIENCE

Learners' experience with games was also correlated with certain aspects of gameplay. Making errors in the game was positively correlated to frequency of playing games ( $r(59)=.320, p=.016$ ), playing first person shooters ( $r(59)=.279, p=.037$ ), and considering oneself to be a "gamer" ( $r(59)=.376, p=.004$ ). Posttest scores on the SJT was correlated with playing first person shooters ( $r(87)=.302, p=.02$ ).

## 9.7. DISCUSSION AND CONTRIBUTIONS

In this study, I found that the type of dialog used by agents can significantly affect learners' attitudes towards the agent. Social informational dialog (SID), composed with the use of the conversational strategies of narrative, affirmation, and self-disclosure, was rated as more social in the manipulation check, while the task-based comparison dialog (TID) felt focused on the task. Confirming *H1*, I found that SID influences learners to have more social perceptions of virtual agents who employ this dialog. In particular, SID participants liked and trusted the agent more, they felt more like they were working on a team with her, they felt like she knew them and they knew her better. She influenced them to conform to her perspective in the responses they gave. Confirming *H2*, I found that agents who use SID increase learners' social goals for interaction. SID participants desired acceptance and recognition from Zahora; they cared whether she liked them. These social outcomes provide a contribution to the literature on learner-agent interactions, increasing understanding of how varying agent characteristics influence learner perceptions. These results also make design contributions, serving as a guide for instructional designers in creating social pedagogical agents.

On the other hand, the task dialog seemed to have a greater feeling of authoritative instruction. Learners in the TID condition reported that they were demonstrating an American perspective back to the agent, perhaps trying to teach her something in return (an alternative explanation is that they were being contrary in responding to her proclamations of Iraqi perspectives!). Immediately after meeting with Zahora, learners in the task condition also reported significantly greater goals for desiring her help. This difference in social goals between agents is an interesting contribution that may suggest possible uses of dialog types, depending on the goals instructional designers wish to support. It is also, to my knowledge, the first use of such a social goals questionnaire in learner-agent interactions. This dissertation makes a

contribution through its exploration of the responses to this questionnaire, by uncovering dimensions of the scales (in particular acceptance and assistance goals) that are relevant to virtual learning environments. This survey approach has utility in understanding agent effects on learner intentions.

However, in post-study interviews, many participants in the TID condition stated that they did not actually find Zahora to be more helpful than the documents they read in preparation for each meeting: “She was kind of helpful. But not so much as the packets of information we got.” They said that if these documents had not been available, they would have found Zahora to be a very useful resource. Instead, “She wasn't very helpful. It wasn't that I didn't like her so much. She chimed in with information at a time when I didn't need it.” On the other hand, students in the social condition did not make any comparisons between Zahora and the documents they read. Instead, most of them claimed that she was very useful in social ways; e.g., making them feel more comfortable in the rest of their interactions (“It was great to start off slow, she let you feel relaxed unlike if you went straight into the negotiation”), and letting them practice the social part of the meetings (“It was a nice introduction to the game, and something that wasn't so intense... yeah, there were no objectives, just try to get to know this woman”).

These qualitative observations may also help explain the contradiction with Study 4's findings on trust. In the current study, learners in the SID condition expressed more trust in the agent. While this finding was in line with my hypothesis (*H1*), in Study 4 I had found that TID learners trusted Zahora more. As is hypothesized in (Mulken, Andre, & Muller, 1999), I believe this may indicate that ratings of trust are positively correlated to utility of information. In Study 4, the agent was evaluated outside of the full educational context, and the TID dialog's more authoritative and less personal tone led students to trust her in the absence of confirmatory evidence. In the current study, the subsequent utility of the communicative practice the SID agent gave increased

learners' ratings of trust. These two findings constitute a contribution to the literature, suggesting that social relationships with agents increase trust, but only if the agent is perceived as useful. Further exploration of the concept of trust in learner-agent interactions is warranted.

Zahora also influenced students' perceptions of other agents in the environment, even when their dialog had not been modified. Learners with a SID-based Zahora felt significantly greater entitativity and trust with subsequent meeting partner Hassan. The SEM model gives further evidence that the strength of learners' perceptions of Zahora was proportional to their attitudes towards the other agents, regardless of condition. Specifically, participants with higher ratings of entitativity with Zahora reported similar higher ratings with both subsequent characters. The SEM model also shows that feeling greater entitativity with each character leads to greater trust in that character. Confirming that the interaction with Zahora influenced the rest of their experience, participants remarked that "Being able to interact socially with Zahora in the beginning, it really made you feel more trusting of the [other] characters." These results also represents a substantial contribution to the literature, suggesting that once a social relationship with one agent has been established, learners' social perceptions carry over to other interactions in the learning environment.

Although the SEM shows that trust and entitativity for both of the other agents were increased by learners' social perceptions of Zahora, these effects were most pronounced with Hassan, and effectively hidden in between-subject ANOVAs with Farid. Beside condition, feeling entitativity or trust with Hassan was not associated with any learner characteristics or attitudes. I instead attribute this outcome to the significant difference in the "likeability" of these two characters. Hassan was intentionally designed to be more difficult, less cooperative, and overall less team-driven, an impression confirmed by post-study survey questions and interviews (e.g., "Hassan is really a jerk!"). Farid, on

the other hand, was received quite positively, aligning strongly with his intended design. Therefore, for Hassan, who had the lowest “likability” of the three characters, the magnitude of the manipulation effect was more visible. This has interesting implications for agent design. In a typical dialog-based learning environment for STEM skills, all pedagogical agents might be designed to use SID. However, to transfer to real-world encounters in the domain of interpersonal skills, students must learn to deal with “difficult” people. Our results suggest a SID manipulation early on can raise entitativity ratings of a difficult character from negative to neutral, and may induce sufficient motivation in the learner to push through the engagement.

An open question is how far this influence extends. Did Zahora influence learners’ feelings only about other agents working on their team? Or perhaps all agents in this context? What about transfer to humans outside of the environment? One way in which my dissertation attempts to answer this question is to look at participants’ change in attitudes. In particular, participants who trusted and identified with the agents in this study decreased their feelings that Arabic people have values and skills different from those required to be successful in America. In other words, learners became more positive about the Arab culture when they trusted and identified with the agents, all of whom were associated with the Arab culture. This result indicates that participants experienced an attitude change which may extend beyond the environment, possibly leading to more positive interactions in the world.

The results also suggested that the agent can affect attitudes about the military. Trust in Zahora was associated with an increase in the belief that civilians have great respect for the military. Such “projective” attitudes scales, while asking about society at large, are used to obtain a truer measure of participants’ own attitudes, as described in Section 9.4.2. These contributions show that use of an agent whom learners trust can effect a change in attitude towards the context in which the agent exists.

*H3* proposed that learning intercultural skills from a social simulation would be more successful than reading a similar document. This was confirmed in the in-game, behavioral posttest. Learners in both BiLAT conditions were more successful in achieving the negotiation outcome than the control condition. Only the social condition, however, made fewer errors than the control while negotiating with the agent. As was suggested in Study 3, this is further indication of a more careful, social approach to playing the game. On the other hand, learners who self-identified as gamers, who played more games and in particular first person shooters, were much more likely to commit errors when meeting with Aziz.

I also found that learning from a social simulation was significantly more fun and more engaging than reading the same material from a document. In post-experiment interviews, many of the learners in the control group said they were bored (“I was kind of losing focus, especially towards the end”), and in observations during the experiment, many were easily distracted. On the other hand, several of the participants in the BiLAT conditions asked if they could take the game home or find it on the web, as they wanted to continue playing (“They should make this into a product that people can buy!”) and suggested that the experience did not end in the lab (“I thought it was really interesting. It made me think about some issues, that I'm going to go home and think about more”).

In this study I did not find any significant differences in learning between groups on the two “paper” measures, the Culture Assimilator and the Situational Judgment Test. This is not to say, however, that no learning differences exist. The particular implementation of these measures suffered from several issues, including a difference in difficulty between forms, and high pre-test scores. The search continues for accurate measures that cover the wide variety of knowledge that is being taught by these systems.

Finally, in this chapter I explored learner characteristics that may affect interactions with an intercultural virtual agent. As described above, many of these characteristics were suggested by Abbe, Gulick, & Herman (2008) as being antecedents of effectiveness in a cross-cultural setting. Others were biographical information that I believed to be relevant to a virtual environment. Surprisingly, many characteristics did not seem to affect learner interactions. Contrary to findings by Abbe et al., demographics information such as foreign experience, number of languages spoken, and self-reported negotiating skill, and gender were rarely if ever correlated with outcomes. On the other hand, characteristics such as extroversion, mastery/performance orientation, and the newly introduced Cultural Intelligence scale showed up frequently in correlations. The difference may be partially explained that this study was conducted in a learning environment, while their work seeks to explain effectiveness in cross-cultural settings. Characteristics that predict aptitude for learning cross-cultural skills in an instructional environment may not be identical to those that predict skill application in the field. While the work done in this study on learner characteristics is exploratory rather than causal, it contributes a direction for productive (and unproductive) characteristics to utilize as explanatory factors in future research on intercultural training in virtual learning environments.

