Journal of Educational Psychology

Learning Intercultural Communication Skills With Virtual Humans: Feedback and Fidelity

H. Chad Lane, Matthew Jensen Hays, Mark G. Core, and Daniel Auerbach

Online First Publication, September 9, 2013. doi: 10.1037/a0031506

Journal of Educational Psychology

0022-0663/13/$12.00 DOI: 10.1037/a0031506

Learning Intercultural Communication Skills With Virtual Humans: Feedback and Fidelity

H. Chad Lane, Matthew Jensen Hays, Mark G. Core, and Daniel Auerbach
University of Southern California

In the context of practicing intercultural communication skills, we investigated the role of fidelity in a game-based, virtual learning environment as well as the role of feedback delivered by an intelligent tutoring system. In 2 experiments, we compared variations on the game interface, use of the tutoring system, and the form of the feedback. Our findings suggest that for learning basic intercultural communicative skills, a 3-dimensional (3-D) interface with animation and sound produced equivalent learning to a more static 2-D interface. However, learners took significantly longer to analyze and respond to the actions of animated virtual humans, suggesting a deeper engagement. We found large gains in learning across conditions. There was no differential effect with the tutor engaged, but it was found to have a positive impact on learner success in a transfer task. This difference was most pronounced when the feedback was delivered in a more general form versus a concrete style.

Keywords: virtual humans, intelligent tutoring systems, sense of presence, feedback, intercultural communication

Pedagogical agents are animated characters that inhabit virtual learning environments and usually play the role of tutor (Haake & Gulz, 2009; Johnson, Rickel, & Lester, 2000) or peer (Y. Kim & Baylor, 2006). In these roles, the agent typically works alongside the learner to provide guidance (Arroyo, Woolf, Royer, & Tai, 2009), hold conversations (Graesser & McNamara, 2010), and encourage and motivate (Baylor, 2011), among many other forms of possible scaffolding. The role of pedagogical agents in virtual learning environments continues to expand. One use of pedagogical agents is replacing a human role player. Thus, instead of the agent assisting the learner with problems, it is the interaction itself with the agent that is intended to have educational value. Here, the agent is usually a *virtual human* playing a defined social role, with learners also playing a role and using specific communicative skills to achieve goals. For example, to prepare for an international business trip, a learner might meet with a virtual foreign business partner from the country of interest to negotiate a fictional contract agreement.

The technology challenge is to simulate social encounters in realistic ways and in authentic contexts. The pedagogical challenge is to design scenarios in ways that achieve the learning goals, maintain a high level of real-world fidelity, and stay within an ideal window of challenge (whatever that may be). The basic problems of doing this with virtual humans are eloquently stated by Gratch and Marsella (2005):

> These "virtual humans" must (more or less faithfully) exhibit the behaviors and characteristics of their role, they must (more or less directly) facilitate the desired learning, and current technology (more or less successfully) must support these demands. The design of these systems is essentially a compromise, with little theoretical or empirical guidance on the impact of these compromises on pedagogy. (p. 256)

The natural tendency is to build simulations to maximize realism since authentic practice opportunities are essential both for learner motivation and transfer to real-world contexts (Sawyer, 2006). However, some questions have been raised regarding the definition of realism as it applies to human communicative behaviors. Human variability due to personality and cultural differences suggest that virtual humans may have a small amount of flexibility to adapt to learners' needs while remaining realistic (Wray et al., 2009). Further, the design of virtual human scenarios can have a profound influence on the efficacy of the resulting learning experiences and should be carefully constructed to exercise the targeted communicative skills (Ogan, Aleven, Jones, & Kim, 2011).

In this article, we describe a game-based system for teaching intercultural communication skills and an associated intelligent tutoring system (ITS). We then present two studies investigating issues related to *fidelity* and *feedback,* both of which are important factors in virtual learning environments with virtual role players. The goal is to identify the influences of these factors on learner behaviors and on their acquisition of new communication skills. The article ends with a summary of the results, limitations of our studies, and a discussion of future research topics.

H. Chad Lane, Matthew Jensen Hays, Mark G. Core, and Daniel Auerbach, Institute for Creative Technologies, University of Southern California.

The project described here was sponsored by the U.S. Army. Statements and opinions expressed do not necessarily reflect the position or the policy of the U.S. Government, and no official endorsement should be inferred. We thank Julia Kim, Eric Forbell, and the BiLAT team for their contributions. We also thank Bob Wray and Brian Stensrud at SoarTech for their support.

Correspondence concerning this article should be addressed to H. Chad Lane, University of Southern California–Institute for Creative Technologies, 12015 Waterfront Drive, Los Angeles, CA 90094. E-mail: lane@ict.usc.edu

The Acquisition of Intercultural Communication Skills

Social skills (or equivalently, interpersonal skills) form the foundation for both simulation of communicative skills (using a virtual human) and for teaching communicative competence. Although no clear consensus has emerged on a single definition of social skills, most include the notions of choosing *appropriate* and *effective* communicative actions for a given context (Segrin & Givertz, 2003). Because of our specific focus on intercultural communication, we adopted the more precise definition of social skills as "the ability of an interactant to choose among available communicative behaviors in order that [she or] he may successfully accomplish [her or] his own interpersonal goals during an encounter while maintaining the face and line of his fellow interactants" (Wiemann, 1977, p. 198). It is worth noting that what constitutes success in a social interaction is not always obvious, rational, or consistent. Further, how interpersonal and communicative goals are established may or may not be evident (Spitzberg & Cupach, 2002).

Despite the peculiarities of human communication, the concept of social skills can be broken down in many different ways. One of the simplest is to consider two fundamental processes: *message reception* (Wyer & Adaval, 2003) and *message production* (Berger, 2009). Message reception refers to one's ability to both interpret social signals of others (such as speech and nonverbal behaviors) and infer meaning from the communicative acts conveyed by those social signals. The receiver must both have (a) the motivation to interpret and process the message and (b) the knowledge necessary to comprehend it (Wyer & Adaval, 2003).

Challenges to successful decoding of a message can come from contextual and pragmatic sources in the immediate environment, as well as from internal biases or beliefs. For example, assumptions one makes on the basis of stereotypes can greatly impede message reception. On the message production side, similar challenges arise. How one forms a message (consciously or not) depends again on context, beliefs, biases, and so on. Automated communicative skills are deeply rooted and, thus, difficult to modify in ways that enhance the odds of producing more effective outgoing messages. Nonetheless, the acquisition of novel communicative skills has been shown to follow the same patterns as other cognitive skills (Greene, 2003), and so the same techniques used to promote learning should apply. For example, it is known that repeated practice opportunities with feedback are an essential component in the development of expertise (J. R. Anderson, Corbett, Koedinger, & Pelletier, 1995; Kluger & DeNisi, 2004; Shute, 2008). We have applied these foundational principles in our work by providing a virtual practice environment for intercultural communication skills with automated feedback.

