Affective Learner-Centered Design Framework for Virtual Human Educational Technologies

Peter Khooshabeh and Jonathan Gratch
Institute for Creative Technologies
University of Southern California
Playa Vista, CA 90094
{khooshabeh, gratch} @ ict.usc.edu

Abstract:
In this paper we describe an affective, learner-centered framework for building learning technologies. Specifically, we outline an ontology of social effects that includes the necessary and sufficient criteria that virtual human technologies need to have in order to design compelling training and learning applications focused on teaching tacit, soft skills for interpersonal communication. Finally, we focus on recent projects from our research that exemplify applications of the ontology of social effects to implementing virtual human learning technologies.

Keywords: Affective computing, interpersonal communication, influence and persuasion, virtual agents

Introduction

Students at every level of the educational system are enrolled in coursework that is mainly intended to teach traditional topical content. Most of these fundamental courses range from mathematics, language arts, history, science, and possibly music. When students achieve their formal educational goals, most of them leave the academic environments with their degrees in hand; ready to be a part of society at large. However, these formal learning experiences do not necessarily teach students the tacit soft skills they need in order to successfully operate in teams and groups. Our work proposes a design framework for building learning technologies to support teaching skills that are important in interpersonal communication.

Learning technologies tend to focus on teaching students cognitive skills and strategies to perform knowledge-based tasks. For example, in the introductory chapter to the Cambridge Handbook of Multimedia Learning, Mayer (2005) points out the role of educational technology for teaching mechanics of how a pump works. The two main factors that are relevant to the design of multimedia environments to support traditional learning, such as the physical dynamics of a pump, are how to balance the presentation of words, sounds, and visuals. The whole field of learning sciences has coupled nicely with educational psychology and devised both theoretical and applied principles for the effective design of these multimedia learning systems. While this is all well and good, there are is principled approach to building effective social skills training, especially for learning how to effectively work in groups in the 21st century. In addition to words, sounds, and visuals, we propose that virtual human interaction participants are more important in the context of teaching interpersonal communication skills such as effective listening and gauging other participants’ state of mind.
In order to support future plans for learning technologies in the 21st century, immersive training environments must include computer-generated virtual humans with whom the learner can interact, and in their most general form, such virtual humans must have the ability to act, react, and counter-react to verbal and non-verbal stimuli, use natural language processing, have appropriate facial expressions and gestures, respond based on the situation and be capable of showing emotion. A major challenge in developing effective virtual human training applications is determining the correct level of fidelity with which they are designed, which refers to the level of realism in the simulation of various components of the virtual human experience. Some components of fidelity are the visual appearance of the virtual human character, the behavioral actions such as gestures and other facial animations, and communicative realism such as natural language expressiveness.

Although virtual environments must be realistic enough to allow learning, they do not have to replicate the physical world in all respects. In fact, the relevancy principle from multimedia design suggests that irrelevant details in learning technologies will introduce extraneous cognitive load to a task. Therefore, virtual human fidelity must be driven by psychological principles of how learning takes place and the necessary factors in multimedia to maximize learning transfer. For some training objectives, low-fidelity characters may be sufficient, whereas for others, even small departures from reality could produce negative training. Unfortunately, the relationship between virtual human characteristics and training objectives is poorly understood and the current research effort is directed at closing this gap.

This paper is directed at gaining insight into the relationship between virtual human fidelity (both visual and behavioral) and learning outcomes with the domain of interpersonal-skills training (see Figure 1). Rather than exploring generic concepts like realism or believability, we adopt a functionalist approach. We argue that realism is important only to the extent that it evokes objective responses and subjective impressions in trainees that are relevant to the intended learning task. In line with this perspective, a central focus of the proposed work is to document the various social effects that virtual humans can have on trainees and the relevance of these effects to the learning of specific skills. We posit that these social effects are an essential element for effective interpersonal skills training.

Contemporary virtual human technology falls short of the intelligence, flexibility, and interactivity of human role-players, and will continue to remain so for many years. Nonetheless, a wide range of deployed applications demonstrate that virtual humans can currently achieve effective training and provide a number of advantages over other forms of training (Beal, Johnson, Dabrowski, & Wu, 2005; Hill et al., 2006; Johnsen, Raij, Stevens, Lind, & Lok, 2007; Kenny, Parsons, Gratch, & Rizzo, 2008; Lane, Schneider, Michael, Albrechtsen, & Meissner, in press; Marsella, Johnson, & LaBore, 2003). Our work focuses on the best practices related to virtual human technologies that have social effects. We also highlight other critical features of virtual human technologies that makes them particularly amenable for effective interactions as learning technologies.
Affective Learner-Centered Design Framework