# 10. DISCUSSION, CONTRIBUTIONS, AND FUTURE DIRECTIONS

My thesis lies at the intersection of the fields of human-computer interaction, learning sciences, and intercultural training. Virtual training environments that are populated with enculturated agents are a promising new way to teach intercultural competence skills. My work seeks to understand what benefits there are to learners in interacting with these virtual agents in social ways, and if there are benefits, how do we encourage learners to do so. In Study 1, I first explored expert strategies and misconceptions in the domain of intercultural negotiation, while validating the content of the virtual learning environment with experts in both the military and Arab culture. In Study 2, I investigated novice misconceptions and errors made during practice while utilizing a think-aloud procedure. Both novices and experts had difficulty separating the social interaction with the virtual agents that populate the environment from the game mechanics that propel the experience. Therefore, in Study 3, I manipulated the environment to prompt learners with explicit social goals for the interaction. I found that this explicit prompting was not successful in increasing learning from the environment; however, I did find that learners who reported their own social goals for interacting with the agents did learn more from the interaction. Given that explicit prompting was not as successful as desired, I focused on ways in which learners' social orientation might be influenced in a more implicit fashion. As described in Chapter 7, I developed a model of social informational dialog that tied conversational strategies to interpersonal outcomes that influence this social orientation, as well as learning outcomes. In Study 4 I piloted a new agent developed using this dialog model, in comparison to an agent using a task-based model. In the final study,

described in Chapter 9, I found that the same agent significantly influenced learners' perceptions of the agent, lead to a reduction of social errors in future negotiations, and found a change in attitudes towards both Arab culture and the military. In the remainder of this chapter, I discuss my findings in more detail, outline the contributions this work makes to each of the three domains I touch on, and suggest possible future directions for this work.

## 10.1. LEARNING SCIENCES

The work in this dissertation contributes to the learning sciences by developing an understanding of the role of social goals in learning interpersonal skills. I do this in three ways: studying the explicit prompting of social goals, investigating learning interpersonal skills from a virtual environment, and exploring learners' interactions in the environment.

### 10.1.1. EXPLICIT PROMPTING OF SOCIAL GOALS

Studies that investigate causal effects of learners' achievement goals (e.g., performance or mastery goals; see Dweck & Leggett, 1988, McNeil & Alibali, 2000) generally use short, explicit instructions prior to a learning task to influence students' objectives. I found, however, that with social goals, explicit prompting of learners' goals through brief instruction was not beneficial to learning. In fact, a third of the students in the explicit social goals condition in Study 3 did not heed the manipulation, based on the goals they reported after the study. These students may not have understood how to achieve the social goal they were given, or may not have wanted to achieve it.

What the data showed instead was that learners with *self-reported social goals* for the interaction had greater learning gains than learners without such goals. The results of this study reconcile two sets of seemingly conflicting evidence. Okita et al. (2008) and Rosé and Torrey (2005) suggest that students should

believe that virtual characters are being controlled by a human to benefit from interacting with them. However, Reeves and Nass (2006) suggest that social responses to virtual characters are automatic, even when there is no indication that there is a human in the loop. One explanation for the discrepancies between these results is that learning is not an automatic process like emotion but requires attention and processing (Cohen, Ivry, & Keele, 1990). Students believe they are taking a socially relevant action when they interact with a human, and thus pay more attention and feel more accountable. In Study 3, I found that only learners with self-reported social goals had higher learning gains. It may be that in tasks that require deep processing, having social goals for interacting with virtual agents (giving them social agency) can help learners achieve the same benefits as learning from a human.

In addition to an explicit goal manipulation, this work makes a contribution to methodology for the study of social goals. In Study 5, I shifted from self-report of goals to investigate the first use of a validated social goals questionnaire (McCollum, 2006) in understanding learner-agent interactions. Based on the data collected in this study I refactored the social goals scale into dimensions that may be the most relevant to interactions in virtual learning environments, namely: acceptance, power, belongingness, and assistance (see Appendix C). Use of this social goals questionnaire could benefit understanding of interactions in both learner-agent as well as peer-peer collaborative settings.

#### 10.1.2. LEARNING INTERPERSONAL SKILLS FROM A VIRTUAL ENVIRONMENT

It is important to determine whether virtual learning environments for training interpersonal skills, which require a huge investment of resources to develop, are beneficial to learning. While a few studies have been conducted in such environments and shown that students do learn (e.g., Johnson & Wu, 2008), to my knowledge, none have used a validated assessment and compared to an ecological control condition to understand the unique advantages of VLEs.

In Study 3, I found on a validated assessment that learners do indeed learn from practice in a virtual environment for interpersonal skills training. As described above, I also found that those learners who reported having social goals had significantly higher learning gains.

I then compared use of a simulation to an ecological control group that read a document on the same material. In Study 5, I found that training in a social simulation has a significant learning benefit over an ecological control in a practice-based posttest. All learners using the simulation were more successful in achieving the negotiation outcomes than the control condition. Only the social condition, however, made fewer errors than the control condition while negotiating with a new agent. This may be indicative that learners with a social orientation take a more careful, social approach to interacting with virtual agents, which is further discussed in the section below. Learners also found the virtual training environment significantly more fun and more engaging than reading. Increased engagement is an important benefit of using a virtual environment for learning. The more likely learners are to utilize a training tool outside of traditional learning time, the more likely they are to see learning gains.

In Study 5, I did not find any significant differences in learning between groups on the two “paper” measures. However, the measures as used in this study suffered from several issues, including a difference in difficulty between forms, and high pre-test scores. Because the in-game practice measure is a behavioral measure, it is arguably more important and more ecologically valid than the two multiple choice tests mentioned here.

### 10.1.3. INTERACTIONS IN THE LEARNING ENVIRONMENT

Beyond posttest assessments, it is important to develop a formative understanding of the productive ways in which learners interact with virtual

agents. I found that learners who report a social goal interact with virtual agents in a qualitatively different way from those who do not have such a goal. In an identical amount of time in the game, they took fewer actions, which may indicate that they spent more time reflecting on each action to consider their partners' perspective. They took fewer total actions relating to business and a higher percentage of social actions. Additionally, they took fewer unique actions, signifying less exploration of the conversation space (seemingly avoiding dialog actions that could potentially be seen as offensive).

Together, these patterns indicate that learners with social goals take what I believe is a social orientation towards gameplay. This outlook is in contrast to a prominent view of learning from gameplay, which involves exploring a risk-free, task-oriented environment for discovery, as suggested by the PsychoSocial Moratorium principle and other theories of game-based learning (Gee, 2003; Johnson, 2006). This principle states that games are a place where learners take risks they would not normally be comfortable taking in the real world. In BiLAT, learners might manifest this principle by intentionally offending the virtual character, or experimenting with all available actions in an attempt to understand the boundaries of acceptable behavior. In fact, learners do engage in these behaviors, as can be seen in the verbal protocol data presented in Study 2. Given the context of the virtual environment, learners who play more games may be at greater risk of indulging in mindless behaviors. In Study 5, I found that learners who self-identified as gamers, who played more games and in particular first person shooters, were much more likely to commit errors while interacting with an agent in a behavioral posttest.

This risk-free gameplay is at odds with a social orientation, in which learners would carefully consider their partner's perspective, attempt to avoid giving offense, and avoid exploration that would take them into unknown territory of culturally acceptable behaviors. This implies that learners with social goals are engaging in mindful social interaction, rather than mindless trial and error

behaviors. In mindful play, learners likely feel more accountability for their actions, and thus reflect more on the consequences and benefits of each action, and make a greater effort to understand the feedback they receive from the agent. This dissertation proposes a characterization of social gameplay, and indicates that encouraging mindful social gameplay through implicit means is beneficial to learning.

## 10.2. INTERCULTURAL COMPETENCE TRAINING

Beyond contributions to the learning sciences in general, the work in this dissertation contributes to understanding in the particular domain of interest, intercultural competence. I do this in three ways: enumerating successful strategies and misconceptions held by experts and novices, exploring what learner characteristics are associated with intercultural competence, and validating the content and investigating learning gains from an intercultural simulation.

### 10.2.1. SUCCESSFUL STRATEGIES AND MISCONCEPTIONS

Understanding the strategies that experts would apply in the real world, and documenting misconceptions and errors that novices and experts make in a VLE can help us find targets for improvement in intercultural training, and in the interfaces of the virtual training environments themselves. Work exploring these questions in the context of a VLE has not previously been published. Through interviews with experts, I found that a suite of interpersonal skills were important for successful negotiations, which highlights skills such systems can focus on developing. In general, experts indicated that the social interaction, especially small talk and appropriately opening a meeting, is critical to success. It is also important to have a good knowledge of your meeting partner prior to meeting. At the same time, experts strongly emphasized the

need for flexible or adaptable thinking, as new knowledge is gained or external circumstances change from moment to moment. Intercultural simulations, including BiLAT, currently provide an excellent environment for practicing small talk and specific cultural skills (e.g., when and what gifts to offer). Focusing on adaptable cultural thinking, for instance by training students on how to make appropriate cultural explanations for observed behavior, is a promising next step.

It was also important to identify any misconceptions or potential barriers to learning that still remained at an expert level. One area for improvement I identified was that even experts may approach the environment with a “gamer mentality” that prevents them from engaging with the agents. Both experts and novices attempted to explore actions that would deliberately provoke the agents, and often blamed the interface for their failures in communication. They often became overly focused on game mechanics rather than thoughtfully considering the social implications of their actions. Also, it appeared that negotiation skill and positive cultural attitudes do not always coincide. Supporting positive and open attitudes while addressing these unproductive behaviors are excellent targets for intercultural training simulations.