Virtual Humans as Role Players

Live role playing has a long history in education (Kane, 1964) and for teaching social interaction skills (Mendenhall et al., 2006; Segrin & Givertz, 2003). There are problems, however, with the approach. First, role playing in classrooms is not situated in a realistic context, which potentially limits transfer of the learned skills. Second, when peers act as role players, the attitudes, conversational content, and so forth of the role play may not be authentic or realistic. Third, expert human role players are generally believed to be the best option but are not cost effective and can be prone to inconsistency and fatigue. Although virtual humans have significant limitations, they undoubtedly address some of these complex issues (Cassell, Sullivan, Prevost, & Churchill, 2000; Lim, Dias, Aylett, & Paiva, 2012).

Empirical Support for Learning With Virtual Humans

Can virtual humans be effective role-players? Seminal work presented by Reeves and Nass (1996) in *The Media Equation* showed that people bring many of their usual assumptions about human–human interaction to computer-based interactions. Further, evidence is mounting that this result holds even more strongly when the computer presents a virtual agent (Gratch, Wang, Gerten, Fast, & Duffy, 2007; Pfeifer & Bickmore, 2011; Zanbaka, Ulinski, Goolkasian, & Hodges, 2007). In other words, people treat virtual humans as if they are real. Further, characters who provide *personalized* interactions are known to increase feelings of social presence, which in turn enhance learning (Moreno & Mayer, 2004). Learning can also be enhanced when learners choose to adopt social goals (e.g., "come to know your partner") while interacting with virtual humans (Ogan, Kim, Aleven, & Jones, 2009). Together, these results suggest that virtual humans can induce feelings of social presence in learners, that these feelings are enhanced through personalization and simulation of social and relational behaviors, and, ultimately, that we should expect a concomitant improvement in learning.

Early studies of the efficacy of virtual-human-based systems to teach intercultural skills seem to support this conclusion. Significant gains in overall learning were found for Tactical Iraqi (Surface, Dierdorff, & Watson, 2007) as well as Bilateral Negotiation Trainer (BiLAT; Durlach, Wansbury, & Wilkinson, 2008; J. M. Kim et al., 2009; Lane, Hays, Auerbach, & Core, 2010). Unfortunately, these and similar studies of other virtual learning environments for culture do not compare the systems with traditional (e.g., classroom-based) intercultural training, so it is not yet known if they are more effective than classroom-based learning (Ogan & Lane, 2010).

Virtual humans have been used successfully as role players in many contexts. For example, virtual agents have served as patients in clinical training (Johnsen, Raij, Stevens, Lind, & Lok, 2007), persons of interest in police officer training (Hubal, Frank, & Guinn, 2003), modelers of healthy play for children with autism (Tartaro & Cassell, 2008), victims and perpetrators of bullying in school settings (Aylett, Vala, Sequeira, & Paiva, 2007; Sapouna et al., 2010), and modelers of coping behaviors for mothers of children with serious illness (Marsella, Johnson, & LaBore, 2000). A key question for the intelligent virtual agent community is whether effectiveness will also increase with increased sophistication of the agents.

BiLAT: Teaching Bilateral Negotiation With Cultural Awareness

The context for our work is BiLAT, a game-based simulation for practicing the preparation, execution, and understanding of bilateral meetings in a cultural context (J. M. Kim et al., 2009). As part of an overarching narrative, learners prepare and meet with a series of virtual humans to solve problems in a fictional Iraqi city. Even though BiLAT's overall scope is much broader, our focus is on face-to-face meetings between learners and virtual characters and the basic intercultural communicative skills necessary to build trust and reach agreements. BiLAT meetings emphasize both message production and reception skills as discussed earlier.

In BiLAT, learners meet with one or more characters to achieve a set of predefined objectives. For example, the learner may need to convince a high-ranking local official to stop imposing certain taxes on his people or reach an agreement about who will provide security at a local marketplace. In all cases, the learner is required to adhere to Arab business cultural expectations and norms (Nydell, 2006), establish a relationship through building trust, and apply integrative negotiation techniques. Specifically, BiLAT is designed as a practice environment for learning win/win negotiation techniques, which promotes the idea that negotiation counterparts should proactively strive to meet each other's needs as well their own (Fisher, Ury, & Patton, 1991). To achieve this in BiLAT, learners must also apply their understanding of the character's culture to modify their own communicative choices (Landis, Bennett, & Bennett, 2004). BiLAT's focus is on the communicative intent and the structure of meetings and does not seek to teach new languages.

Screenshots of the BiLAT interface are shown in Figure 1. On the left is one of several navigation screens used in the game. On the right is the meeting screen, where learners spend much of their time during play. To communicate with the virtual character (i.e., apply message production skills), the learner selects from a menu of about 70 conversational actions that can vary between scenarios. For example, the learner can engage in small talk (e.g., "talk about soccer"), ask questions (e.g., "ask who is taxing the market" and "ask if he enjoys travel"), and state intentions (e.g., "say you are interested in finding a mutually beneficial agreement"), among other possibilities. Physical actions are also available (e.g., "remove sunglasses" or "give medical supplies"). Corresponding dialogue text is displayed in a dialogue pane while the character responds with synthesized speech and animated gestures.

BiLAT characters possess culturally specific models of how they expect meetings to progress. This progression includes expectations for an opening phase, a social period, a business period, and a closing social period. These phases are derived from live role playing sessions and cognitive task analysis performed with subject-matter experts early in the development of BiLAT (J. M. Kim et al., 2009). An example of a knowledge component taught by BiLAT is to *follow the lead of your host*. If a learner chooses an action that is not appropriate for the current phase of a meeting, the character will respond negatively. Trust, which is directly affected by the ability of the learner to take appropriate and effective actions, is a major factor in whether BiLAT characters will be agreeable or difficult. When trust is not established, it is often impossible to achieve all necessary agreements because the character will not be as interesting in working together. This means that learners often need multiple follow-up meetings with the same character to achieve objectives and to try different strategies for building trust.

Intelligent Tutoring in BiLAT

The intelligent tutoring system in BiLAT provides feedback to learners as they interact with characters. It is based on knowledge components that were identified during the initial cognitive task analysis and uses them to maintain a learner model and generate the content for feedback messages (Lane et al., 2008). Help can come in the form of *feedback* about a previous action (e.g., explain a reaction from the character by describing an underlying cultural difference) or as a *hint* about what action is appropriate at the given time. Both types of messages appear in the BiLAT dialogue pane (shown in Figure 1). Further, the system implements an adjustable model–scaffold–fade algorithm that reduces coaching support with increased time and successful interactions (Collins, Brown, & Newman, 1989).

