Individualized Cultural and Social Skills Learning with Virtual Humans

H. Chad Lane
University of Southern California
Los Angeles, CA  USA

Robert E. Wray
Soar Technology, Inc.
Ann Arbor, MI  USA

1 Introduction

Pedagogical agents usually play the role of tutor (Johnson, Rickel, & Lester, 2000) or peer (Y. Kim & Baylor, 2006) in virtual learning environments. In these roles, the agent works alongside the learner to solve problems, ask questions, hold conversations, and provide guidance. Over the last decade or so, a new breed of pedagogical agents have emerged that do not play the role of expert or peer, however, but rather act as the object of practice. That is, instead of helping on the side, it is the interaction itself with the virtual character that is intended to have educational value. Here, the agent is usually a virtual human playing a defined social role in an interaction that requires the learner to use specific communicative skills to achieve some goal. For example, to prepare for an international business trip, one 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 accuracy, 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)

What are the implications of pedagogical demands on virtual human design? How can virtual humans facilitate learning? This chapter seeks to lay out the space of what is possible in using virtual humans to promote the acquisition of social and intercultural skills and describe a framework for adapting their behaviors to better meet the specific needs of learners.

The tension between fidelity of simulations built for educational purposes and the demands of learning is not new. Authentic practice opportunities are essential both for learner motivation and transfer to real-world contexts (Sawyer, 2006); however, substantial risk is associated with not providing the guidance novices need (Kirschner, Sweller, & Clark, 2006). Beginning learners generally lack the cognitive and metacognitive resources to effectively comprehend, process, or encode their learning experiences. The most common methods of providing the needed guidance are through a human instructor or an intelligent
tutoring system (Lane, et al., 2008); however, here we focus primarily on complementary methods that involve providing support for learning through the virtual human role players. Inspired by anecdotal statements from expert human role players who reported adjusting their behaviors based on observations of learners, we outline the dimensions of what is adjustable in virtual humans and discuss some examples of how virtual human role players might similarly adapt to meet specific learner needs. This line of research falls under the broader notion of pedagogical experience manipulation which involves adjustment of the learning experience based on the evolving needs of the learner (Lane & Johnson, 2008; Wray, et al., 2009).

*** INSERT TABLE 1 ABOUT HERE ***

Orchestrated practice can often be viewed as having two loops: an outer loop that involves selecting problems and an inner loop that involves taking a series of steps to solve that problem (Vanlehn, 2006). We define pedagogical experience manipulation as encompassing two key methods: configuration of the learning experience (the outer loop) and dynamic tailoring of the learning experience as it unfolds (the inner loop). Both of these activities can be applied in service to supporting a learner’s affect and/or cognitive learning (table 1), which interact in complex (and not yet fully-understood) ways (D'Mello & Graesser, in press; Kort, Reilly, & Picard, 2001). Dynamic tailoring comprises not only more familiar methods of directly supporting learning objectives, but also emerging methods that employ narrative adaptation and user-interface manipulations (Magerko, Stensrud, & Holt, 2006; Mott & Lester, 2006; Riedl, Stern, Dini, & Alderman, 2008) to engage and motivate (left column of table 1).

The chapter continues in the next section with a discussion of how humans acquire social and intercultural skills. Developmental models are presented that describe how people tend to progress when learning new intercultural and social skills. This discussion is followed by an overview virtual human research, including discussion of both the technology required to build them, empirical findings regarding their use in learning contexts, and consideration of various methods to adaptively control their behaviors to promote learning. It concludes with pertinent open questions and suggestions for future research on virtual humans.

2 Acquiring interpersonal and intercultural communication skills

Whether learning to solve algebra equations, to shoot free throws, or to play music, there is remarkable consistency in how human beings acquire and develop expertise – in general, people are more alike than they are different (Pashler, McDaniel, Rohrer, & Bjork, 2008). Skill acquisition is believed to occur in three general stages: (1) encoding of declarative knowledge, (2) strengthening of knowledge through practice, and ultimately, (3) automaticity (Anderson, 2005; Proctor, 2006; VanLehn, 1996). Stage 2, also known as the associative stage, is when learners confront impasses, make errors, and correct their understanding of the domain knowledge. Learning improves according to a power law – i.e., as the amount of practice increases, accuracy in executing the targeted skill improves and time to do it decreases according to a power function (Anderson, 2005, pp. 282 - 286). Studies have shown that these patterns seem to be universal, across a multitude of domains, cultures, and ages.