#### 10.2.2. CONTRIBUTING CHARACTERISTICS TO INTERCULTURAL COMPETENCE

Also of interest to the intercultural training community is to understand what characteristics contribute to success in intercultural interactions. For organizations from the Army to the Peace Corps, it is important to have measures that can predict an applicant’s potential in this field. Abbe, Gulick, & Herman (2008) have investigated characteristics that explain effectiveness in cross-cultural settings. I investigated these same characteristics along with other demographic information in the context of a learning environment. Contrary to findings by Abbe et al., demographics information such as foreign experience, number of languages spoken, and self-reported negotiating skill, and

gender were rarely if ever correlated with learning, social, or attitudinal outcomes. On the other hand, characteristics such as extroversion, mastery/performance orientation, and the previously unused Cultural Intelligence scale showed up frequently in correlations. While the work done on learner characteristics in my studies is exploratory rather than causal, it contributes a new focus for productive (and unproductive) characteristics as explanatory factors in future research. It also suggests data that may contribute to an improved learner model for use in intelligent tutoring systems in this domain. For example, social agents may be developed to take a different strategy with extroverts than with introverts.

### 10.2.3. SIMULATION VALIDATION AND LEARNING GAINS

According to the Handbook of Intercultural Training (Landis, Bennett, & Bennett, 2003), there is still a great need for empirical work that determines whether intercultural training can be effective. Before investigating whether a simulation can produce learning results for students, it is important to determine that the content they are learning is a valid representation. BiLAT is a state-of-the-art environment for intercultural negotiation training which has been used by many learners, but the cultural content had not been validated subsequent to development. In order to do this, I conducted interviews with experts from the Army and Iraqi nationals. The experts indeed found the experience to be convincing and engaging, the cultural behaviors and appearance of the agents to be authentic, and the phases and social interactions in the negotiation accurate. Given the current lack of acceptance of technology in the field of intercultural education, it is important to show that such tools can in fact provide learners with a realistic view of another culture.

The next step was to understand the effects of the training tools on learning. BiLAT is one of the few environments to date that has been quantitatively assessed in any manner (see Durlach, Wansbury, & Wilkinson, 2008; Lane et

al., 2008; Hays et al., 2009). Prior studies on virtual training for intercultural skills, however, generally either do not compare to a control condition that is external to the simulation, or confound learning from the practice environment and a preparatory video. Showing that a system for training intercultural competence can produce increases on a validated assessment is a contribution given the current state of evaluation in the field (Ogan & Lane, 2010), which relies mainly on qualitative metrics or self-report of learning.

In Study 3, I looked only at learning within a social simulation. I used two assessments that had been validated by other researchers: the Situational Judgment Test (see Appendix B), which had previously been used as a measure of learning, and the Culture Assimilator (see Appendix A), which had not previously been used as an assessment with a virtual learning environment. The Situational Judgment Test describes an intercultural scenario and asks learners to judge whether particular actions are appropriate or inappropriate to take in that situation. Although previous studies showed that BiLAT increased gains on this assessment, my data indicate that these gains most likely came from viewing the preparatory video rather than practice in the simulation (see Hays, Ogan, & Lane, 2010). This suggests that the SJT is more suitable for assessing declarative intercultural knowledge that can be learned from a video or presentation. The Culture Assimilator also describes an intercultural scenario, but then asks learners to make an appropriate cultural explanation for the events in that scenario. Based on the results from this study, I found that the Culture Assimilator, on the other hand, shows potential as a quantitative measure of deeper knowledge such as that learned in a virtual practice environment.

In Study 5, I evaluated whether intercultural negotiation is better learned from practice with enculturated virtual agents than an ecological, non-simulation control, namely, reading a high-quality text on the subject of intercultural negotiations. As discussed above, I did find a significant win over the control in

the practice-based posttest. Learners who had used the virtual environment were more successful in achieving the negotiation outcome and made fewer errors than the document-reading control. A main goal of intercultural training is to prepare learners for interacting in the real world. Achieving more objectives and making fewer social errors are both critical outcomes that heavily influence intercultural interactions. My results suggest that using a virtual environment for training can help learners achieve these outcomes.

Finally, I also wanted to discover whether *savoir-être*, or holding positive attitudes and openness to new cultures, is an important disposition that contributes to the skills of performing appropriate cultural behaviors (*savoir-faire*). While I did not find a correlation between the two in Study 5, I did find that attitudes did change based on students' perceptions of the agents in the environment. I found that trusting and identifying with the Iraqi agents was associated with a statistically significant decrease in negative feelings about Arabic people. This result is an indication that learners training in a virtual environment can experience an attitude change which may extend beyond the environment, possibly leading to more positive interactions in the world.

### 10.3. VIRTUAL AGENTS

The work in this dissertation also contributes to understanding human-agent interactions. I do this in three ways: formalizing a dialog development process and creating a model for social informational dialog, investigating the social effects that this model has on learner-agent interactions, and uncovering the attitude change associated with differing perceptions of the agents.

#### 10.3.1. DIALOG MODEL AND DEVELOPMENT PROCESS

This dissertation contributes a formalization of a process for the generation of agent dialog. In the *Preparation* stage, I developed a character bible and a set of

conversation topics for the agent. While the original designers of BiLAT (professional game designers) based their hand-written dialog on real-world scenarios described by experts, the *Dialog Generation* stage of my process instead used role-players to produce a large corpus of utterances. In the *Iterative Development* stage, these utterances were then crafted into agent and learner dialog, with which I explored user-based research methods with rapid prototyping tools to refine the discourse. Other recent and current work in this area has focused on machine-generation of dialog for narrative-based agents (Riedl & Young, 2003; Cassell & Bickmore, 2003). While this is also an important goal for the community, there are many reasons why researchers may want to have more control over their dialog; for instance, the incorporation of the unique perspectives that role-players of different backgrounds may provide. Formal procedures for writing such dialog with user input have not been described. The process presented in this thesis, which includes descriptions of the pitfalls and successes encountered, can be used by any researcher for the development of dialog for agents in narrative-based learning environments.

This development process went hand in hand with the creation of a model of social informational dialog (SID), linking conversational strategies to desirable interpersonal outcomes. Agent dialog models are of interest to both the intelligent virtual agent as well as the educational technology communities. The model described in this dissertation differs from other dialog models in its instructional nature, as well as the incorporation of conversational strategies directly into the dialog rather than as separate utterances added to a task-based model (e.g., Cassel & Bickmore, 2003; Gratch, Wang, Okhmatovskaia, Lamothe, Morales, & Morency, 2007; Wang, Johnson, Rizzo, Shaw, & Mayer, 2005). In this work, I identified three interpersonal outcomes that were likely to lead to stronger social ties between an agent and a learner: trust, entitativity, and shared perspective. Through a review of communications theory and the

data generated in the Iterative Development stage of the process, I then identified three conversational strategies that should lead to these interpersonal outcomes; narrative, affirmation, and self-disclosure. The extensive literature on human-human communication has yet to be fully explored for theories that might affect the development of relationships with virtual agents, and this work focuses on three that might affect learning outcomes as well.

### 10.3.2. SOCIAL EFFECTS IN LEARNER-AGENT INTERACTIONS

The goal of such a dialog model is to change user perceptions of the agent and their interactions. In Study 5, I compared an agent using the SID model of conversational strategies and interpersonal outcomes to a task-based (TID) model of dialog. I found that learners perceive a significant difference between task and social dialog; SID feels more social and less task-focused. SID causes greater entitativity, shared perspective, liking of the agent, and feelings of knowing one another. It also influences social goals in specific ways, by increasing goals for acceptance and belonging. Task dialog, on the other hand, seemed to have a greater feeling of authoritative instruction. It increases participants' desire for assistance from the agent. It also engenders a contrasting, home-culture perspective, perhaps an indication that learners are trying to teach the agent something in return.

More complex are the effects of SID on trust. In Study 4, the agent was evaluated outside of the full educational context, and the more authoritative and less personal tone of the TID agent led students to trust her significantly more in the absence of confirmatory evidence. In Study 5, however, learners in the TID condition expressed that the agent was not more helpful than the documents they read in preparation for each meeting. Learners in the SID condition, on the other hand, had significantly greater trust in the agent, and claimed that she was very useful in social ways. As is hypothesized by Mulken,

Andre, and Muller (1999), I believe this may indicate that ratings of trust in the agent are tied to utility of information. These two findings constitute a contribution to the literature, suggesting that social relationships with agents increase trust, but only if the agent is providing information perceived as useful. Further exploration of the concept of trust in learner-agent interactions is warranted.

The use of SID also influences attitudes towards other agents whose dialog is not based on SID. In Study 5, learners with a SID-based Zahora felt significantly greater entitativity and trust with subsequent meeting partner Hassan. The SEM model constructed in this study gives further evidence that the more SID influenced perceptions of a SID-based agent, the more it influenced perceptions of the other agents in the environment. SID engendered greater feelings of entitativity, which in turn, boosted perceptions of trust. The SEM model shows this effect carries forward to future interactions with other characters with both trust and entitativity. This result also represents a substantial contribution to the literature, suggesting that once a social relationship with one agent has been established, student social perceptions and attitudes carry over to other interactions in the learning environment.