In line with research on the use of human role-players for interpersonal skill training (e.g., Peterson, 2005; Spence, 2003), we claim that virtual humans provide several potential benefits over traditional classroom instruction:

Social Effects

Numerous studies by various different research groups show that people react socially towards virtual agents and avatars (c.f. e.g. Bailenson, Blascovich, Beall, & Loomis, 2003; Cassell et al., 2002; Bickmore, Gruber, & Picard, 2005; Gratch, Wang, Gerten, Fast, & Duffy, 2007; Krämer, 2005; Reeves & Nass, 1996; Nass, Moon, Morkes, Kim, & Fogg, 1997, see Krämer, 2008, for an overview). We posit that these social effects are an essential element for effective interpersonal skills training. For example, to learn to overcome fear of public speaking in a virtual environment, it is important that trainees feel anxious talking to a virtual audience; or to teach a health-care provider to deliver bad news, it is important that a trainee feels, and learns to overcome, the social pressure to withhold painful truths.

Interactivity

Interpersonal skills are the skills one uses to interact with other people, so it is not surprising that there can be clear benefits to allowing trainees to practice these skills in an interactive way, when compared with other methods such as traditional classroom instruction or instructional videos (Peterson, 2005; Staub, 1971). This is in line with constructivist and guided-experiential learning theories that deem interaction to be central to the learning process and regard knowledge as constructed by the learners through a process of making sense of their interactions with learning material (Clark, Yates, Early, & Moulton, 2010; Merrill, 1991).

Feedback and Guidance

In order to support meaningful social interactions with trainees, virtual humans must sense, understand, and shape aspects of the trainee’s behavior. Within the context of virtual human research, this information is primarily useful to inform the agent’s behavior. However, it can be equally valuable to repurpose this information in support of learning objectives.
Consistency and Dissemination

One challenge with human role-players is ensuring that they consistently behave in a manner that supports the underlying pedagogy. For example, medical role-players require extensive and expensive preparatory training, often using paid actors, yet there are still concerns with ensuring consistency across different role-players. Consistency is especially difficult as some behaviors necessary for training may not be under voluntary control. For example, medical role-players cannot easily simulate the dilated pupils and loss of muscle control associated with traumatic brain injury. Social behaviors, in particular, often occur outside conscious awareness and gain their diagnostic power by virtue of being involuntary (Dijk, Koenig, Ketelaar, & de Jong, 2011). Virtual humans are not limited by these physiological and cognitive constraints and can be programmed to consistently present exact stimuli. Further, one can systematically manipulate a character's race, gender, age and culture as needed for training. Of course, being digital further affords the advantage that virtual humans can be easily copied and distributed across multiple training sites, ensuring uniform training.

Ontology of Social Effects

Although we claim that all of these factors (social effects, interactivity, guided feedback, and consistency and dissemination) play a central enabling role in interpersonal-skill training, this paper will focus primarily on the role of social effects in learning and how these effects might be affected by the level-of-realism provided by a virtual human. This choice of focus is motivated by the observation that social effects are a unique and arguably controversial advantage of virtual humans when compared with other forms of interactive learning technology (Mayer, 2005; van Mulken, André, & Muller, 1998).

A high-level of realism is not entirely necessary to exhibit social effects—e.g., as in the classic Heider and Simmel (1944) where people readily make causal attributions of the social effect of ostracism with a simple animation of geometric objects. Several researchers have documented that social effects can arise when technology incorporates even minimal social cues. For example, Nass and colleagues have shown that a voice on a computer can evoke gender stereotypes (Nass, Moon, & Carney, 1999) and even something as minimal as the label on a monitor can trigger group associations (Moon, 1998). Unfortunately, such demonstrations of the existence of social effects are insufficient to guide the design of effective training and learning systems as the impact of specific social effects on specific learning objectives remains unclear. Virtual humans can vary in realism along a number of dimensions. For example, Figure 2 organizes several systems in terms of their visual realism (i.e., the extent to which they capture the physical appearance of humans) and interactional realism (i.e., the extent to which they faithfully model the verbal, nonverbal, and dynamic nature of human social interactions). It remains unclear how these “dimensions of reality” impact the ability of a virtual human to socially influence human users with the consequence that some effects are difficult to replicate across contexts and settings (Tourangeau, Couper, & Steiger, 2001), and even when the effects are robust, they may not consistently translate into learning gains (Moreno, 2004; van Mulken, et al., 1998).
To address these limitations and to work toward a theoretical framework that can inform the design of effective interpersonal-skills training systems, we:

- Identify a target set of interpersonal skills for which virtual humans hold potential to teach. This list is not intended to be exhaustive, but rather to be representative and to ground the research in a range of concrete and relevant tasks.
- Clarify the potential benefits of virtual human technology when contrasted with traditional classroom instruction.
- Characterize and organize the various dimensions of realism that can potentially impact these effects.
- Empirically investigate alternative frameworks proposed in the literature that can inform the effective design of virtual humans for interpersonal-skills training.