In this section, we briefly review related work on in the fields of communication and intercultural competence that focuses on learning. Although the skills in question are not as well-defined as many of those cited above, there is still little reason to believe that the development of interpersonal and intercultural skills is any different.
2.1 Acquisition of social 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, Segrin & Givertz (2003) state that “one can distill most of these definitions of social skills (and their associated aliases) to the ability to interact with other people in a way that is both appropriate and effective” (p. 136). Given the more specific focus below on intercultural communication, we adopt 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 and what interpersonal goals are adopted are not always obvious (Spitzberg & Cupach, 2002).

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, 2003). The assumption is that humans engage in communication when a desire to receive, deliver, or exchange content (of some kind) exists. Further, this process is usually governed by communicative goals (conscious or not) and an understanding of the kinds of plans that achieve them. These processes can fail in an almost countless number of ways and for equally as many reasons. Message reception and production skills are, nonetheless, the building blocks for participating in communication.

Message reception refers to one’s ability to interpret social signals of others and infer meaning from the communicative actions of others. Ultimately, the receiver must both have (1) the motivation to interpret and process the message, and (2) 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 based on 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, etc. Automated communicative skills are deeply rooted and thus, difficult to modify in ways that enhance the odds of a producing a message that will be successfully decoded by the hearer.

However, studies on how people acquire novel communicative skills – i.e., ones that require modification (or “reprogramming”) of automated, even seemingly innate skills – have shown that learning generally follows a power law in a way similar to acquisition of any cognitive skill (Berger, 2003). Thus, although making fundamental changes to our communicative abilities may seem like a learning challenge that poses more hurdles than many others, the same underlying principles for learning any new cognitive skill seem to apply. One such principle is that practice with feedback is an essential component for learning (Anderson, Corbett, Koedinger, & Pelletier, 1995), which is why the development of virtual humans is particularly relevant for intercultural and social skill learning.

2.2 Development of intercultural competence

Very good examples of biases and beliefs that have high potential to hinder message reception and production come from the domain of intercultural competence. A key aim of many intercultural training programs is to support learners in recognizing how their own beliefs influence their communicative
choices, and how those choices are seen and interpreted through the eyes of others. 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 (Landis, Bennett, & Bennett, 2004). Many of these methods are based on a classroom instruction model and seek to leverage peer interaction and debate to engage learners. Typically, the goal is to induce changes in knowledge, skills, and/or attitudes. Knowledge includes basic facts about a new culture, such as common values and beliefs, preferences for physical contact, and typical eating and drinking patterns. Skills usually refer to the learner’s ability to interact with someone from the new culture, including communicating their 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. Evaluations of intercultural training programs also tend to focus on these three dimensions (Mendenhall, et al., 2006).

The best training programs are grounded in underlying models of individual intercultural development and human learning. Such models seek to provide a theoretical account for the changes that occur as one experiences a new culture and integrates new knowledge into existing beliefs and understandings. For example, United States Peace Corps volunteers are told to expect four levels of cultural awareness when they begin a new assignment in a foreign country (1999):

1. **Unconscious incompetence**: minimal awareness of cultural difference or mistakes made; a state of “blissful ignorance”.
2. **Conscious incompetence**: basic realization of cultural difference; minimal understanding of underlying reasons or their significance.
3. **Conscious competence**: increased understanding of differences; deliberate behavioral adjustments are made to reduce cultural errors and misunderstandings.
4. **Unconscious competence**: culturally appropriate behavior is more or less automatic; one’s “instincts have been reconditioned.”

By reifying these stages for consideration by the learner, Peace Corps educators are promoting the idea of self-awareness, which is essential for intercultural growth. For example, to move from unconscious to conscious incompetence, one must begin to realize that what seems “normal” may be considered strange by people from another culture. It must be recognized that any peculiar observed behavior is likely in reaction to the learner. For people with underdeveloped self-awareness, there is a real risk these connections will fail to be made, leading them to conclude the strangeness they perceive is inevitable and beyond their understanding.

A similar, but more empirically tested model than the Peace Corps’ can be found in Milton Bennett’s *Developmental Model of Intercultural Sensitivity* (DMIS) (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 the positioning of one’s own culture in relation to others. An ethnocentric orientation implies that one perceives 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.

Theoretical models of intercultural development are not only important for designing training, but also could form a foundation for tracking a learner in an intelligent learning environment (Ogan & Lane, in press). An environment that adapts virtual human behaviors to meet the needs of the learner will clearly benefit from an estimate of that learner’s current stage of development. In the next section we outline the space of configurability in current virtual human implementations, and then in section 4, use the DMIS and Peace Corps models in examples to demonstrate how individualization could be accomplished.