As interesting as learners' ratings of trust, entitativity, and shared perspective, was their spontaneous behavior in responding to chat messages purportedly from the agent. To reinforce the manipulation, Zahora sent learners an encouraging note prior to interacting with each of the other characters. In an informal qualitative analysis, the responses received from participants in the two dialog conditions were very different. Most replies from the task condition were simple and curt, or a request for more information, confirming the hypothesis that task dialog is seen as more authoritative. On the other hand, the SID agent received replies with effusive, verbose comments including expressions of gratitude and future social goals. This result gives a richer picture of the social response motivated by SID-based dialog.

These social outcomes provide a contribution to the literature on learner-agent interactions, increasing understanding of how varying agent characteristics influences learner perceptions. These results also make design contributions, serving as a guide for instructional designers in creating social pedagogical agents. To create an agent that engenders entitativity, trust, and shared perspective, an effective way to do so is to ensure the agent's dialog contributions include a high proportion of narrative, self-disclosure, and affirmation. The difference learners report in their social goals for interaction also suggests potential uses of these dialog models, depending on the goals instructional designers wish to support. For instance, an agent who will act as a resource, such as a biologist in an ecology simulation, might use task dialog to encourage learners to come ask for help. On the other hand, an agent that acts as a fellow “student” in a peer pressure simulation for teens might use social informational dialog to increase students' desire to belong to their group, making the experience more difficult and realistic. Better understanding the role of goals, both social and academic, in designing learner-agent interactions is an important one for future research.

### 10.3.3. ABILITY OF AGENTS TO INFLUENCE LEARNER ATTITUDES

Finally, a somewhat unexpected outcome is that learner perceptions of an agent are associated with a significant change in attitude towards the context in which the agent exists. In particular, I found that trusting and identifying with the agents in Study 5 was associated with a statistically significant decrease in negative feelings about Arabic people, and an increase in positive attitudes towards the military. Even more importantly, *distrust* of the agents was associated with the reverse: an increase in negative opinions about Arabs and a decrease in positive opinions about the military. These effects were much larger than I anticipated. While it is encouraging that positive perceptions of the agents themselves seemed to transfer to the context in which these agents

exist, these results suggest that much care needs to be taken in the design of such agents to avoid unintended effects. An “untrustworthy” French business agent, for example, might lead to students leaving with more negative attitudes towards the French than when they arrived.

#### 10.4. FUTURE DIRECTIONS

The work presented in this dissertation offers many avenues for extending this research. The case is not yet closed on the success of learning from intercultural training environments. While I found that learning does occur while using such an environment, the lack of definitive measures and uncertainty about what constitutes an appropriate control condition leave much room for further exploration. One “gold standard” for assessment that might be considered in future studies is a behavioral measure conducted by a trained role-player. Although this is a resource-intensive method, gaining an understanding of how these skills translate from the virtual environment to real life is critical.

There is also still much work to be done in modeling agent dialog. There have been many successes in modeling agent affect, personality, culture, and other characteristics (Swartout, Gratch, Hill, Hovy, Marsella, Rickel, & Traum, 2006; Gratch & Marsella, 2001; also see Rehm, 2010). All of these features may play a part in generating dialog. Given the intricacies of human-human communication, however, there are many factors that remain. One of the most interesting directions for training interpersonal skills may be to give the agent a theory of mind (Premack & Woodruff, 1978); that is, not just to track the agents’ own state of trust and rapport with the learner, but to model the agents’ beliefs about the current state of the learner. Just as the learner should be always evaluating the agents’ trust in him and working to improve it, the agents’ behavior may change based on his current understanding of the learner. For

example, intercultural interactions in Iraq are changing rapidly as the people modify their behavior to be more acceptable to what they perceive as American practices. Enabling an agent with a theory of mind could help create more accurate agent behaviors and dialog, as well as assist in training flexible and adaptable thinking.

A final avenue of fruitful research is the continuation of a general model linking human-human communication theory, desirable interpersonal outcomes, and learning results. This dissertation has begun that work by looking to structural equation models to start positioning the disparate results into one coherent whole, but much work remains to be done. As of yet, this dissertation reports one of the few studies that have empirically linked interpersonal outcomes like rapport to learning results (Cassell, white paper), but as the field moves forward it will become more important to integrate our knowledge into a general model.

# II. REFERENCES

- Abbe, A., Gulick, L., and Herman, J. (2008). Cross-Cultural Competence in Army Leaders: A Conceptual and Empirical Foundation. *ARI Study Report 2008-01*. U.S. Army Research Institute for the Behavioral and Social Sciences.
- Aleven, V., & Koedinger, K.R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science*, 26 (2), 147-179.
- Allik, J., & McCrae, R.R. (2004). Towards a geography of personality traits: Patterns of profiles across 36 cultures. *Journal of Cross-Cultural Psychology* 35:13-28.
- American Council on the Teaching of Foreign Languages, (1996). *Standards for Foreign Language Learning: Preparing for the 21st Century*. New York: ACTFL.
- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, 84, 261-271.
- Anderson, J. R., Corbett, A., Koedinger, K., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences*, 4(2), 167-207.
- Ang, S., Van Dyne, L., Koh, C., Ng, K. Y., Templer, K. J., Tay, C., & Chandrasekar, N. A. (2007). Cultural Intelligence: Its Measurement and Effects on Cultural Judgment and Decision Making, Cultural Adaptation and Task Performance. *Management and Organization Review*, 3(3), 335-371.
- Aylett, R., Paiva, A., Vannini, N., Enz, S., André, E. (2009). But that was in another country: Agents and intercultural empathy. In *Proceedings of the International Conference on Autonomous Agents and Multi Agent Systems* 329-336.
- Babcock, L., Laschever, S. "Women Don't Ask: Negotiation and the Gender Divide", Princeton University (2003).
- Bailenson, J.N. & Yee, N. (2005). Digital Chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological Science*, 16, 814-819.
- Bateman, C. *Game writing: narrative skills for videogames*. Hingham, Massachusetts: Charles River Media. (2007).
- Baylor, A., & Kim, Y. (2004). Pedagogical Agent Design: The Impact of Agent Realism, Gender, Ethnicity, and Instructional Role. *Intelligent Tutoring Systems*: 592-603.

- Bennett, M.J. (1993). Towards ethnorelativism: A developmental model of intercultural sensitivity. In R.M. Paige (Ed.) *Education for the Intercultural Experience* (pp. 27-71). Yarmouth: Intercultural Press.
- Bickmore, T., & Cassell, J. (2001). Relational agents: a model and implementation of building user trust. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems ACM*, New York, NY, 396-403.
- Birukou, A, Blanzieri, E., & Giorgini, P. (2010). Implicit culture framework for behaviour transfer. In E.G. Blanchard and D. Allard (eds), *Handbook of research on culturally-aware information technology: Perspectives and models*, IGI Global, Hershey PA.
- Blanchard, E.G. (2009). Adaptation-oriented culturally-aware tutoring systems: When adaptive instructional technologies meet intercultural education. In H. Song and T. Kidd (eds), *Handbook of research on human performance and instructional technology*, Information Science Reference, Hershey PA.
- Blanchard, E.G., Mizoguchi, R., Lajoie, S.P. (2010). Structuring the cultural domain with an upper ontology of culture. In E.G. Blanchard and D. Allard (eds), *Handbook of research on culturally-aware information technology: Perspectives and models*, IGI Global, Hershey PA.
- Bochner, AP, Ellis, C., Tillmann-Healy, LM.: Relationships as stories: Accounts, storied lives, evocative narratives. In K. Dindia & S. Duck (Eds.), *Communication and personal relationships* (pp. 12-29). Chichester: John Wiley & Sons, Ltd. (2000)
- Bollen, K. (1989). *Structural Equation Models with Latent Variables*, Wiley.
- Bonvillain, N. (2008). *Language, culture, and communication: The meaning of messages*, 5<sup>th</sup> edition. Pearson Prentice Hall, Upper Saddle River NJ.
- Brooke, D., Postmes, T., Jetten, J., Dyson., S. (in preparation). Can We put "I" in "Team"? How the Group Dynamics of Social Identity Formation Influence Team Performance.
- Brown, P, & Levinson, S. (1987). *Politeness: Some universals in language usage*. Cambridge University Press, Cambridge MA.
- Bruckman, A. (1995). *The MediaMOO Project: Constructionism and Professional Community. Convergence*, 1:1, Spring.
- Bruner, J. *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press. (1986).
- Burak, A., Keylor, E., & Sweeney, T. (2005). PeaceMaker: A Video Game to Teach Peace. In *Intelligent Technologies for Interactive Entertainment* (pp. 307-310).
- Burt, R., (1992). *Structural holes*. Cambridge, MA: Harvard University Press.
- Byram, M. (1997). *Teaching and Assessing Intercultural Communicative Competence*. Clevedon: Multilingual Matters.
- Cassell, J. (1999). "Embodied Conversation: Integrating Face and Gesture into Automatic Spoken Dialogue Systems." In Luperfoy (ed.), *Spoken Dialogue Systems*. Cambridge, MA: MIT Press.