Technical Barriers

In line with the claims above, we argue that addressing the problem of how to design virtual humans to teach interpersonal skills can be profitably decomposed into two related questions: 1) how does the level of visual and interactional realism of a virtual human impact its ability to produce social effects in human trainees; and 2) how do these social effects facilitate social skills training? The former has been an active area of investigation within the media psychology and human-computer interaction communities. Although no clear theoretical consensus has emerged to guide virtual human design, our work aspires to address current limitations in this existing body of research. The latter question is only recently beginning to attract attention and we will largely be breaking new ground, although we plan to leverage the extensive body of research about the role of social cues on other forms of learning, such as concept learning. We are unaware of existing work that explores these two questions in combination.

How do dimensions of realism impact social effects?

There are many theories and approaches that attempt to provide an explanation for the occurrence of social effects in human-computer interaction. For example, the innovation hypothesis states that social reactions towards a computer are a temporal phenomenon due to the novelty of the situation. This novelty effect vanishes once the user becomes accustomed to the interaction with the technology (Epley, Waytz, & Cacioppo, 2007; Kiesler & Sproull, 1997). According to the deficit hypothesis, social effects occur due to deficits on the part of the human, such as a lack of knowledge, inexperience of youth, or psychological or social dysfunctions (c.f., Barley, 1988; Turkle, 1984). The Threshold Model of Social Influence by Blascovich and colleagues (Blascovich, 2002; Blascovich, Loomis, Beall, Swinth, Hoyt, & Bailenson, 2002) suggests that people only respond socially to other people – or, if situated in a virtual reality, to characters they believe are controlled by another person (i.e., an avatar). According to this view, a virtual human, however, would not elicit social responses unless the behavior is sufficiently realistic so that the user cannot distinguish the virtual human from a human-controlled avatar. The Ethopoeia concept by Nass and colleagues (Nass, Moon, Morkes, Kim, & Fogg,
1997, Nass & Moon, 2000), on the other hand, argues that effects depend merely on the presence of social cues, regardless of whether they are believed to be generated by a person or a computer. When social cues are present, the Ethopoeia theory suggests that people will treat computers as social actors.

Unfortunately, there are several challenges in using this prior research to inform the design of virtual humans. Each of these theories has different implications for virtual human design, yet none have definitive empirical support, creating uncertainty on how to proceed. Further, much of this prior research has focused on dimensions of reality that are less relevant to the virtual human research effort. For example, most studies have investigated the impact of visual realism. Behavioral realism, to the extent that it has been modeled at all, has examined low-level behaviors such as gaze or random body movements (Hoyt, Blascovich, & Swinth, 2003), rather than complex behaviors such as speech and nonverbal communication explored in this virtual human project. The result is that studies often manipulate the “mere presence” of social cues (van Mulken, et al., 1998) rather than considering how realistically an agent simulates when such cues are produced (e.g., does it accurately model when a person might smile in a social context?). Finally, most of this research has explored social effects as an end in themselves, and not considered the impact of realism on learning.

How do social cues and social effects facilitate interpersonal-skills learning?

There is growing interest in the use of social cues within educational technology, though their use has largely focused on “cognitive skills” such as the teaching of math, science, or language arts (D'Mello, Picard, & Graesser, 2007; Leelawong & Biswas, 2008; Lester, Towns, Callaway, Voerman, & FitzGerald, 2000) and little is known about their use in interpersonal-skills training. With regard to cognitive skills, educational technologists argue that social cues can enhance cognitive learning by increasing student motivation and engagement. For example, students are often motivated to perform better in the presence of a teacher through social pressures (such as the desire to please or the fear of criticism) and they might be equally motivated by a virtual teacher that evokes these social effects. Several studies support this intuition. Mayer (2005) reports on a series of studies that manipulated social characteristics such as voice, personalization, and visual appearance of characters and demonstrated clear learning benefits in domains ranging from teaching about concepts such as lightning, environmental science, mathematics word problems, and electric motors.