Assessment of learner actions is driven by a model of intercultural interactions for Arab business meetings. We defined a typing system for the lowest level elements in the knowledge component hierarchy to facilitate the ITS' understanding of the different categories of message production. These include *required steps*, *usual steps* (but not required), *rules of thumb,* and *avoids* and are identified as tags on steps in recipes for achieving certain communication or negotiation goals. For example, the knowledge components include recipes for greeting, socializing, eliciting the perspective of the counterpart, asking about local events, and more. Which tags belong in which recipes was completed as a joint authoring effort between researchers and subject-matter experts.

These scenario-independent recipes were then mapped into communicative actions available in the game. This allowed the ITS to track learner actions in terms of knowledge components and evaluate actions as positive or negative instances of understanding those components. This authoring task was also performed jointly between researchers and subject-matter experts. It was often necessary to assign two links to some actions that had both positive and negative elements. For example, if a learner promises to give a character what she or he wants, the relationship with that char-

Figure 1. Screenshots of Bilateral Negotiation Trainer, a serious game for intercultural communication.

acter may be enhanced, but the promise could lead to problems down the road (e.g., the character's neighbors may grow jealous and demand the same favors). These trade-offs were highlighted by the ITS when they occurred—there were usually reasons to take the action (or respond) and reasons not to do it, and the best choice depended on the payoff and specific problem being solved.

Measuring Learning From BiLAT and the ITS

In the experiments described, we used two measures to evaluate learning produced by BiLAT and the ITS. The first measure was a *situational judgment test* (SJT). In general, SJTs present several domain-relevant scenarios, each of which is accompanied by several actions that the learner might perform in response to the scenario (Legree & Psotka, 2006). The participants provided Likert-scale ratings for each action (0 = *very poor action*, 5 = *mixed/OK action*, 10 = *very good action*). There were eight total scenarios and 28 total actions in the SJT (these items were provided by an external team at the U.S. Army Research Institute). The following is an example situation and to-be-rated actions:

> 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.
>
> (0–10) ___ Revisit any results of the meeting that were unsatisfactory and try to work them out.
>
> (0–10) ___ Make sure Hamad clearly understands all agreements. If the meeting runs a little longer than scheduled, it is okay.
>
> (0–10) ___ Spend some social time together and remind Hamad that his friendship is valuable.

To score the participant responses, we used ratings provided by three subject-matter experts. Understanding of the domain knowledge is defined as the degree that a participant's responses correlate with the experts' responses (Legree & Psotka, 2006). The test was administered in a pretest–posttest design, and so learning was defined as the increase in the correlation from pretest to posttest. Because the situational judgment test focused on the participants' ability to recognize and understand concepts about intercultural interaction, it measured learning at the lower levels of Bloom's taxonomy of cognitive skills (L. W. Anderson & Krathwohl, 2001).

The second measure was an in-game posttest that focused on a new issue (supply thefts from an Iraqi hospital rather than the market). During the participants' meetings with a hospital administrator, no feedback was provided. For each action that a participant selected during these meetings, we examined the probability that it was inappropriate. Participants who made fewer errors were said to have learned more about intercultural interaction than were participants who made more errors. We also examined the probability that the participant was able to successfully negotiate with the hospital administrator. Although it was a binary measure, success indicated that the participants were able to build up trust and consider their partner's needs effectively. Because the in-game posttest measured the participants' ability to apply what they learned about intercultural interaction, it measured learning at the middle levels of Bloom's taxonomy of cognitive skills (L. W. Anderson & Krathwohl, 2001).

Experiment 1: Fidelity and Presence

Although not the only method, one approach to measuring engagement is by investigating to what a degree a system can establish a *sense of presence*. One way to induce a sense of presence is to provide greater visual and auditory fidelity (e.g., more realistic graphics and sound; Lombard & Ditton, 1997). Intuitively, it seems that greater sensory fidelity should also promote better training; this is a common point of emphasis in training system design requirements. However, recent studies on the effect of presence on training suggest that engagement and effective outcomes are enhanced by greater sensory fidelity, but learning does not necessarily improve (Rowe, Shores, Mott, & Lester, 2011). An exception is the case in which a specific task domain requires high-sensory-fidelity simulation (e.g., a flight simulator), but most systems with greater sensory fidelity are not necessarily better trainers as a result.

Experiment 1 was designed to determine whether a *social* simulator must have high visual and auditory fidelity in order to effectively engage and instruct. We therefore created two versions of the system. Both versions had the high *social* fidelity of the standard BiLAT experience: rich characters, extensive dialogue, intricate character backgrounds, and so forth. But only one version had the rich visual and auditory experience; the other used a static, primarily text-based interface. On one hand, because BiLAT is essentially a social-skills trainer, the difference in visual and auditory fidelity may not have affected either presence or learning. On the other, given the tendency of people to treat virtual human interactions as being real (Gratch et al., 2007), we anticipated some advantages for the high fidelity version of BiLAT, including a deeper sense of presence and realism.

Method

Participants. The participants were 46 U.S. citizens (recruited from college campuses) who received $60 as compensation for approximately 3 hr of participation.

Measures. We used the SJT in a pretest–posttest design, as described previously. We also used the in-game posttest described and analyzed, for each participant, the number of actions they took and the amount of time they deliberated between actions. Participants who took more actions and deliberated for less time were thought to be less engaged or to be taking the experience less seriously.

We added a new measure to capture how engaged the participants were while playing BiLAT: the Temple Presence Inventory (TPI). The TPI is a validated battery of self-report Likert-scale ratings that attempt to measure how engaged or immersed one is in a multimedia experience (Lombard & Ditton, 1997). We used two subscales from the TPI: the *Social* subscale and the *Spatial* subscale. An example of a Social subscale item is "How often (1 = *never*, 7 = *always*) did it feel as if someone you saw/heard in the environment was talking directly to you?" An example of a Spatial subscale item is "How much (1 = *not at all*, 7 = *very much*) did it seem as if you could reach out and touch the objects or people you saw/heard?" (Items on the Spatial subscale that addressed sound or animation were removed.)