- Cassell, J. (2009). "Culture as Social Practice: Being Enculturated in Human-Computer Interaction". In C. Stephanidis (Ed.) Proceedings of HCII, (published as Universal Access in HCI, Part III. Berlin Heidelberg: Springer-Verlag), pp. 303–313.
- Cassell, J., & Bickmore, T. (2003) "Negotiated Collusion: Modeling Social Language and its Relationship Effects in Intelligent Agents" *User Modeling and User-Adapted Interaction* 13(1-2): 89-132
- Chou, C-Y., Chan, T-W., Lin, C-J.: Redefining the learning companion: the past, present, and future of educational agents. *Computers & Education*, 40, 255-269. (2003)
- Cohen, A., Ivry, R., & Keele, S.W. (1990). Attention and structure in sequence learning. *Journal of experimental psychology. Learning, memory, and cognition* , 16 (1).
- Cohen, A. D., Paige, R. M., Shively, R. L., Emert, H., & Hoff, J. (2005). *Maximizing study abroad through language and culture strategies: Research on students, study abroad program professionals, and language instructors*. Final Report to the International Research and Studies Program, Office of International Education, DOE, Minneapolis, MN.
- Cooper, C.R., Denner, J. (1998). Theories linking culture and psychology: Universal and community-specific processes. *Annual Review Psychology* 49:559-584.
- Cozby, P. C.: Self-disclosure: A literature review. *Psychological Bulletin*. (1973)
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*. 16, 297-334.
- Cushner, K., & Brislin, R. (1995). *Intercultural Interactions: A Practical Guide*. Second Edition. Thousands Oaks, CA: Sage Publications, Inc.
- D'Mello, S. K., Picard, R., & Graesser, A. C. (2007). Toward an affect-sensitive AutoTutor. *IEEE Intelligent Systems*, 22, 53–61.
- De Dreu, C. K. W., Weingart, L. R., & Kwon, S. (2000). Influence of social motives on integrative negotiation: A meta-analytical review and test of two theories. *Journal of Personality and Social Psychology*, 78, 889-905.
- Deaton, J., Barba, C., Santarelli, T., Rosenzweig, L., Souders, V., McCollum, C. Seip, J, Kerr, B, & Singer, M. (2005). Virtual Environment Cultural Training for Operational Effectiveness (VECTOR). *The Journal of Virtual Reality*. 8(3) 156-157.
- Despain, W.: *Professional Techniques for Video Game Writing*. (2008)
- Doerner, D. (2003). The mathematics of emotions. In Proceedings of the 5<sup>th</sup> International Conference on Cognitive Modeling. 75–79.
- Donoghue, S. (2000). Projective techniques in consumer research. *Journal of Family Ecology and Consumer Sciences*, 28, 47-53.
- Dweck, C.S., & Leggett E.L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95 256-273.

- Dweck, C.S., Davidson, W., Nelson, S., & Enna, B. (1978). Sex differences in learned helplessness: II. The contingencies of evaluative feedback in the classroom. III. An experimental analysis. *Developmental Psychology*, 14, 268-276.
- Earley, C.P. & Mosakowski, E. (2004). Cultural intelligence. *Harvard Business Review*, 82(10): 139-146.
- Eccles, J.S. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75-146). San Francisco: Freeman.
- Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. Cole (ed), *Nebraska Symposium on Motivation*, 19. University of Nebraska Press, Lincoln.
- Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. *Journal of Educational Psychology*, 100, 613-628.
- Endrass, B, Rehm, M, & André, E. (2010). Towards culturally-aware virtual agents systems. In E.G. Blanchard and D. Allard (eds), *Handbook of research on culturally-aware information technology: Perspectives and models*, IGI Global, Hershey PA.
- Ericsson, K., & Simon, H. (1993). *Protocol Analysis: Verbal Reports as Data* (2nd ed.). Boston: MIT Press.
- Falsetti, J., & Schweitzer, E. (1995). SchMOOze University: A MOO for ESL/EFL students. In M. Warschauer (Ed.), *Virtual connections: On-line activities and projects for networking language learners* (pp. 231-232). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.
- Fiske, S. & Taylor, S. (1991). *Social Cognition*. New York: McGraw-Hill.
- Ford, M.E, & Nichols, C.W. (1987). A taxonomy of human goals and some possible applications. In M.E. Ford & D.H. Ford (ed), *Humans as Self-Constructing Living Systems: Putting the Framework to Work*, p.289-311, Hillsdale, New Jersey, Erlbaum.
- Furstenberg, G., Levet, S., English, K., & Maillet, K. (2001). Giving a Virtual Voice to the Silent Language of Culture: The Cultura Project. *Language Learning & Technology*, 5(1), 55-102.
- Furstenberg, G. (1993). Comprehending Authentic French via Videodisk - A la Rencontre de Philippe. In J. Johnston & R. B. Kozma (Eds.), *When Lectures Fail: Educational Computing in the Humanities* (pp. 135-150). Ann Arbor, MI: University of Michigan Press.
- Gaertner, S. L., Dovidio, J. F., Anastasio, P. A., Bachman, B. A., & Rust, M. C. (1993). The common ingroup identity model: Recategorization and the reduction of intergroup bias. In W. Stroebe & M. Hewstone (Eds.), *European Review of social Psychology*, Vol. 4, pp. 1-26.
- Galinsky, A., & Mussweiler, T. (2001). Promoting good outcomes: Effects of regulatory focus on negotiation outcomes. *Journal of Personality and Social Psychology*, 81, 657-669.
- Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York: Palgrave/Macmillan.
- Gelfand, M., Brett, J. Eds. *The Handbook of Negotiation and Culture*. Stanford. (1983).

- Gratch, J. & Marsella, S. (2001). "Tears and Fears: Modeling emotions and emotional behaviors in synthetic agents," in *Proceedings of the 5th International Conference on Autonomous Agents*, Montreal, Canada, June 2001
- Gratch, J., Wang, N., Okhmatovskaia, A., Lamothe, F., Morales, M & Morency, L.P. (2007). Can virtual humans be more engaging than real ones? *12th International Conference on Human-Computer Interaction*, Beijing, China 2007
- Hall, J. (1984) *Nonverbal Sex Differences: Communication Accuracy and Expressive Style*. Johns Hopkins University Press.
- Hall, L., Vala, M., Hall, M., Webster, M., Woods, S., Gordon, A., & Aylett, R. (2006) FearNot's appearance: Reflecting Children's Expectations and Perspectives. In: *Gratch, J; Young, M; Aylett, R; Ballin, D. and Olivier, P,(eds) 6th International Conference, IVA 2006*, Springer, LNAI 4133, pp 407-419.
- Hays, M. J., Lane, H. C., Auerbach, D., Core, M. G., Gomboc, D., & Rosenberg, M. (2009). Feedback specificity and the learning of intercultural communication skills. In *Proceedings of the 14<sup>th</sup> International Conference on Artificial Intelligence in Education* (pp. 391-398).
- Hays, M. J., Ogan, A., & Lane, H. C. (2010). The evolution of assessment: Learning about culture from a serious game. In *Proceedings of the Workshop on Intelligent Tutoring Systems for Ill-Defined Domains at the 10th International Conference on Intelligent Tutoring Systems*.
- Henderson, J., Fishwick, P., Fresh, E., Futterknecht, R., & Hamilton, B.D. (2008). Immersive learning simulation environment for Chinese culture, In *Proceedings of Interservice/Industry Training, Simulation and Education Conf (I/ITSEC)*, paper 8334.
- Hill, R.W., Belanich, J., Lane, H.C., Core, M.G., Dixon, M., Forbell, E., Kim, J., & Hart, J. (2006). Pedagogically Structured Game-based Training: Development of the ELECT BiLAT Simulation, in the *Proceedings of the 25th Army Science Conference (ASC 2006)*, Orlando, FL.
- Hoffman, M. L. (1977). Sex differences in empathy and related behaviors. *Psychological Bulletin*, 1977b, 84, 712-722.
- Hofstede, G. (1984). The cultural relativity of the quality of life concept. *Academic Management Review*, 9(3):389-398.
- Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations*, 2<sup>nd</sup> ed. Sage, London.
- Hofstede, G. (2010). Scores of Hofstede's cultural dimensions. [http://www.geert-hofstede.com/hofstede\\_dimensions.php](http://www.geert-hofstede.com/hofstede_dimensions.php) . Accessed 18 February 2010.
- Horwitz, E., Horwitz M., & Cope, J.. (1986). Foreign language classroom anxiety. *Modern Language Journal*, 70, 125-32.
- House, R.J., Hanges, P.J., Javidan, M., Dorfman, P., Gupta, V. (2004). *Culture, leadership and organizations: The GLOBE study of 62 societies*, Sage, Thousand Oaks CA.
- <http://www.census.gov/Press-Release/www/releases/archives/population/012496.html>

<http://www.statcan.gc.ca/pub/91-551-x/91-551-x2010001-eng.pdf>

Huang, H-H., Cerekovic, A., Pandzic, I.S., Nakano, Y., Nishida, T. (2009). Toward a multi-culture adaptive virtual tour guide agent with a modular approach. *International Journal of AI & Society*. 24(3):225-235.

Hubal, R.C., Kizakevich, P.N., Guinn, C.I., Merino, K.D., & West, S.L. (2000). The virtual standardized patient: Simulated patient-practitioner dialogue for patient interview training. *Studies in Health Technology and Informatics*, 70, 133-138.

Ip, G. W. M., Chiu, C. Y., & Wan, C. (2006). Birds of a feather and birds flocking together: Physical versus behavioral cues may lead to trait- versus goal-based group perception. *Journal of Personality and Social Psychology*, 90, 368-381.