3 Virtual humans as role players

Live role playing has a long history in education (Kane, 1964) and because it is interactive and situated, is a common strategy for teaching social interaction skills (Segrin & Givertz, 2003). There are problems, however, with the approach. Role playing in classrooms is not situated in a realistic context, and when done with peers, authenticity concerns are raised. Expert human role players are generally believed to be the best option, but are not cost-effective and can be prone to inconsistency (between different role players and due to fatigue). Virtual agents and humanoid robots have been developed to fill a variety of roles and with many different goals in mind. Generally, the purpose is to provide naturalistic communication with human users to establish social presence and support some shared goal. Rich and Sidner (2009) document several prominent examples and describe the expanded space of interaction afforded by humanoid robots and agents in the areas of engagement, emotion, collaboration, and social relationships (pp. 30-31). In this section, however, we review several systems that use virtual role players built specifically for learning.

3.1 Examples of virtual human role-players

Because of a widespread and increasing desire for international travel, cultural learning, intercultural communication, and language learning are popular targets for virtual human-based training. For example, BiLAT is a serious game that situates the learner in a narrative context to prepare and meet with a series of virtual humans to solve problems in a fictional Iraqi city. To succeed, it is necessary to understand the underlying narrative context, adhere to cultural expectations, build trust with town officials, and negotiate mutually beneficial agreements (J. M. Kim, et al., in press; Lane, et al., 2008). A similar structure is used in the Tactical Language family of serious games where the focus is on conversational language, communicative (including gestures), and intercultural competence (Johnson & Valente, 2008). Here, various instructional activities, such as listening and speaking practice, are integrated with immersive learning in 3D, simulated environments from various countries. Both of these systems maintain student models and carefully orchestrate scenarios to be appropriate (e.g., in terms of difficulty) for learners as they move through the game.

Another prominent example is the use of virtual humans for clinical training. Again, human role playing is common in doctor training programs to implement “standardized” patients that exhibit consistent symptoms for practice. Controlled studies have shown that virtual patients can be used to train psychiatric students in the classification of post-traumatic stress disorder (PTSD) cases. “Justin” and “Justina” are two such characters that exhibit symptoms of PTSD in their responses to clinician questions (Kenny,
In related work, virtual patients have been shown to be as effective as human role players for clinical interviewing skills, including body positioning and eye gaze (Johnsen, Raji, Stevens, Lind, & Lok, 2007).

Many other learning contexts have been used with virtual human role players to teach different kinds of communicative skills in social contexts. Some examples are police officer training in handling the mentally unstable (Hubal, Frank, & Guinn, 2003), healthy play for children with autism (Tartaro & Cassell, 2008), anti-bullying and coping behaviors for school children (Aylett, Vala, Sequeira, & Paiva, 2007; Sapouna, et al., 2010), and coping behaviors for mothers of children with serious illness (Marsella, Johnson, & LaBore, 2000). Across the wide spectrum of these applications, most of the individualization that occurs is (1) at the learner’s discretion, and (2) at the scenario level (e.g., to select appropriate characters to meet with). In other words, learner preferences are incorporated in many instances, but learners are grouped into broad categories for which specialized modes of the system may run. In the sections that follow, we provide some background on how people learn with virtual humans and then explore how the level of individualization might be pushed down into the dynamic behaviors of the characters themselves, towards true individualization.

3.2 Foundations of learning with virtual humans

Can virtual humans be effective role-players? Building on results from The Media Equation, which shows that people treat computers like people (Reeves & Nass, 1996), there is increasing evidence that this result holds even more strongly with virtual humans (Gratch, Wang, Gerten, Fast, & Duffy, 2007; Zanbaka, Ulinski, Goolkasian, & Hodges, 2007). In other words, people treat virtual humans as if they are real. Further, studies on relational agents have shown that people can form longer-term bonds with virtual humans that engage in social dialogue, display empathy, discuss future encounters, and more (Bickmore & Picard, 2005). 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 learning were found for Tactical Iraqi (Surface, Dierdorff, & Watson, 2007) as well as BiLAT (Durlach, Wansbury, & Wilkinson, 2008; J. M. Kim, et al., in press; Lane, Hays, Auerbach, & Core, in press). Unfortunately these studies do not compare the systems to traditional intercultural training, so there is no way to determine if they are more effective than classroom-based learning.

3.3 Adaptability of virtual humans

Given the richness and complexity of face-to-face human interaction, it should be no surprise that the space of adjustability in virtual humans is vast. Below, we first consider the distinction between nonverbal and verbal behaviors, and then touch on underlying models of cognition, emotion, and language.

*** INSERT FIGURE 1 ABOUT HERE ***
Figure 1. Expressions of skepticism, anger, umbrage, and defensiveness by ICT virtual humans (Kenny, et al., 2008; Swartout, et al., 2006).