Design and procedure. After providing consent, the participants completed the pretest SJT (online, administered by a survey-hosting website). Within a few days, the participants arrived at our Institute and were randomly assigned to one of the two between-subjects conditions. In the three-dimensional (3-D) condition, the participants played BiLAT with the rich, immersive interface previously described. In the 2-D condition, the participants played BiLAT with a nonimmersive, static, text-focused interface (shown in Figure 2). The 2-D interface had neither animation nor sound but was otherwise equivalent to the 3-D interface. That is, the characters and coach functioned identically in both conditions. The participants then completed the in-game posttest (using the same interface as they used with the market scenario). They then completed the two subscales of the TPI described previously. Finally, they completed the posttest SJT.

Results

Presence and immersion. The primary results from Experiment 1 are split into Table 1 and Table 2. Table 1 presents the participants' self-reported presence, the number of meetings they conducted with each character, and the number of actions they took in each meeting.

As can be seen, there was a main effect of interface on self-reported presence. The 3-D interface yielded higher spatial presence ratings than did the 2-D interface: $t(44) = 3.091$, $p = .003$, partial $\eta^2 = .178$. The 3-D interface also yielded higher social presence ratings than did the 2-D interface: $t(44) = 2.542$, $p = .015$, partial $\eta^2 = .128$.

There was also a main effect of interface on how the participants interacted with the virtual characters. (A software error corrupted the logs for two participants. Their data did not contribute to this analysis.) The participants conducted more meetings in the 2-D interface than in the 3-D interface: $t(42) = 3.143$, $p = .003$, partial $\eta^2 = .190$. During each meeting, the participants performed more actions in the 2-D interface than in the 3-D interface: $t(42) = 2.546$, $p = .015$, partial $\eta^2 = .134$. Summed across meetings, participants performed nearly 50% more actions in the 2-D interface than in the 3-D interface in approximately the same amount of time. A similar pattern of results appeared in the in-game posttest. The 3-D interface, it appears, caused people to think more about their actions than did the 2-D interface.

Learning. Table 2 presents the participants' SJT scores and their performance on the in-game posttest.

Declarative knowledge. A repeated-measures mixed analysis of variance (ANOVA) revealed that there was not a main effect of interface on the participants' pretest–posttest gain: $F(1, 44) < 1$, *ns*. A main effect was not obscured by pretest differences between the two conditions; participants assigned to the 2-D interface did not score reliably higher than those assigned to the 3-D interface: $t(44) = 1.330$, $p = .191$. Thus, although the 3-D interface created more presence, it did not produce learning gains at the lower levels of Bloom's taxonomy (L. W. Anderson & Krathwohl, 2001).

Application and transfer. There was also not a main effect of interface on the probability of successful negotiation during the in-game posttest: $F(1, 44) = 1.208$, $p = .278$. Finally, there was not a main effect of interface on the probability of making an error during the in-game posttest: $t(40) = 1.536$, $p = .132$. Along with the SJT data, it seems clear that greater visual fidelity—and the spatial and social immersion it generates—does not appear to have a substantial effect on learning cross-cultural interactions as addressed in BiLAT.

Overall gains. We conducted additional analyses of the SJT data to examine the overall gains produced by interacting with BiLAT and the ITS. A repeated-measures ANOVA revealed that there was a main effect of practice on the improvement from pretest to posttest. Correlation with subject-matter experts increased from pretest ($M = 0.56$, $SE = 0.03$) to posttest ($M = 0.72$, $SE = 0.08$): $F(1, 44) = 40.039$, $p < .001$ partial $\eta^2 = .476$. Overall, it appears that BiLAT and the ITS can effectively increase knowledge about how to interact in a intercultural context.

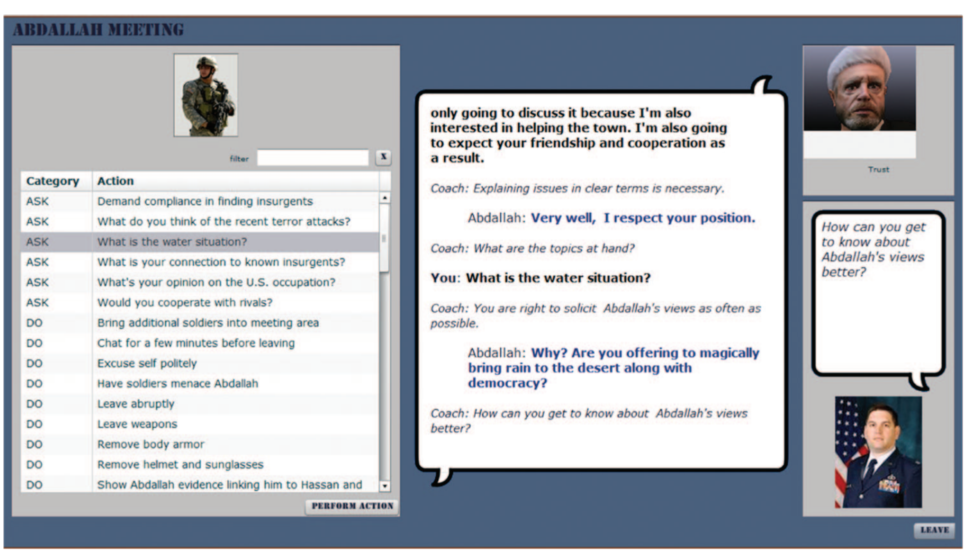

Figure 2. A screenshot of the two-dimensional interface for Bilateral Negotiation Trainer.

Discussion

The 3-D version of BiLAT (with animated virtual humans and synthesized speech) produced a greater sense of presence than did the 2-D interface. However, according to the SJT, there were no differences in declarative knowledge gains between the two conditions. Thus, the 3-D interface did not appear to improve BiLAT's teaching efficiency.

However, there were several differences in the how users in the two conditions interacted with the characters. Learners in the 3-D environment deliberated longer and, correspondingly, needed fewer actions in order to succeed. Research on rapport with virtual humans has shown that people react to virtual humans as if they are real (Gratch et al., 2007). One possible explanation for the interface-driven behavioral differences is that learners were more concerned about the impacts of their choices and thus thought them through more carefully. They may have been using that time to generate better mental simulations of the conversation. They may also have been establishing better expectations or generating better hypotheses about the mental state of their meeting partner. Future studies would be necessary to determine why users deliberate longer with embodied characters and how they are using that time.

Experiment 2: Formative Feedback

The results of Experiment 1 suggested that the content—not the appearance—of the system appeared to be responsible for learning. We therefore designed Experiment 2 to focus on the effectiveness of that content by examining the hints and feedback provided by the ITS. Some participants received formative feedback (Shute, 2008), which emphasizes productive revisions of knowledge. This feedback was very conceptual in nature (e.g., "Be sure to avoid appearing overly defensive or protective"). Other participants received very helpful, but very specific, assistance (e.g., "You are still in full combat gear, including your helmet and sunglasses").