If social cues can enhance the effectiveness of cognitive-skills training, one could expect the effects to be even stronger in the context of social-skills learning. This is because, whereas social factors are argued to play only an indirect role in cognitive learning (i.e., social cues enhance learning motivation which translates
into more effort, which translates into better learning), social factors are intrinsic to the task of interpersonal-skills training. For example, standard techniques for teaching people to overcome social phobias require the student to be placed in fear-inducing situations and these methods would be inapplicable if virtual social contexts didn’t induce fear. Unfortunately, there has been little rigorous research on the effectiveness of interpersonal skills training virtual systems, with the exception of some focused efforts on training developmentally challenged populations such as Asperger’s syndrome (Cobb, Kerr, & Glover, 2001). It remains both a technical and theoretical challenge to understand how to train healthy, normal populations effective social and interpersonal skills in order to perform tasks such as negotiation, networking, and conflict resolution.

Empirical testbeds Using the learner-centered Design framework

Because virtual humans model considerably richer social behavior than has heretofore been explored in prior research on the social effects of computers, we need to develop empirical test beds that allow us to meaningfully explore dimensions of behavioral realism. For example, theories of human social behavior argue that behavior is contingent on the behavior of others (e.g., people mimic each other’s expressions and gestures) and contingent on the social context (e.g., the meaning of a smile depends on the actions that proceeded it). Therefore we need to develop a suite of experimental tasks that allow the systematic manipulation of such contingent factors and enable the elicitation and measurement of social effects and skills. Test beds that we anticipate using in the course of this research include:

- The Rapport Agent: The rapport agent is a test bed for exploring the interpersonal effects of contingent nonverbal behavior. It has been successfully used to study the impact of agent nonverbal behavior on interpersonal rapport (Gratch et al., 2006), speech fluency (Gratch, Wang, Gerten, & Fast, 2007), memory (Wang & Gratch, 2009), gender stereotypes (Kulms, Krämer, & Gratch, 2011), cultural stereotypes (Dehghani, Khooshabeh, Huang, Gratch, & Hovanesian, 2011) and intimate self-disclosure (Kang, Gratch, Wang, & Watt, 2009).

- Lunar Survival Task: This test bed has been developed to explore the importance of natural-language fidelity on social effects. It allows subjects to interact with a virtual human through natural language and negotiate with the agent over which items to bring along following a crash landing on the moon. The test bed has been used to study the factors that impact social persuasion (Khooshabeh, McCall, Gandhe, Gratch, & Blascovich, 2011). This task has some ecologically validity because it puts participants in a simulated high-stress situation, not unlike those experienced by high-level leaders and decision makers.

- Social Dilemma Agents: This test bed has been developed in collaboration with USC’s Marshall School of Business to explore the importance of context in determining the social effect of emotional expressions. It allows subjects to interact with a virtual human through a menu-driven interface on a number of standard laboratory tasks used to study social decision making including prisoner’s dilemma and multi-issue bargaining. It has been used to study factors of virtual human behavior that impact users
decisions on these tasks (de Melo, Carnevale, & Gratch, 2011; de Melo, Gratch, & Carnevale, 2011) as well as their physiology (heart-rate, blood-pressure, and vascular contractility) and eye movement behavior (Khooshabeh et al, in prep).

- Learning transfer: A crucial test of whether training technologies meet their objective is when students are tested outside of instructional settings. This is the concept of transfer and it shows how learning in a specific context generalizes to novel contexts. There are different types of transfer, each of which has a unique way of measuring it. In the problem-solving literature, a transfer test measures learning beyond simple retention of the material presented during instruction (Mayer, 1999). Other aspects of transfer are near and far transfer. Near transfer is a situation where learners perform better at tests that are relevant and closely related to the instructional setting. An example of near transfer is when students take an exam and the questions are different from the homework practices but related. Far transfer is when learners perform better in contexts that are markedly different from those of the instructional setting. An example of far transfer is evident from the work of Green, Pouget, and Bavelier (2010) where they showed that playing an action video game for 50 hours improved performance in a tone location discrimination task, which is very different from the video game practice task. As part of developing the above test beds, we will develop transfer tasks that can assess the generality of learned knowledge.

Applications of an Affective Learner-centered Design Framework

Using an affective, learner-centered design framework for building learning technologies can make it possible to meet many educational grand challenges. One of these is to increase the number of students in science, technology, engineering, and mathematics (STEM) disciplines. Learner-centered design principles can address fundamental questions related to the social, technical, and cognitive factors that lead to successful learning environments for the STEM disciplines. We apply social psychological topics to understand and improve 21st century skills such as interpersonal communication as well as the role of emotions in decision making. This work will improve our understanding of the interplay between communication and learning.

References