Johnson, D. W. & Johnson, R. T. (1990). Cooperative learning and achievement. In S. Sharan (Ed.), *Cooperative learning: Theory and research* (pp. 23-37). NY: Praeger.

Johnson, S. Everything bad is good for you: How today's popular culture is actually making us smarter. New York: Riverhead Books. (2006)

Johnson, W. L., & Valente, A. (2008). Tactical Language and Culture Training Systems: Using artificial intelligence to teach foreign languages and cultures. In *Proceedings of Innovative Applications of Artificial Intelligence* (pp. 1632-1639).

Johnson, W. L., & Valente, A. (2008b). Collaborative authoring of serious games for language and culture. In *Proceedings of SimTecT*.

Johnson, W.L., & Wu, S. (2008). Assessing aptitude for learning with a serious game for foreign language and culture. In *Proceedings of Intelligent Tutoring Systems*, 520-529.

Johnson, W. L., Mayer, R. E., André, E., Rehm, M. (2005). Cross-cultural evaluation of politeness in tactics for pedagogical agents. In C. Looi, G. McCalla, B. Bredeweg, and J. Breuker, (eds), *Proceedings of Artificial Intelligence in Education*. IOS Press, 298-305.

Johnson, W.L. (2007). Serious Use of a Serious Game for Language Learning. In *Proceedings of the Conference on Artificial Intelligence in Education 2007*.

Jones, E.E., & Nisbett, R.E. (1987). The actor and the observer: Divergent perception of the causes of behavior. In E.E. Jones (Ed.), *Attribution: Perceiving the causes of behavior* (pp. 79-94). Hillsdale, NJ: Erlbaum.

Jordan, P., Makatchev, M., Pappuswamy, U., VanLehn, K., & Albacete, P. (2006). A natural language tutorial dialogue system for physics. In: G. Sutcliffe & R. Goebel (eds.), *Proceedings of the 19th International FLAIRS Conference*. AAAI Press.

Kashima, Y. (2000). Conception of culture and person for psychology. *Journal of Cross Cultural Psychology*. 31(1):14-32.

Kim, J, Hill, R.W., Durlach, P, Lane, H.C., Forbell, E., Core, M., Marsella, S., Pynadath, D., & Hart, J. (2010). BiLAT: A game-based environment for practicing negotiation in a cultural context. *International Journal of Artificial Intelligence in Education*.

- Kim, Y., & Baylor, A. L. (2006). A Social-Cognitive Framework for Pedagogical Agents as Learning Companions. *Journal of Educational Technology Research and Development*, 54(6), 569-596.
- Kirkman, B.L., Lowe, K.B., Gibson, C.B. (2006). A quarter century of culture's consequences: a review of empirical research incorporating Hofstede's cultural values framework. *Journal of International Business Studies*, 37:285-320.
- Kline, P. (1983). *Personality: Measurement and theory*. London: Hutchinson.
- Koedinger, K., Anderson, J., Hadley, W., & Mark, M. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education*, 8, 30-43.
- Kramsch, C. (1993). *Context and Culture in Language Teaching*. Oxford: Oxford University Press.
- Kumar, R., Ai, H., Beuth, J., & Rosé, C. (2010). Socially-capable Conversational Tutors can be Effective in Collaborative-Learning situations. In *Proceedings of the Conference on Intelligent Tutoring Systems 2010*, Pittsburgh, PA.
- Landis, D., Bennett, J., & Bennett, M. J. (2003). *Handbook of Intercultural Training* (3rd ed.). Sage Publications, Inc.
- Lane, H.C., Core, M.G., Gomboc, D., Karnavat, A., & Rosenberg, M. (2007). Intelligent tutoring for interpersonal and intercultural skills. In *Proceedings of Interservice/Industry Training, Simulation, and Education Conference*, paper 1514.
- Lane, H.C., Hays, M.J., Core, M.G., Gomboc, D., Forbell, E., Auerbach, D., & Rosenberg, M. (2008). Coaching intercultural communication in a serious game. In *Proceedings of the 16th International Conference on Computers in Education*, 35-42.
- Lang, A., Bolls, P., Potter, R., Kawahara, K. (1999). The effects of production pacing and arousing content on the information processing of television messages. *Journal of Broadcasting & Electronic Media*, 43, 4, p 451 – 475.
- Lave, J., & Wenger., E. (1991) *Situated learning: Legitimate peripheral participation*. Cambridge University Press, New York.
- Leach, C. W., van Zomeren, M., Zebel, S., Vliek, M., Pennekamp, S. F., Doosje, B. Ouwerkerk, J. P., & Spears, R. (2008). Group-level self-definition and self-investment: A hierarchical (multicomponent) model of in-group identification. *Journal of Personality and Social Psychology*, 95, 144-165.
- Lee, K. M., Jung, Y., Kim, J., & Kim, S. R. (2006). Are physically embodied social agents better than disembodied social agents?: The effects of physical embodiment, tactile interaction, and people's loneliness in human-robot interaction. *International Journal of Human-Computer Studies* 64, 10 (October 2006), 962-973.
- Lee, J., & Marsella, S. (2006). Nonverbal Behavior Generator for Embodied Conversational Agents. In *Proceedings of the 6th International Conference on Intelligent Virtual Agents*, 243-255.

- Legree, P. & Pstotka, J. (2006). Refining situational judgment test methods. In Proceedings of the 25<sup>th</sup> Army Science Conference. Orlando, FL.
- Levin, D. Z.. (1999). *Transferring knowledge within the organization in the R&D arena*. Unpublished doctoral dissertation, Northwestern University.
- Lickel, B., Hamilton, D. L., Wierzchowska, G., Lewis, A., Sherman, S. J., & Uhles, A. N. (2000). Varieties of groups and the perception of group entitativity. *Journal of Personality and Social Psychology* 78, no. 2: 223–246.
- Litman, D., & Silliman, S. (2004). ITSPROKE: An Intelligent Tutoring Spoken Dialogue System. In Companion Proceedings of the Human Language Technology Conference (HLT/NAACL), Boston, MA.
- Lord, F.M. (1980). Applications of item response theory to practical testing problems. Lawrence Erlbaum Associates, Hillsdale, New Jersey.
- Lou, Y., Abrami, P. C., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research*, 71(3), 449-521.
- Lynch, C., Ashley, K., Alevan, V., & Pinkwart, N. (2006). Defining ill-defined domains; a literature survey. In V. Alevan, K. Ashley, C. Lynch, & N. Pinkwart (Eds.), Proceedings of the Workshop on Intelligent Tutoring Systems for Ill-Defined Domains at the 8th International Conference on Intelligent Tutoring Systems (pp. 1-10).
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integration model of organizational trust. *Academy of Management Review*, 20: 709-734.
- McAuley, E., Duncan, T. E., & Russell, D. W. (1992). Measuring Causal Attributions: The Revised Causal Dimension Scale (CDSII). *Pers Soc Psychol Bull*, 18(5), 566-573.
- McCollum, D. L., (2006). Students' social goals and outcomes. *Academic Exchange Quarterly*, 10(2), 17-21.
- McCrae, R. R., & Costa, P. T., Jr. (1996). Toward a new generation of personality theories: Theoretical contexts for the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 51-87). New York: Guilford.
- McNeil, N. M., & Alibali, M. W. (2000). Learning mathematics from procedural instruction: Externally imposed goals influence what is learned. *Journal of Educational Psychology*, 92, 734–744.
- McQuiggan, S., Rowe, J., Lee, S., & Lester, J. (2008). Story-Based Learning: The Impact of Narrative on Learning Experiences and Outcomes. In *Proceedings of the 9th International Conference on Intelligent Tutoring Systems*, Montreal, Canada, pp. 530-539.
- McSweeney, B. (2002). Hofstede's model of national cultural differences and their consequences: A triumph of faith – a failure of analysis. *Journal of Human Relations*. 55(1):89-118.