Prior studies have found that learners who struggle during training eventually prosper as a result (Bjork, 1994; VanLehn, 1988). We believed that the specific feedback would be more helpful during practice and be appealing to learners (since it told them exactly what to do), but that the conceptual feedback would require the participants to deliberate more and think more deeply about the principles of cross-cultural interaction. We therefore predicted a greater increase in declarative knowledge (greater SJT improvement) and better transfer to new contexts (greater in-game posttest performance) for participants in the formative feedback condition.

Table 1
Temple Presence Inventory and Meeting–Interaction Data From Experiment 1 by Condition

	TPI				No. of interactions			
	Social		Spatial		Meetings		Actions	
Interface	Mean	*SE*	Mean	*SE*	Mean	*SE*	Mean	*SE*
2-D	2.77	0.19	2.30	0.20	13.67	0.91	17.70	0.90
3-D	3.49	0.18	3.21	0.21	10.30	0.60	15.09	0.54

Note. Participants' self-reported presence, the number of meetings they conducted with each character, and the number of actions they took in each meeting. TPI = Temple Presence Inventory; Social = Social subscale of the TPI; Spatial = Spatial subscale of the TPI; *SE* = standard error; 2-D = two-dimensional interface; 3-D = three-dimensional interface.

Table 2
Situational Judgment Test and In-Game Posttest Data From Experiment 1 by Condition

	SJT				In-game posttest			
	Pretest		Posttest		*p* (success)		*p* (error)	
Interface	Mean	*SE*	Mean	*SE*	Mean	*SE*	Mean	*SE*
2-D	.59	.04	.72	.03	.37	.11	.32	.02
3-D	.52	.04	.72	.03	.35	.10	.27	.02

Note. Participants' Situational Judgment Test (SJT) scores and their performance on the in-game posttest. 2-D = two-dimensional interface; 3-D = three-dimensional interface; *SE* = standard error.

Method

Participants. The participants were 47 U.S. citizens (recruited from college campuses) who received $60 as compensation for approximately 3 hr of participation.

Design and procedure. After providing consent, the participants completed the pretest SJT online (cf. Experiment 1). Within a few days, the participants arrived at our institute and were randomly assigned to one of the two between-subjects conditions. Some of the participants used BiLAT with a coach that provided hints and feedback that were exclusively *specific* to in-game actions. The coach for the other participants provided hints and feedback that were exclusively *conceptual*. The two versions of the coach[1] otherwise behaved identically in the two conditions (e.g., they chose when to provide feedback or hints based on the same policies).

The participants then completed the in-game posttest. They were then compensated and dismissed. A week later, the participants were e-mailed a link to the posttest SJT; 46 of the 47 participants completed it after an average of about 2 days.

Results

The primary results from Experiment 2 are presented in Table 3.

Declarative knowledge. There was not a main effect of feedback type on the participants' pretest–posttest gain: $F < 1$, *ns*. Their acquisition of declarative knowledge appears to not have been influenced by specific versus conceptual feedback.

Application and transfer. On the in-game posttest, there was not a main effect of coach type on the probability of a successful meeting outcome: $t < 1$, *ns*. However, there was a main effect of feedback type on the probability of making an error: $t(40) = 2.049$, $p = .05$, partial $\eta^2 = .095$. Even with equivalent declarative knowledge, the participants who encountered the conceptual coach were better able to interact with the new character in order to solve

[1] Typically, the coach follows a simple policy to decide whether to provide specific or conceptual feedback during practice (which is to try general first and then shift to concrete if the learner struggles).

Table 3
Effects of Feedback Type on Situational Judgment Test Scores and In-Game Posttest Performance

	SJT				In-game posttest			
	Pretest		Posttest		*p* (success)		*p* (error)	
Coach type	Mean	*SE*	Mean	*SE*	Mean	*SE*	Mean	*SE*
Specific feedback	.58	.04	.75	.02	.77	.09	.20	.02
Conceptual feedback	.56	.04	.74	.02	.76	.10	.15	.02

Note. SJT = Situational Judgment Test; *SE* = standard error.

the problem. This pattern of results is consistent with the notion that formative feedback can be superior to simple performance-oriented feedback (Shute, 2008). Further, the disparity between the in-game posttest results and the SJT is consistent with our belief that these two measures operated at different levels of Bloom's taxonomy of learning (L. W. Anderson & Krathwohl, 2001) taxonomy (Hays, Ogan, & Lane, 2010).

Overall gains. On the SJT, there was a main effect of practice with BiLAT. A repeated-measures ANOVA revealed that the participants' SJT scores increased from pretest to posttest: $F(1, 44) = 61.169$, $p < .001$, partial $\eta^2 = .582$. As in Experiment 1, it appears that the participants learned from their in-game experience.

Discussion

In Experiment 2, we tested the hypothesis that general feedback would be better for learning. The results suggested that conceptual feedback transfers more readily than does concrete feedback. Although we cannot conclude that concrete feedback never has a place (extreme versions of the ITS were used in the experiment), this study does suggest that for intercultural communication skills, a reasonable default setting is to use conceptual feedback first and then shift to concrete if future performance gains are not observed.

General Discussion

We sought to build a virtual environment for teaching intercultural communication skills with virtual humans in a specific context (i.e., Arab business practice). Our approach included the use of modern game technologies (i.e., a 3-D game engine) and an intelligent tutor to scaffold learners as they interacted with those characters. We conducted two experiments in which we found large overall gains in declarative knowledge as a result of interacting with this system. We found that the visual fidelity of the interface had significant impacts on learner behaviors, perceptions, and in-game success. We also found that conceptual feedback enhanced learners' ability to apply the targeted knowledge. At least in the context of intercultural communication, the nature of feedback had a much greater influence on learning than did visual fidelity.

Fidelity

For social simulations and virtual humans, the choice of where to invest development time is challenging: It is difficult to ignore any single dimension but also difficult to develop elaborate models for all relevant aspects of the communicative skill. Our studies suggest that for the message reception/production model of communication with a menu-based simulation of social interactions, learning of declarative knowledge is not affected by the richness of the sound and animations. However, given the complexities of social interactions, there are analogs to other domains that require higher fidelity simulations, like flight training. For example, virtual human agents used for teaching recognition of nonverbal behaviors (e.g., in deception detection training) would require a higher level of visual fidelity to properly capture and teach the subtle elements that are part of the knowledge being covered.

In BiLAT, learners are practicing the decision making involved in intercultural communication and learning what differences require attention. Variations in speech and nonverbal behaviors are not as critical, given these goals, and so the fidelity important to BiLAT learning has most to do with the content of the characters utterances, which is driven by the underlying models. The result that a richer interface engendered longer deliberations suggests that future studies are needed in order to understand the nature of how this time is being used: if virtual humans and high-fidelity graphics can be linked to greater attention to consequences of actions or more self-explanations, then future studies should seek to determine if these do in fact contribute to learning.