Nonverbal behaviors. Observable, nonverbal behaviors during interactions with virtual humans are often a primary focus in studies of their communicative competency and fluidity. For example, the role of eye gaze, nodding, and gestures play a significant role in generating feelings of rapport in users (Gratch, et al., 2007). When no attempt is made to align nonverbal behaviors with the utterances of users (“non-contingent” responses), feelings of distraction and disfluency in speech follow. The implication for learning with virtual humans is that if their nonverbal behaviors are unnatural to the point of being a distraction, learning may be hindered.

Nonverbal behaviors play a large part in the expression of emotion and it is possible to convey a great deal of implicit feedback through them. There is staggering complexity that emerges from facial expressions alone, but also through gaze, body positioning and movement, and gesturing (examples are shown figure 1). Such signals also come in varying levels of intensity, as measured by onset, duration, and length (Ekman, 1993), and so these all represent adjustable parameters that would enable the system to dampen or amplify nonverbal backchannel feedback from the virtual human.

Content. The information conveyed and the words used to encode a message represent another critical dimension in the space of configurability. A message may have more or less content, more or less meaning, more or fewer emotive words, more or less explanatory content, and be at a higher or lower rate or tone. The “best” choice of content depends heavily on many factors, including the context of the simulated social situation (e.g., business vs. casual), the culture and personality of the virtual human (e.g., reticent vs. talkative), the familiarity of the character with the user, and more.

Cognitive, communicative, and emotional models. The most sophisticated virtual humans engage in complex, task-based reasoning and engage in coping behaviors that are based on appraisals of situations (i.e., whether something is desirable or should be avoided). Beliefs are formed using underlying representations of the dialogue, their intentions, desires, the task domain, their emotions – communicative behaviors, nonverbal and verbal, flow from these representations (Swartout, et al., 2006). For example, a virtual human in a highly distressed state will display very different nonverbal and verbal behaviors than one who is more hopeful or joyful.

Figure 2 shows one possible visualization of an implementation of emotional and personality variables in virtual humans (Marsella & Gratch, 2009; Traum, Swartout, Gratch, & Marsella, 2008). Sliders to the left indicate low intensity values, while those to the right represent high values. In the figure, the virtual human is set to a maximal state of distress, with low values for all other emotions. Emotions can change over the course of an interaction based on the learner’s ability to build trust and on what is revealed, promised, etc. and the display updates as emotions change. Personality traits are more static and lead directly to different patterns of interaction. The colors of the bars for the emotion dimensions indicate the general categories (red and yellow are more negative while green is positive), whereas the colors on the personality variables change over the course of an interaction depending on the “stance” of the virtual human. For example, the blue bar for “defensiveness” indicates a positive stance is being taken towards the current course of action (which is tracked in the dialogue manager).
This interface also allows a system user (such as an instructor) to modify emotion and personality settings. All of the sliders can be adjusted before a meeting (i.e., configuration) or during (i.e., tailoring). Significantly different values in the character’s emotion and personality values can lead to dramatically different reactions, beliefs, and results. Below, we discuss a few cases of using this adjustability to address learning goals.

*** INSERT FIGURE 2 ABOUT HERE ***

**Figure 2.** Adjustable emotional parameters for virtual humans with emotional models.

Also, what a human user says also has influence on these various states. For example, a threatening utterance might trigger a withdraw intention, which increases the likelihood of compliance, and culminates in terseness of language, folding of arms, and so on. Of course, the information that a speaker intends to communicate may vary greatly from how the message is encoded, as well as how it is decoded by the receiver. Perceived or actual misunderstanding between a learner and a virtual human role player can have a profound effect on the learner’s evolving understanding of the skills being practiced. Emotions are believed to play an important role in the adaptivity present in human cognition, and so computational models of emotion may provide a substrate for individualization of virtual human behaviors (Gratch & Marsella, 2005).

## 4 Adaptive training with virtual humans

Given the dimensions of adjustability discussed in the previous section, how could virtual human behaviors be adapted to promote learning during an interaction? To override the default behavior requires a belief on the part of the system that the learner will benefit from a particular adaptation as well as a model of what adaptations might pose some threat to the quality or believability of the experience. As discussed in section 1, this highlights the tension between fidelity and learning.

### 4.1 Learner modeling for social and cultural knowledge

Because of fundamental challenges presented by computationally modeling culture itself, modeling ones’ acquisition of cultural knowledge poses equal, if not greater challenges. Systems reviewed earlier maintain learner models that track progress in language and culture learning (Johnson & Valente, 2008) as well as social norms related to social interaction in international business (Lane, et al., 2008). Outside of the learning context, modeling a user’s culture has been used to personalize user interfaces. For example, using Hofstede’s cultural dimensions (2001), Reinecke & Bernstein (2009) showed that it was possible to predict which user interfaces would be rated most preferable for individual users. The idea of adapting learning environments based on the cultural background of the learner is equally as important. For example, various versions of the Tactical Language systems are available for speakers with different cultural backgrounds so that translations are accurate and cultural differences can be appropriately addressed by the system (Johnson, 2009).