- Mendenhall, M.E., Stahl, G.K., Ehnert, I., Oddau, G., Osland, J.S., Kühlmann. (2004). Evaluation studies of cross-cultural training programs: A review of the literature from 1988 to 2000. In D. Landis, J. M. Bennett, & M. J. Bennett (Eds.), *Handbook of Intercultural Training*. 129-144.
- Mesquita, B., Frijda, N.H., & Scherer, K.R. (1997). Culture and emotion. In P. Dasen and T.S. Saraswathi (eds), *Handbook of cross-cultural psychology vol.2: Basic processes and developmental psychology*, Allyn & Bacon, Boston, 255-297.
- Miller, C.A., Ott, T., Wu, P., Vakili, V. (2010). Politeness and etiquette modeling: Beyond perception to behavior. In E.G. Blanchard and D. Allard (eds), *Handbook of research on culturally-aware information technology: Perspectives and models*, IGI Global, Hershey PA.
- Mintzberg, H. (1973). *The nature of managerial work*. New York: Harper Row.
- Mirrielees, E. (2008). *Story Writing*. Wildside Press LLC.
- Moon, Y., "Intimate self-disclosure exchanges: Using computers to build reciprocal relationships with consumers," Harvard Business School, Cambridge, MA. (1998).
- Mulken, S. v., Andre, E., & Muller, J. (1999). An Empirical Study on the Trustworthiness of Life-Like Interface Agents. In the *Proceedings of HCI 1999*, 152-156.
- Nazir, A., Enz, S., Lim, M.Y., Aylett, R., Cawsey, A. (2009). Culture–personality based affective model. *International Journal of AI & Society* 24(3):281-293.
- Neale, M. A., & Bazerman, M.H. (1983). The role of perspective-taking ability in negotiating under different forms of arbitration. *Industrial and Labor Relations Review*, 36, 378-388.
- Newcity, Janet, (1999) "Description of the 1998–1999 TISS Survey on the Military in the Post Cold War Era". As of February 3, 2011:  
<http://www.docstoc.com/docs/41544229/Description-of-the-1998-1999-TISS-Surveys-on>
- Nisbett, R.E., & Norenzayan, A. (2002). Culture and cognition. In: Medin D & Pashler H (eds), *Stevens' Handbook of Experimental Psychology, Third Edition, Volume Two: Memory and Cognitive Processes*. John Wiley & Sons, New York.
- Nydell, M. *Understanding Arabs: a guide for modern times* (2006).
- Ogan A, Aleven, V, & Jones, C. (2010). Advancing development of intercultural competence through supporting predictions in narrative video. *International Journal of Artificial Intelligence in Education*. Amsterdam: IOS Press.
- Ogan, A, & Lane, H.C. (2010). Virtual learning environments for culture and intercultural competence. In E. Blanchard and D. Allard (eds), *Handbook of Research on Culturally-Aware Information Technology: Perspectives and Models*, IGI Global, Hershey, PA.
- Okita, S.Y., Bailenson, J., & Schwartz, D. L. (2008). Mere Belief of Social Action Improves Complex Learning In S. Barab, K. Hay, D. Hickey (Eds.), *Proceedings of the 8th International Conference for the Learning Sciences*. New Jersey: Lawrence Erlbaum Associates. Utrecht, The Netherlands.

- Patrick, H. (1997). Social self-regulation: Exploring the relations between children's social relationships, academic self-regulation, and school performance. *Educational Psychologist*, 32, 209-220.
- Pelz, D. C., Andrews, F. M.: *Scientists in organizations: Productive climates for research and development*. New York: Wiley. (1966)
- Perry, P.P, Hechter, F.J., Menec, V.H., & Weinberg, L.H. (1993). Enhancing achievement motivation and performance in college students: An attributional retraining perspective. *Research in Higher Education*, 34(6), 687-723.
- Pettigrew, T.F., & Meertens, R.W. (1995). Subtle and blatant prejudice in Western Europe. *European Journal of Social Psychology*, 25(1): 57-75.
- Phares, E. J. *Locus of Control in Personality*. Morristown NJ: General learning press. (1976)
- Premack, D. G. & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1, 515-526.
- Prentice, D. A., & Miller, D. T. (1999). The psychology of cultural contact. In D. A. Prentice & D. T. Miller (Eds.), *Cultural divides: Understanding and overcoming group conflict* (pp. 1-19). New York: Russell Sage Foundation.
- Pynadath, D. V. & Marsella, S. C. (2005). PsychSim: Modeling Theory of Mind with Decision Theoretic Agents. In Proceedings of the International Joint Conference on Artificial Intelligence, pp. 1181-1186.
- Raybourn, E. M., & Waern, A. (2004). *Social Learning Through Gaming*. In Extended Abstracts of CHI Proceedings 2004. ACM Press.
- Reeves, B., & Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media like Real People and Place*. Cambridge University.
- Rehm, M. (2010). Developing enculturated agents: Pitfalls and strategies. In E.G. Blanchard and D. Allard (eds), *Handbook of research on culturally-aware information technology: Perspectives and models*, IGI Global, Hershey PA.
- Rehm, M., André, E., Bee, N., Endrass, B., Wissner, M., Nakano, Y., Nishida, T., & Huang, H-H. (2007). The CUBE-G approach – Coaching culture-specific nonverbal behavior by virtual agents. Proceedings of the 38th Conference of the International Simulation and Gaming Association (ISAGA), Nijmegen, 2007.
- Reinecke, K., Schenkel, S., & Bernstein, A. (2010). Modeling a user's culture. In E.G. Blanchard and D. Allard (eds), *Handbook of research on culturally-aware information technology: Perspectives and models*, IGI Global, Hershey PA.
- Rickel, J., & Johnson, W.L. (1997). Integrating pedagogical capabilities in a virtual environment agent. In Proceedings of the 1<sup>st</sup> International Conference on Autonomous Agents. 30-38.
- Riedl, M. Arriaga, R., Boujarwah, F., Hong, H., Isbell, J., & Heflin, L.J. (2009). Graphical social scenarios: Toward intervention and authoring for adolescents with high functioning autism. In *Proceedings of the AAAI Fall Symposium on Virtual Healthcare Interaction*, Arlington, VA.

- Riedl, M., & Young, R.M. (2003). Character-focused narrative generation for execution in virtual worlds. In *Proceedings of the International Conference on Virtual Storytelling*, 47-56.
- Rosé, C. P., & Torrey, C. (2005). Interactivity versus Expectation: Eliciting Learning Oriented Behavior with Tutorial Dialogue Systems, *Proc. Intl Conf. on HCI (Interact '05)*.
- Rothwell, S.L. (2010). Information technology and the development of a global safety culture: A nuclear perspective. In E.G. Blanchard and D. Allard (eds), *Handbook of research on culturally-aware information technology: Perspectives and models*, IGI Global, Hershey PA.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80. (Whole No. 609).
- Ryan, R.M., & Deci, E.L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychology* 55:68-78.
- Savicki, V. (2008). *Developing intercultural competence and transformation: Theory, research, and application in international education*. Sterling, VA: Stylus.
- Schank, R. C., & Berman, T. R. (2002). The pervasive role of stories in knowledge and action. In M. C. Green, J. J. Strange, & T. C. Brock (Eds.), *Narrative impact: Social and cognitive foundations* (pp. 287-313). Mahwah, NJ: Lawrence Erlbaum.
- Schank, R. C. (1990). Case-Based Teaching: Four Experiences in Educational Software Design. *Interactive Learning Environments* 1(4): 231–253.
- Scharifian, F. (2003). On cultural conceptualizations. *Journal of Cognitive Culture*. 3(3):187-207.
- Schunk, D. H., Pintrich, P. R., & Meece, J. *Motivation in Education: Theory, Research, and Applications* (3rd ed.). Prentice Hall (2007).
- Silvera, D. H., Martinussen, M., & Dahl, T. I. (2001). The Tromso Social Intelligence Scale, a self-report measure of social intelligence. *Scandinavian Journal of Psychology*, 42, pp. 313-319.
- Spiro, P., Glymour, G., & Scheines, R. *Causation, Prediction, and Search*, 2nd Edition, MIT Press (2000).
- Stewart, O., & Chakraborty, J. (2010). Culturally determined preferences: Automatic speech recognition (ASR) systems versus live help. In E.G. Blanchard and D. Allard (eds), *Handbook of research on culturally-aware information technology: Perspectives and models*, IGI Global, Hershey PA.
- Swartout, W., Gratch, J., Hill, R., Hovy, E., Marsella, S., Rickel, J., & Traum, D. (2006). Toward virtual humans. *AI Magazine* 27(2), 96–108.
- Sykes, J.M., Wendland, L., & Moore, P. (2008). *CROQUELANDIA: Helping learners develop authentic intercultural communication skills in a synthetic world*. Paper presented at EDUCAUSE Learning Initiative Annual Meeting, San Antonio, TX.
- Sykes, J. (in preparation) Learner requests in Spanish: Examining the potential of multiuser virtual environments for L2 pragmatic acquisition. *CALICO Journal*.

- Taylor, D. M., & Moghaddam, F. M. Theories of intergroup relations: International social psychological perspectives. 2nd. Ed. New York: Praeger. (1994).
- Tomlinson, B. (2005). Social characters for computer games. *International Journal of Interactive Technology Special Issue on Social Learning through Gaming*. Vol. 2, No. 2, p. 101-115.
- UNESCO (2007) UNESCO Guidelines on Intercultural Education. <http://unesdoc.unesco.org/images/0014/001478/147878e.pdf> . Accessed 18 February 2010.
- Urdu, T., & Maehr, M. (1995) Beyond a Two-Goal Theory of Motivation and Achievement: A Case for Social Goals. *Review of Educational Research*, Vol. 65, No. 3, 213-243.
- Wang, N., & Johnson, L. W. (2008). The Politeness Effect in an intelligent foreign language tutoring system. In the *Proceedings of the Conference on Intelligent Tutoring Systems*.
- Wang, N., Johnson, W. L., Rizzo, P., Shaw, E., & Mayer, R. E. (2005). Experimental evaluation of polite interaction tactics for pedagogical agents. In *Proceedings of the 10th International Conference on Intelligent User interfaces*. ACM, New York, NY, 12-19.
- Weingart, L. R., Bennett, R. J., & Brett, J. M. (1993). *The impact of consideration of issues and motivational orientation on group negotiation process and outcome*. *Journal of Applied Psychology*, 78(3), 504-517.
- Wentzel, K. R. (1989). Adolescent classroom goals, standards for performance, and academic achievement: An interactionist perspective. *Journal of Educational Psychology*, 81 (2), 131-142
- Wentzel, K.R. (1991). Social Competence at school: Relation between social responsibility and academic achievement. *Review of Educational Research*, 61 (1), 1-24.
- Wheless, L., & Grotz, J., (1977). "The Measurement of Trust and Its Relationship to Self-Disclosure," *Human Communication Research*, vol. 3, pp. 250-257.
- Wiggins, J. (1979). "A psychological taxonomy of trait descriptive terms," *Journal of Personality and Social Psychology*, vol. 37, pp. 395-412.
- Wunderle, W. (2007). Through the Lens of Cultural Awareness: A Primer for US Armed Forces Deploying to Arab and Middle Eastern Countries. Department of the Army.
- Yahja, A. (2006). Simulation Validation for Societal Systems. Unpublished dissertation.
- Yee, N., & Bailenson, J.N. (2006). Walk A Mile in Digital Shoes: The Impact of Embodied Perspective-Taking on The Reduction of Negative Stereotyping in Immersive Virtual Environments. In *Proceedings of PRESENCE 2006: The 9th Annual International Workshop on Presence*.