Feedback

Presence is often defined as forgetting that one is having a mediated experience. Thus, it is important to understand (a) if the use of unsolicited feedback interrupts this experience and (b) if that positively or negatively impacts learning. We found no evidence that the use of feedback (from a disembodied ITS) impacted the learner's sense of presence in either environment. How to deliver feedback optimally is an ongoing question for learning science researchers. Our study found benefits for using more general feedback that, we posit, required the learner to interpret the help and apply it to his or her own situation (i.e., it is *formative*). As our study only tested the extremes, an ITS that properly balances concrete feedback with conceptual feedback is more likely to be effective for the most learners. Future studies should focus on various algorithms for comparing different timing and content settings.

Limitations and Future Work

There were several technical and methodological limitations of the present studies.

Modality. Since communicating with BiLAT characters is accomplished through menu-based selection of actions rather than free speech input, learners are limited in what they can say and are most likely influenced by the choices that are available. This is vastly different than being required to generate utterances as they would in normal conversation. On the other hand, this design choice reflects current limitations of speech input and natural language understanding and does provide some structure for novice learners (J. M. Kim et al., 2009). Thus, because our measures focus on culture at the same level of abstraction as the game, it is unclear whether BiLAT practice with coaching would transfer to more realistic contexts. Because this is a critically important question, it suggests further study using a more elaborate (and expensive) a posttest measure using human role players or perhaps virtual humans that are capable of understanding free speech input.

Feedback. As discussed, feedback in BiLAT is delivered as text. An important lesson learned about the delivery of feedback resulted from early testing when we discovered that many learners were not reading the coaching messages. The reason, we found out, was that the BiLAT display included a readout of how much "trust" the character felt toward the learner. Thus, people played by selecting an action, *listening* to the character's response, and watching for changes in the trust meter. Because the messages were rendered only as text in the dialogue pane (on the other side of the screen from the trust meter), they often went unseen. We resolved this issue by hiding the trust meter in all of our studies and drawing attention to the coach and how it worked when introducing the participants to the system. It may be that having a trust meter or other visible "score" may have benefits for learning (e.g., rapid self-assessments), and thus our studies are limited in that they do not address the interaction of learning and game-like elements, like score, narrative, or challenge.

Further, our studies focus only on limited forms of feedback delivery. In Experiment 2, we considered two extreme versions of feedback—either completely general or completely concrete. It is likely that some balance between these is needed, and thus, continued study of how the generality of specificity of feedback should vary with context is needed. Also, the settings on our model–scaffold–fade algorithm were based on trial-and-error and observation of learners interacting with the system. Our goal was to strike a balance and provide the best level of support given the successes (or failures) of the learner, yet we had no theoretical support for the settings we used on this algorithm. Although there is substantial literature on the form of feedback (Shute, 2008), in general, we have found little guidance—empirical or theoretical—regarding the timing and optimal rates of fading of tutor support. This suggests future studies varying our algorithm and investigating the impact of these on performance and learning.

Measures. Unfortunately, both of the measures we used to gauge changes in learners have drawbacks. Although the in-game posttest does detect changes in learners' understanding of some culturally based rules of interaction, it is conducted within the environment used to teach those rules. As a result, it may only reflect shallow learning (i.e., learning to play the game rather than learning about the culture) or basic evidence of the existence of smooth learning curves. Because the ultimate test of learning is in face-to-face interactions with people from the target culture, in-game performance measures are inherently suggestive, at best.

Also, as discussed, although the SJT was designed by an external team, it may not be sufficiently precise to detect learning that occurs during BiLAT meetings. Further, it includes content from components of BiLAT that are not part of the tutoring system, such as preparation (i.e., research on counterparts) and broader scenario issues, like following up on commitments and social network changes based on the overarching narrative in the scenarios. It should also be noted that both these measures focus exclusively on the *message production* side of social interaction. Thus, even though the ITS does address *message reception* skills, our studies had no chance to detect any changes in a learner's ability to process and understand the utterances from the virtual human characters.

Another potential limitation of the SJT is that it uses identical prompts on the pretest and posttest. One could therefore argue that the overall gains from pretest to posttest we have reported merely reflect learning from the test. We took care to reduce this possibility by modifying some of the prompts to avoid divulging additional information (Asher, 2007). Also, because the SJT responses are numeric rather than potentially informative solutions (as in multiple-choice tests), it is unlikely that the participants used the SJT to guide their BiLAT experience so that their posttest score would be improved. Nevertheless, multiple counterbalanced versions of the SJT would be a more empirically sound measure.

Many other measures were possible, such as perspective-taking instruments (Paige, 2006) and measures to gauge perceived relationships with the agent (Ogan et al., 2011). Such measures are extremely important in intercultural development because much of it involves adjustment of one's own perspective on self, others, and more (Bennett, 1993), and so in future studies, investigators should more carefully address the role of feedback and fidelity on these factors while respecting the practical limits on testing time used during controlled studies.

Conclusion

This article began with the question of how virtual human role players might be used to enhance the learning of communication skills and highlighted the dearth of guidelines, principles, and empirical evidence for their design. Broadly, the results of our studies support the limited, but growing, body of literature (Durlach et al., 2008; Surface et al., 2007) that virtual humans can be used effectively to improve intercultural communication knowledge and skills. Generally, learners in both of our experiments showed gains in declarative knowledge from pretest to posttest. The key takeaway messages from these studies are that (a) the fidelity of such systems should be a function of the domain knowledge being taught and (b) feedback can be given in such a way that it enhances future performance and does not distract from the immersive nature of the system. Although our studies were not specifically "design" studies, further investigation of precise manipulations of virtual human content, behavior, and interaction modalities is definitely necessary. As with many advanced technologies (games, mobile devices, and so on), the number of available systems from the commercial and research sectors is rapidly growing, and so there is an urgent need for empirically derived guidelines and principles for using and scaffolding learning with virtual humans.

Many open questions remain about the use of virtual humans in social skills training and education. We believe future work is needed to develop and test new measures of learning and perspective change and to understand the role of feedback in these environments. As virtual humans continue to approach live human role players in realism, continued experimental research that focuses on the nature of these interactions, the sophistication of their implementations, and the role of supporting technologies such as intelligent tutoring is certainly merited.

References

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *Journal of the Learning Sciences, 4,* 167–207. doi:10.1207/s15327809jls0402_2

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of bloom's taxonomy of educational outcomes*. New York, NY: Longman.