To date, however, no intelligent systems model cultural learning or intercultural development at the level of those suggested by the theoretical models of the Peace Corps (1999), Bennett (1993), or any others (Landis, et al., 2004). Having an estimate of a learner’s stage of intercultural growth can have profound implications on his or her interpretation of observed behaviors, however. For example, a learner at the earliest stage of “unconscious incompetence” will need support in recognizing when cultural differences are apparent and with interpreting signals from others. Also, as discussed earlier, rushing to behavior
change too quickly can negatively impact growth. Specifically, a learner may very well be able to learn a new social behavior with minimal instruction; however, teaching that behavior without the underlying conceptual knowledge increases the chance of shallow learning (Bennett, 1993). It should be done at a later stage when an explanation for why it is appropriate and what cultural values drive the change can be provided and accepted by the learner.

Performing diagnosis to determine which developmental stage a learner may be experiencing poses significant challenges. A variety of psychometric measures have been proposed and evaluated that seek to place individuals at a particular stage of intercultural growth (Paige, 2006). These have proven useful for tracking the perspectives and opinions of learners over extended periods of time during intercultural training, but the results have not yet been integrated meaningfully into educational software for cultural learning. In addition, Bennett (1993) describes various behavioral cues that represent advancement through DMIS stages, as well as behaviors that suggest impasses. Automated detection of these behaviors in a virtual environment would provide stronger evidence for classification of a learner’s stage than self-report measures alone and so these advances represent important next steps for the field of intelligent learning environments for culture. In the next two sections, we outline what such an integration would enable and how it is essential for meaningful adaptation of virtual human behaviors.

4.2 Techniques for tailoring the behavior of virtual human behaviors

Although the focus in this chapter is on the turn-by-turn behaviors of virtual humans, pedagogical experience manipulation applies much more broadly and can be motivated by many factors. Adjustments can be made based on pedagogical strategies (such as fading learning support with time), for purposes of orientation, to manage user engagement, challenge, emotional state, and difficulty, among others (Wray, et al., 2009). Thus, adaptation of specific virtual human behaviors should be viewed in service to these broader categories of learning support.

Specific adaptations of virtual human behavior are discussed below. These are intended to address the following categories of support for learning social and cultural skills:

- To support recognition of errors or when ideal actions are taken.
- To give a hint about an action that will have a positive outcome.
- As an explanation for observed reactions, emotional state changes, or opinion shifts.
- To elicit a self-explanation from the learner to identify relevant domain principles.

Although intended as a framework for delivering feedback implicitly (through a character), these categories are intentionally similar to the goals behind the use of explicit feedback. In fact, implicit and explicit feedback can be viewed as having a shared goal: to provide formative feedback with the aim to convey information “to the learner that is intended to modify his or her thinking or behavior for the purpose of improving learning” (Shute, 2008, p. 154). Although this support can certainly be provided by an intelligent tutoring system (Johnson, Vilhjalmsson, & Marsella, 2005; Lane, et al., 2008), there may be benefits if pedagogical goals can be achieved organically. For example, it may be less of an interruption to the flow of the experience, or the learner may feel as if she or he did not need excessive support in order to succeed. There are tradeoffs associated with these benefits which are addressed again in the final section of the chapter.
To provide support for learning through the adaptation of virtual humans, there needs to be (1) a pedagogical purpose (a defined curriculum or learner model can provide this), (2) a tailoring tactic available that achieves the purpose, and (3) heuristics for detecting when and if the adaptations represent a threat to the fidelity of the experience. Tailoring tactics can occur either at the surface level, by adjusting nonverbal behaviors or content, or by adjusting the underlying models in ways that may produce different surface behaviors (see figure 2). Below, we organize the suggested tactics according to the pedagogical purposes described above, and link them to idealized learner model contents.

**Recognition.** A learner who is in the early stages of exposure to a new culture – especially if it is the first new culture they have experienced – will only have their own culturally biased methods for recognizing and interpreting feedback from people of that culture. For example, if a learner is in the Peace Corps’ “unconscious incompetence” stage, she or he has yet to consistently identify when communicative actions fail or succeed. To bring greater attention to such failures or successes, a tactic to **amplify** an existing behavior may be helpful. For example,

- If a learner brings up a taboo topic, angry (or saddened) facial cues and hand gestures could be exaggerated to capture the attention of the user.
- The content of utterances could also be modified to use emotionally charged vocabulary that draws attention to an inappropriate (or just sensitive) topic mentioned by the learner. A prototype system has been implemented using such an agent-based dynamic tailoring architecture (Wray, et al., 2009).
- For positive feedback, similar adjustments of facial responses (e.g., smiling) as well as gestures (e.g., hand to heart or a handshake) could be amplified to reinforce the learner behavior.