# APPENDIX A

## SAMPLE ITEMS FROM THE CULTURE ASSIMILATOR ASSESSMENT

Robert, an Englishman, has recently arrived in a Middle Eastern country and obtained a position as a private English teacher. He is required to obtain a work permit, and so presents himself at the appropriate government office to apply. He is told to fill out a form and return in a few days. When he returns and asks if the permit is granted, he is told there are some problems and to return in a few days. On two more visits he meets the same response and exasperatedly asks another teacher if this is normal. He is told that such delaying tactics are frequent and that he can avoid them by giving the official a small amount of money to expedite the process. Robert becomes very indignant at this and declares he will never resort to such bribery. However, after several more fruitless visits he slips the official some money and is subsequently granted his permit. He feels very bitter about the incident, however, and constantly denounces the corruption of these people to his fellow expatriates. How would you interpret the official's action so as to make it more acceptable to Robert?

- The official is not being discriminatory, as everybody is obliged to pay such bribes. Robert should not take it so personally.
- The payment could be regarded as equivalent to a tip for services, such as that given to a waiter or porter.
- The official does not demand any large sums of money, so he is not really doing anything seriously wrong.
- Such behavior is probably not seen as unethical by the official, so Robert should not try to impose his culturally influenced values upon someone from another culture.

"Adjustment to Japan has been much easier than I thought it would be," Ted Owens told his wife about a year after their move from the United States. Ted had been sent by an automobile company in Detroit to see if he could establish production facilities for transmission systems that would be built in Japan and imported into the United States. Having been told that negotiations take a long time in Japan, he was not disappointed that it had taken a year for a major meeting to be set up with his key Japanese counterparts. But the Japanese had studied the proposal and were ready to discuss it this morning, and Ted was excited as he left for work. At the meeting, people discussed matters that were already in the written proposal that had been circulated beforehand. Suddenly, it occurred to Ted that there was an aspect of quality control inspection that he had left out of the proposal. He knew that the Japanese should know of this concern because it was important to the long-range success of the project. Ted asked the senior person at the meeting if he could speak, apologized for not having already introduced the quality control concern he was about to raise, and then went into his addition to the proposal. His presentation was met with silence, and the meeting was later adjourned without a decision having been made on the whole manufacture-importation program. Because Ted thought that a decision would be made that day, he was puzzled. What was the reason for Ted's difficulty?

- Ted had brought up quality control, an issue about which the Japanese are very proud. The Japanese thought that Ted was questioning their commitment to quality control.
- Expecting a decision in a year was still unrealistic. Ted should have been more patient.
- Ted had brought up an issue on which there had not been prior discussion among the people somehow involved in that specific issue.
- Ted had asked the senior person about speaking; in actuality, there was a younger person present who was in charge and Ted should have deferred to that person.

Mr. Wong and Mr. Chang have known each other for a good number of years. They both have several children about the same age. Mr. Chang has two brothers who live in the States and have small family businesses. Both Mr. Chang and Mr. Wong have businesses that seem fairly prosperous. In fact it was the capital that Mr. Chang provided that enabled Mr. Wong's brothers to get started in their own flourishing livelihood. In recent times, however, Mr. Chang has had several difficulties. Interest rates are up and the unstable character of his country's monetary unit has caused him much financial loss. He has the opportunity to invest in a very promising venture but at the moment does not have enough cash. This venture could pull his whole operation up and over the slump he finds himself in, but his brothers do not have the large sum that he needs. Even though they seem to be doing well, the cash flow situation prevents additional investments. He considers carefully and then goes to Mr. Wong. Mr. Wong gladly lends him the money, and the venture does indeed turn out very profitably for Mr. Chang, who is then able to repay Mr. Wong with the interest immediately. Later that year, when their sons are applying for colleges, Mr. Wong calls on Mr. Chang (whose other son is already attending a university in the States), to help get his son into the school by using his connections. Mr. Chang acquiesces and secures a place for Mr. Wong's younger son. A few months later, Mr. Wong has a nephew who will be going to the States by himself. He asks Mr. Chang if perhaps his brothers could help him out and give him a job when he gets there. Again Mr. Chang complies without hesitation. Their relationship continues along in this manner with Mr. Wong calling on Mr. Chang for assistance in several more instances and Mr. Chang in some instances asking for several small favors as well. How can one explain Mr. Wong's attitude?

- Mr. Chang had obligated himself to Mr. Wong and it was his duty to help that also obligated Mr. Wong to him.
- This just shows the spirit of friendship and cooperation of friends helping one another.
- Mr. Chang was just taking as much as he could, but he was planning a way to get back at Mr. Wong.
- Mr. Wong was simply taking advantage of Mr. Chang's good fortune.

# APPENDIX B

## INSTRUCTIONS FOR THE SITUATIONAL JUDGMENT TEST

You will be presented with several short scenarios, each of which is followed by several possible actions. Please rate each action on a scale from zero to ten (i.e., 0 = very poor action, 5 = mixed/OK action, and 10 = very good action). You can assign the same rating to more than one action - you are not ranking them.

## SAMPLE ITEMS FROM THE SITUATIONAL JUDGMENT TEST

The town in Major Cross's area of operations has been relatively calm after a car bombing that killed two Iraqi civilians. Major Cross has just arrived for a meeting with Hamad at Hamad's home. They exchange greetings at the door and walk toward a table. Rate the following ways in which Major Cross could begin interacting with Hamad

- Remove his body armor and helmet.
- Avoid wasting Hamad's time with off-topic discussion; immediately begin discussing the market.
- Bestow some praise on Hamad (e.g., compliment house and/or furnishings).

Major Hale is meeting Achmed for the first time, and they are getting to know each other. Major Hale knows that Achmed, a Shia, is an important businessman with influence in this primarily Shiite area. Major Hale wants to develop rapport with Achmed. Rate the following courses of conversation that Major Hale could have during the meeting.

- Major Hale should explain how much authority he has as a LTC in the US Army
- Major Hale should show interest in Achmed's family's health.
- Major Hale should mention how much he enjoys the local food
- Major Hale should describe why Iran is one of the main reasons why there is still unrest in Iraq

Major Cross and Hamad are wrapping up their meeting, right on schedule. There are only a few minutes left in the allotted time for the meeting. Before the meeting, Hamad explained that he would need to leave at a particular time so that he is able to get to the mosque in time for afternoon prayer. Rate the following ways in which Major Cross could end the meeting.

- Make sure all agreements are clearly understood by Hamad. If the meeting runs a little longer than scheduled, it is OK.
- Revisit any results of the meeting that were unsatisfactory and try to work them out.
- Spend some social time together and remind Hamad that his friendship is valuable.

# APPENDIX C

## REFACTORED SOCIAL GOALS QUESTIONNAIRE

	Component			
	1	2	3	4
I would like to keep promises I made to Zahora.	<b>.661</b>	-.066	.093	.043
I had a goal to be a productive member of this team.	.194	-.082	<b>.718</b>	.057
Gaining the trust of Zahora was important to me.	<b>.721</b>	-.062	.090	.296
I hoped to be accepted by Zahora.	<b>.791</b>	.111	.141	.110
I desired to be liked by Zahora.	<b>.804</b>	.138	.110	.075
I wanted Zahora to view me with a high social status.	.130	<b>.647</b>	-.108	.140
I would like to control Zahora.	-.123	<b>.828</b>	-.021	.004
I consider it important to persuade Zahora.	.194	<b>.553</b>	-.421	.099
I wanted to control the interaction with Zahora.	-.114	<b>.792</b>	.177	-.244
When I have a difficulty, I would want Zahora to help me through it.	.404	.040	.131	<b>.674</b>
It was important for me to get advice from Zahora.	.140	-.097	-.012	<b>.846</b>
I wanted to be mentored by Zahora.	-.109	.326	.506	<b>.582</b>
I wanted to be a member of a team with Zahora.	.484	.004	<b>.645</b>	.103
I hoped to be a good group member.	.497	-.086	<b>.669</b>	.100
I hoped to establish a sense of community with Zahora.	<b>.711</b>	-.084	.208	.069

Rotated Component Matrix for all Social Goals items. Bold indicates that the item is most closely associated to the factor in that column. Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

Renamed Component	# items	$\alpha$
1: Acceptance	5	.821
2: Power	4	.688
3: Belongingness	3	.765
4: Assistance	3	.599