Arroyo, I., Woolf, B. P., Royer, J. M., & Tai, M. (2009). Affective gendered learning companions. In V. Dimitrova, R. Mizoguchi, B. Du Boulay, & A. C. Graesser (Eds.), *Proceedings of the 14th International Conference on Artificial Intelligence in Education* (pp. 41–48). Amsterdam, the Netherlands: IOS Press.

Asher, H. (2007). *Polling and the public: What every citizen should know* (7th ed.). Washington, DC: CQ Press.

Aylett, R., Vala, M., Sequeira, P., & Paiva, A. (2007). FearNot! An emergent narrative approach to virtual dramas for antibullying education. In M. Cavazza & S. Donikian (Eds.), *Virtual storytelling: Using virtual reality technologies for storytelling* (Lecture Notes in Computer Science Vol. 4871, pp. 202–205). Berlin, Germany: Springer. doi: 10.1007/978-3-540-77039-8_19

Baylor, A. (2011). The design of motivational agents and avatars. *Educational Technology Research and Development, 59,* 291–300. doi: 10.1007/s11423-011-9196-3

Bennett, M. J. (1993). Toward ethnorelativism: A developmental model of intercultural sensitivity. In R. M. Paige (Ed.), *Education for the intercultural experience* (2nd ed., pp. 21–71). Yarmouth, ME: Intercultural Press.

Berger, C. R. (2009). Message production processes. In C. R. Berger, M. E. Roloff, & D. R. Roskos-Ewoldsen (Eds.), *The handbook of communication science* (2nd ed., pp. 111–128). Thousand Oaks, CA: Sage.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.

Cassell, J., Sullivan, J., Prevost, S., & Churchill, E. (Eds.). (2000). *Embodied conversational agents*. Cambridge, MA: MIT Press.

Collins, A., Brown, J. S., & Newman, D. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction* (pp. 453–494). Mahwah, NJ: Erlbaum.

Durlach, P. J., Wansbury, T. G., & Wilkinson, J. (2008). *Cultural awareness and negotiation skill training: Evaluation of a prototype semi-immersive system*. Retrieved from the Defense Technical Information Center website at http://www.dtic.mil/dtic/

Fisher, R., Ury, W., & Patton, B. (1991). *Getting to yes: Negotiating agreement without giving in*. New York, NY: Penguin Books.

Graesser, A., & McNamara, D. (2010). Self-regulated learning in learning environments with pedagogical agents that interact in natural language. *Educational Psychologist, 45,* 234–244. doi:10.1080/00461520.2010.515933

Gratch, J., & Marsella, S. (2005). Lessons from emotion psychology for the design of lifelike characters. *Applied Artificial Intelligence, 19,* 215–233. doi:10.1080/08839510590910156

Gratch, J., Wang, N., Gerten, J., Fast, E., & Duffy, R. (2007). Creating rapport with virtual agents. In C. Pelachaud, J. Martin, E. Andre, G. Chollet, K. Karpouzis, & D. Pele (Eds.), *Proceedings of the 7th International Conference on Intelligent Virtual Agents* (pp. 125–138). Berlin, Germany: Springer–Verlag.

Greene, J. O. (2003). Models of adult communication skill acquisition: practice and the course of performance improvement. In J. O. Greene & B. R. Burleson (Eds.), *Handbook of communication and social interaction skills* (pp. 51–91). New York, NY: Routledge.

Haake, M., & Gulz, A. (2009). A look at the roles of look & roles in embodied pedagogical agents: A user preference perspective. *International Journal of Artificial Intelligence in Education, 19,* 39–71.

Hays, M. J., Ogan, A. E., & Lane, H. C. (2010). The evolution of assessment: Learning about culture from a serious game. In C. Lynch, K. Ashley, T. Mitrovic, V. Dimitrova, N. Pinkwart, & V. Aleven (Eds.), *Proceedings of the 4th International Workshop on Intelligent Tutoring Technologies for Ill-Defined Problems and Ill-Defined Domains Held at the 10th International Conference on Intelligent Tutoring Systems (ITS 2010) in Pittsburgh, Pennsylvania* (pp. 37–44). Clausthal-Zellerfeld, Germany: Clausthal University of Technology.

Hubal, R. C., Frank, G. A., & Guinn, C. I. (2003). Lessons learned in modeling schizophrenic and depressed responsive virtual humans for training. In L. Johnson & E. Andre (Eds.), *Proceedings of the 8th International Conference on Intelligent User Interfaces* (pp. 85–92). Miami, FL: ACM Press.

Johnsen, K., Raij, A., Stevens, A., Lind, D. S., & Lok, B. (2007). The validity of a virtual human experience for interpersonal skills education. In D. Gilmore (Ed.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1049–1058). San Jose, CA: ACM Press.

Johnson, W. L., Rickel, J., & Lester, J. C. (2000). Animated pedagogical agents: face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education, 11,* 47–48.

Kane, P. E. (1964). Role playing for educational use. *Speech Teacher, 13,* 320–323. doi:10.1080/03634526409377396

Kim, J. M., Hill, R. W., Durlach, P. J., Lane, H. C., Forbell, E., Core, M., . . . Hart, J. (2009). BiLAT: A game-based environment for practicing negotiation in a cultural context. *International Journal of Artificial Intelligence in Education, 19,* 289–308.

Kim, Y., & Baylor, A. (2006). A social–cognitive framework for pedagogical agents as learning companions. *Educational Technology Research and Development, 54,* 569–596. doi:10.1007/s11423-006-0637-3

Kluger, A. N., & DeNisi, A. (2004). Feedback interventions: Toward the understanding of a double-edged sword. In T. F. Oltmanns & R. E. Emery (Eds.), *Current directions in abnormal psychology* (pp. 76–82). Upper Saddle River, NJ: Pearson Education.

Landis, D., Bennett, J. M., & Bennett, M. J. (2004). *Handbook of intercultural training*. Thousand Oaks, CA: Sage.

Lane, H. C., Hays, M. J., Auerbach, D., & Core, M. (2010). Investigating the relationship between presence and learning in a serious game. In J. Kay & V. Aleven (Eds.), *Proceedings of the 10th International Conference on Intelligent Tutoring Systems* (pp. 274–284). Berlin, Germany: Springer–Verlag.

Lane, H. C., Hays, M. J., Auerbach, D., Core, M., Gomboc, D., Forbell, E., & Rosenberg, M. (2008). Coaching intercultural communication in a serious game. In S. Chen, A. Mitrovic, & R. Mizoguchi (Eds.), *18th International Conference on Computers in Education* (pp. 35–42). Tapei, Taiwan: Asia–Pacific Society for Computers in Education.