Much less obvious manipulations could be made to underlying models that might produce immediate or even delayed feedback. For example, characters in social simulations often model a version of trust of the learner (Johnson, et al., 2005; J. M. Kim, et al., in press; Swartout, et al., 2006). As the interaction progresses, the quality and appropriateness of the learner’s actions cause trust to go up and down. Amplifying the positive or negative impact of certain communicative actions could therefore lead to different outcomes, both immediate and longer-term.

Conversely, in some contexts it may make sense to **dampen** a virtual human reaction in order to draw attention away from a certain topic. For example, if the user commits a minor cultural error and is concentrating on an advanced communicative skill (e.g., negotiation), it may be preferable to delay discussion of the error until a reflective period. In these instances, recognition is less of a priority. The system may have a belief that the user has the skill and may simply be committing a **slip**. That said, systematic and/or multiple slips may indicate deeper misconceptions and there may be value in helping the learner have practice recovering from such slips. So, this pedagogical approach should be represented in the instructional model and the learner model should be equipped to support such decisions.

**Hinting.** Sometimes learners reach impasses and cannot progress without some help. Although less intuitive, some limited options are available for allowing virtual humans to make suggestions to the learner. Here, it might be related to a personality trait such as **helpfulness**, or even to an awareness by the virtual human that the learner is “new” to the (virtual) country and would benefit from some friendly advice. However, what constitutes an impasse or success in social contexts varies tremendously (Spitzberg & Cupach, 2002). Also, the goals of the participants and the context can have profound
implications on the limits of what is considered realistic. For example, it is unlikely that a person of interest in an investigation is going to give tips on how to properly elicit information from him or herself. However, in a cultural or informal social context, there are some possibilities:

- For learners who are at the “conscious incompetence” stage, hints may be given as friendly advice from the character highlighting an alternative communicative action that would produce a better outcome. For example, a virtual child may indicate that he or she wants candy.
- In a more serious context, a virtual human might ask about what it is the learner should do at a specific time (e.g., “Would you like to ask me about the contract now?”), or simply bring up a topic for conversation that would normally be the learners responsibility.

In some contexts it may even be permissible to allow the learner to ask the character what to do. The believability of each of these options depends so much on the context, it is again indicative of the difficulty associated with developing heuristics for detecting threats to fidelity.

**Explanation.** Pride in one’s knowledge and/or culture can be leveraged to provide explanations for observations made by the learner. As usual, these explanations can always come from an external source, like a tutor or web page; however, something directly from a character may be just as—if not more—impactful given that it is situated in the moment of the interaction. Again, concerns of fidelity greatly depend on the social context (e.g., a child will probably not provide a deep explanation of a cultural value like justice or the importance of family). Some examples of implicit explanations include:

- If mastery of a specific cultural topic tracked by the learner model is estimated as low, the content of the character’s utterance could be augmented with some additional information about the relevant concept. This approach has also been implemented using the dynamic tailoring architecture mentioned previously (Wray, et al., 2009).
- A learner in a denial stage of the DMIS (an early ethnocentric stage) may be unwilling to accept that a cultural difference even exists based on basic interactions with the character. In these instances, a clarification a relevant domain concept (or belief of the character) may support acceptance of that difference. If it is more deeply rooted, it may be necessary to amplify the characters emotions to indicate that a lack of recognition of the difference is upsetting.
- For one in an early stage of development, a character may help them establish a causal link between a communicative action and a negative (or positive) result via additional content. For example, the character may state “I’m not sure why you would ask me if I want a drink when you are aware that people in my country do not approve of alcohol.” Similarly, the character could also bring up a communicative action from the past to draw the connection.
- A learner in a later ethnocentric stage may be aware when an error occurs, but unwilling to take blame or perhaps prefers to place the onus on the virtual human to be the one who should adapt. In these circumstances, explanations may be particularly critical.