Legree, P. J., & Psotka, J. (2006, November). *Refining situational judgment test methods*. Paper presented at the Proceedings of the 25th Army Science Conference, Orlando, FL.

Lim, M. Y., Dias, J., Aylett, R., & Paiva, A. (2012). Creating adaptive affective autonomous NPCs. *Autonomous Agents and Multi-Agent Systems, 24,* 287–311. doi:10.1007/s10458-010-9161-2

Lombard, M., & Ditton, T. (1997). At the heart of it all: the concept of presence. *Journal of Computer-Mediated Communication*, *3*(2).

Marsella, S. C., Johnson, W. L., & LaBore, C. (2000). Interactive pedagogical drama. In C. Sierra, M. Gini, & J. S. Rosenschein (Eds.), *Proceedings of the 4th International Conference on Autonomous Agents* (pp. 301–308). Barcelona, Spain: ACM Press.

Mendenhall, M. E., Stahl, G. K., Ehnert, I., Oddou, G., Osland, J. S., & Kuhlmann, T. M. (2006). Evaluation studies of cross-cultural training programs: A review of the literature from 1998 to 2000. In D. Landis, J. M. Bennett, & M. J. Bennett (Eds.), *Handbook of intercultural training* (pp. 129–144). Thousand Oaks, CA: Sage.

Moreno, R., & Mayer, R. E. (2004). Personalized messages that promote science learning in virtual environments. *Journal of Educational Psychology, 96,* 165–173. doi:10.1037/0022-0663.96.1.165

Nydell, M. K. (2006). *Understanding Arabs: A guide for modern times*. Boston, MA: Intercultural Press.

Ogan, A., Aleven, V., Jones, C., & Kim, J. (2011). Persistent effects of social instructional dialog in a virtual learning environment. In G. Biswas, S. Bull, J. Kay, & A. Mitrovic (Eds.), *Proceedings of the 15th International Conference on Artificial Intelligence in Education, Auckland, New Zealand* (Proceedings Series: Lecture Notes in Computer Science Vol. 6738, pp. 238–246). Berlin, Germany: Springer–Verlag.

Ogan, A., Kim, J., Aleven, V., & Jones, C. (2009). Explicit social goals and learning in a game for cross-cultural negotiation. In S. Craig & D. Dicheva (Eds.), *Proceedings of the Workshop on Educational Games at the 14th International Conference on Artificial Intelligence in Education, Brighton, UK* (pp. 51–58). Berlin, Germany: Springer–Verlag.

Ogan, A., & Lane, H. C. (2010). Virtual environments for culture and intercultural competence. In E. G. Blanchard & D. Allard (Eds.), *Handbook of research on culturally aware information technology: Perspectives and models* (pp. 501–519). Hershey, PA: IGI Global. doi:10.4018/978-1-61520-883-8.ch023

Paige, R. M. (2006). Instrumentation in intercultural training. In D. Landis, J. M. Bennett, & M. J. Bennett (Eds.), *Handbook of intercultural training* (pp. 85–128). Thousand Oaks, CA: Sage.

Pfeifer, L. M., & Bickmore, T. (2011). Is the media equation a flash in the pan? The durability and longevity of social responses to computers. In D. Tan, G. Fitzpatrick, C. Gutwin, B. Begole, & W. A. Kellogg (Eds.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 777–780). Vancouver, British Columbia, Canada: ACM.

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

Rowe, J. P., Shores, L. R., Mott, B. W., & Lester, J. C. (2011). Integrating learning, problem solving, and engagement in narrative-centered learning environments. *International Journal of Artificial Intelligence in Education, 21,* 115–133.

Sapouna, M., Wolke, D., Vannini, N., Watson, S., Woods, S., Schneider, W., . . . Aylett, R. (2010). Virtual learning intervention to reduce bullying victimization in primary school: A controlled trial. *Journal of Child Psychology and Psychiatry, 51,* 104–112. doi:10.1111/j.1469-7610.2009.02137.x

Sawyer, R. K. (2006). The new science of learning. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 1–16). Cambridge, MA: Cambridge University Press.

Segrin, C., & Givertz, M. (2003). Methods of social skills training and development. In J. O. Greene & B. R. Burleson (Eds.), *Handbook of communication and social interaction skills* (pp. 135–176). New York, NY: Routledge.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78,* 153–189. doi:10.3102/0034654307313795

Spitzberg, B. H., & Cupach, W. R. (2002). Interpersonal skills. In M. L. Knapp & J. A. Daly (Eds.), *Handbook of interpersonal communication* (pp. 564–611). Thousand Oaks, CA: Sage.

Surface, E. A., Dierdorff, E. C., & Watson, A. M. (2007). *Special operations language training software measurement of effectiveness study: Tactical Iraqi study final report* (Technical Report No. 2007010602). Raleigh, NC: SWA Consulting.

Tartaro, A., & Cassell, J. (2008). Playing with virtual peers: Bootstrapping contingent discourse in children with autism. In P. A. Kirschner, J. van Merriënboer, & T. de Jong (Eds.), *Proceedings of the Eighth International Conference for the Learning Sciences* (Vol. 2, pp. 382–389). Utrecht, the Netherlands: International Society of the Learning Sciences.

VanLehn, K. (1988). Toward a theory of impasse-driven learning. In H. Mandl & A. Lesgold (Eds.), *Learning issues for intelligent tutoring systems* (pp. 19–41). New York, NY: Springer. doi:10.1007/978-1-4684-6350-7_2

Wiemann, J. M. (1977). Explication and test of a model of communicative competence. *Human Communication Research, 3,* 195–213. doi:10.1111/j.1468-2958.1977.tb00518.x

Wray, R., Lane, H. C., Stensrud, B., Core, M., Hamel, L., & Forbell, E. (2009). Pedagogical experience manipulation for cultural learning. In S. Craig & D. Dicheva (Eds.), *Proceedings of the Second Workshop on Culturally Aware Tutoring Systems at the 14th International Conference on Artificial Intelligence in Education, Brighton, UK* (pp. 35–44). Berlin, Germany: Springer–Verlag.

Wyer, R. S., & Adaval, R. (2003). Message reception skills in social communication. In J. O. Greene & B. R. Burleson (Eds.), *Handbook of communication and social interaction skills* (pp. 291–355). Mahwah, NJ: Erlbaum.

Zanbaka, C., Ulinski, A., Goolkasian, P., & Hodges, L. F. (2007). Social responses to virtual humans: Implications for future interface design. In D. Gilmore (Ed.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1561–1570). San Jose, CA: ACM Press.

Received December 16, 2011
Revision received December 3, 2012
Accepted December 10, 2012 ■