An advantage that tailored explanations might have over one from a tutor would be that when used properly, are situated in practice. Thus, a character who conveys a personal narrative with emotional content can do so as part of the relationship building process with the learner. Further, more emotionally laden versions of these messages are likely to be more memorable (Gratch & Marsella, 2005).
Self-explanation. Self-explanations enhance a learner’s ability to integrate new knowledge with existing knowledge. Further, such explanations can be facilitated when prompted (Chi, Leeuw, Chiu, & LaVancher, 1994). How might a virtual human role-player, in the context of a simulated social interaction, elicit self-explanations from a learner? In a way similar to hinting, it may be awkward to imagine a virtual human asking the learner to describe their understanding of the target culture. However, in some contexts it might be plausible:

- A learner in the DMIS stage of minimization (where the learner focuses on how his or her culture is similar to the target culture and minimizes the differences) might be engaged in a conversation by the virtual human to discuss their two cultures openly. In this context, the character could ask “What have you noticed about our cultures that are different?” This might support the learner in understanding that minimization is not equivalent to acceptance, and highlight the need for healthy construal of cultural differences.
- To achieve the integrative benefits of self-explanation, virtual humans could ask the learner to describe their experiences with other virtual humans and use that to bridge into eliciting self-explanations about choices made in those interactions.

Natural language dialogue system technology may need to advance in order to achieve the full effect of these tactics, although it is worth noting that pedagogically effective prompting can be achieved with little or no dialogue modeling (Chi, et al., 1994). Of course, self-explanations are likely to be occurring in the mind of learner anyway (as part of their deliberative processes in the communication), and so it may be worthwhile to look for evidence in the content of their contributions or to reserve discussion of them for a reflective period with explicit feedback.

As mentioned earlier, many of the techniques above address shared goals with those of a tutor that provides extrinsic feedback (Shute, 2008). There may be times when provision of intrinsic feedback through a character is culturally inappropriate (i.e., the “realism filter” detects an out-of-bounds tailoring action), but there is still a good reason to provide the feedback. For example, if the learner commits a minor gaffe but the virtual role player is interested in helping maintaining the learner’s face, this would be an ideal situation for feedback because it should probably not go unnoticed. In sum, it seems that intrinsic and extrinsic feedback, while potentially redundant in many cases, may be complementary in many ways – further research is needed to fully explore their relationship.

Given the short span of time that high-fidelity virtual learning environments have been available (Ogan & Lane, in press), very few systems implement dynamic tailoring or attempt to balance it with explicit feedback from an intelligent tutor. In each of the tailoring examples described in this section, the basic aim is to identify a pedagogical need, identify an appropriate tailoring technique, and then execute it if realism is not threatened. Significant research is needed along each of these dimensions.

5 Conclusion

It has been suggested that individualized learning experiences and personalized curricula may represent the next evolutionary stage for education (Christensen, Horn, & Johnson, 2008; Collins & Halverson, 2009). The vision of the future is one dominated by self-directed learning with computers playing a significant role in the management and provision of the individualization. For decades, artificial intelligence researchers have pursued this vision through the development of adaptive learning
environments that dynamically adjust content and learning support to meet the needs of individual learners (Shute & Zapata-Rivera, 2007). In this chapter, we have explored how these ideas take shape in the context of learning with virtual human role players and how they might be adapted to meet specific learner needs.

For virtual human role players to adapt based on pedagogical aims, it is likely that more sophisticated learner models will be necessary. Building learner models for domains such as cultural learning and interpersonal skills is no simple task, but even crude distinctions can be helpful. For example, if an error is made by a true beginner, the character might bring up the underlying cultural difference in their response (a content adjustment). Other learners would get the standard simulation response. We have developed a prototype dynamic tailoring architecture, that runs in BiLAT (J. M. Kim, et al., in press) to intelligently select and modify character utterances based on their pedagogical content and the evolving skills of the learner (Wray, et al., 2009).

An important, recurring question is whether intrinsic adaptations threaten fidelity and the implications to learning when that happens. For example, if a learner comes to rely on intrinsic scaffolding from virtual characters, then the learning in the virtual environment could have a negative effect on far transfer in the field. Similarly, if learners recognize the characters are secretly “helping”, does it ruin the fantasy of a realistic conversation? If so, how does this realization impact learner affect and motivation? These concerns are legitimate and evaluating how to ensure tailoring actions do not break “realism boundaries” remains an important open question. However, three factors reduce the concerns in practice:

1. **The example of adaptive human role players.** As mentioned previously, expert human role players tend to adapt their behavior to the behavior of their students. We have observed human role-players performing some of the adaptive strategies that are outlined herein. A thorough investigation of the pedagogical adaptations of adaptive human role players for training would potentially be useful for establishing practical “rules of thumb” for maintaining a realistic experience (or when other factors dominate). However, the primary conclusion is that intrinsic adaptations are already part of the “toolkit” of the best method known for this kind of training.

2. **Wide variation in human behavior.** In informal conversation, human behavior exhibits a vast range of reasonable responses. Two different people may have very different reactions to some statement; the same person may react very differently to the same statement depending on context and mood. The consequence of this variability is that the “realism boundaries” are very broad and more difficult to breach than may first be apparent. Further, because people tend to view virtual humans as real, minor “quirks” in their reactions might be overlooked by the student, making the boundaries fluid and forgiving.

3. **Variation in adaptive delivery.** A conversation consists of a large number of back-and-forth interactions between the student and the virtual character. One potential way to improve the likelihood that “realism bounds” are not violated is to give the pedagogical adaptation system a rich, fine-grained model of the student’s capability and a range of adaptation strategies. Any single interaction may have more or less pedagogical content, more or less hinting, more or less non-verbal response. This is the approach taken in the dynamic tailoring architecture mentioned previously. If the virtual character has only a coarse model of the student and a fixed adaptation
strategy, then these limitations make it much easier for a student to recognize (and possibly come to depend on) pedagogical adaptations.

4. The unarticulated relationship between fidelity and learning. As discussed, the primary risk associated with tailoring lies in the potential to violate loosely-defined boundaries of realism. If a character misrepresents a target culture, negative training is certainly possible. However, it is unclear whether and to what extent this actually holds. Just as a batter in baseball will sometimes swing two bats before hitting to make one feel lighter, there may be reason to believe an unrealistically difficult character may provide some benefit. Further, commercial video games routinely violate the rules of fidelity to maximize interest and motivation. It may be premature to assume that fidelity represents a firm boundary in the development of virtual humans – much empirical research is needed to address these questions.

Finally, current cognitive and social science research is producing theories and models with such high fidelity that a specific or very narrow range of possible behaviors is mandated in any non-laboratory situation of more than a few seconds. As a consequence, there is no guarantee of realism in intercultural training domains – with or without pedagogical experience manipulation. Until formal validation models of human cultural and interpersonal behaviors become available, the perceived realism of these systems will be largely dependent on the creativity and acumen of their designers and informal feedback from subject-matter experts.

A second open question is whether pedagogical experience manipulation can be as effective explicit help from a tutoring system. As discussed, use of one form of feedback does not rule out use of the other, so the interplay between explicit feedback (such as that from an intelligent tutor) and implicit feedback from role-playing virtual humans represents another important area for future research. A system providing both methods of support would require some unifying pedagogical model to govern and coordinate the activities of both systems.

Advances also remain necessary in the evaluation of intercultural learning and in the validation, appropriateness and determination of accuracy of fundamental models like the DMIS. A meta-analysis of many training programs suggests they are effective at teaching cultural knowledge and generating learner satisfaction, but generally fall short in skill acquisition and attitude change (Mendenhall, et al., 2006). Although the authors criticize the general lack of rigor in the field and recommend a greater empirical focus, no specific suggestions regarding why skills and attitude effects are more difficult to achieve are given. This may be symptomatic of a general lack of realistic practice opportunities in those training programs and bode well for the future of virtual human based environments. Although many of the reviewed studies qualified in terms of rigor, most lacked a strong, experiential component – lectures, assimilators, discussion, and role-play are indicated as the top four types of programs (p. 134). Given that learning is situated and contextual when interacting with virtual humans, they may also play an important role in the future as an assessment tool to gauge culture specific learning as well as movement through developmental stages (Ogan & Lane, in press).

Although many open questions remain regarding the use of virtual humans as role players for the learning of social and intercultural skills, they represent an important class of technology that should continue to remain relevant given the growing importance of such skills. There is a growing global need for intercultural learning and realistic, sustainable practice opportunities. Building a virtual human requires a
great deal of effort – artwork, animation, speech, dialogue modeling, cognitive modeling, emotional modeling, task modeling, and more – and it may not be sufficient to only seek to make them realistic. This chapter provided a brief overview of how virtual humans are currently being used as role players in systems for social and intercultural development, and how they might become specialized for specific learner needs. This represents a possible glimpse of how they may evolve to meet the learning demands of the next several decades.

Acknowledgments

The project or effort described here has been sponsored by the U.S. Army Research, Development, and Engineering Command (RDECOM) and the Air Force Research Laboratory (contract FA8650-08-C-6837). 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. We thank Mark Core, Brian Stensrud, Laura Hamel, Daniel Auerbach, and Dave Gomboc for the many conversations and emails that helped us formulate and refine (and implement some) of these ideas.

References


Table 1: Dimensions of pedagogical experience manipulation

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Engagement/affect</th>
<th>Domain learning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>provide experiences that motivate &amp; encourage</em></td>
<td><em>provide experiences that address specific learner needs</em></td>
</tr>
<tr>
<td>Dynamic tailoring</td>
<td><em>adjust simulation behaviors to motivate &amp; encourage</em></td>
<td><em>adjust simulation behaviors to address specific learner needs</em></td>
</tr>
</tbody>
</table>