

# Chapter XX

## Design and Evaluation of Embodied Conversational Agents for Educational and Advisory Software

Elisabeth André  
*Universität Augsburg, Germany*

### ABSTRACT

*Embodied conversational agents may take on a diversity of roles in learning and advisory scenarios including virtual teachers, advisors, learning companions, and autonomous actors in educational role play. They promote learner motivation, engagement, and self-confidence, and may help prevent and overcome negative affective states of learners, such as frustration and fear of failure. The chapter will provide guidelines and approved methods for the development of animated pedagogical agents including the extraction of multimodal tutorial strategies from human-human teaching dialogues as well as the simulation and evaluation of such strategies in computer-mediated learning environments.*

### INTRODUCTION

The objective to develop more human-centered, personalized and at the same time, more engaging speech-based interactive systems immediately leads to the metaphor of an embodied conversational agent (ECA) that employs gestures, mimics, and speech to communicate with the human user. During the last decade research groups as well

as a number of commercial software developers have started to deploy embodied conversational characters in the user interface, especially in those application areas where a close emulation of multimodal human-human communication is needed. In this chapter, the potential of embodied conversational agents for educational software is investigated. In addition, advisory software is discussed, but restricted to applications which

aim at achieving a change in people's behavior as, for example, health advisors.

Embodied conversational agents bear the advantage that they enable rich multimodal interactions with learners by employing gestures, mimics, and speech to communicate with the human user. The most obvious role of an embodied conversational agent in educational software is that of a virtual teacher. There is empirical evidence that pedagogical agents lead to an improved perception of the learning task and help to engage learners (see Mulken, André, & Müller, 1998). They promote learner motivation, engagement, and self-confidence, and may help prevent and overcome negative affective states of learners, such as frustration and fear of failure. An interesting variant of a conversational agent in a learning scenario is that of a learning companion. Learning companions ensure the availability of a collaborator and may increase the students' engagement in a task (see Craig, et al., 1999). They provide an interesting new training tool since it would be impossible to create a real classroom setting for individual students that fosters their learning progress best. Educational role-play promotes learning by enabling a learner to actively participate in a drama-based environment. It provides the student with a safe environment for experimental learning and can make learning a more engaging and enjoyable experience.

This chapter will provide guidelines and approved methods for the development of animated pedagogical agents including the extraction of multimodal tutorial strategies from human-human teaching dialogues as well as the simulation and evaluation of such strategies in computer-mediated learning environments. The second section will present prominent educational environments that make use of a diversity of embodied conversational agents including pedagogical agents, virtual helpers, and actors in educational drama. In the third section, empirical studies that investigate the potential benefits of embodied conversational agents for educational or advisory software are reported

on. After that, the design and implementation of pedagogical agents are addressed. In the fourth section, how to make use of multimodal corpora to guide the design of tutoring agents is explained. The fifth section describes how to realize different educational settings ranging from face-to-face communication with a single pedagogical agent to educational role play with several autonomous actors. In the sixth section, technology that is used to enable the user to engage in multimodal dialogue with the tutoring system is reported on. First attempts towards the development of perceptive agents which are able to perceive and interpret communicative signals from the learner, for example, to monitor his or her level of interest, are also reported on. The seventh section provides an overview of techniques to recognize the user's emotional state and to respond to it accordingly. How to use affective non-verbal signals, such as facial expressions and mimics, to provide pedagogically effective and conversationally appropriate responses to contributions received from the students, are discussed. Embodied conversational agents in learning environments need a great deal of social competence that manifests itself in a number of different abilities. In the eighth section, various attempts to enhance pedagogical agents by social competence including the use of strategies of politeness and relational strategies are reported on. The ninth section sketches some future research trends.

### **Roles of Embodied Conversational Agents in Educational and Advisory Software**

There is growing number of learning environments that make use of embodied conversational agents. In the following, some prominent examples are briefly reviewed to give the reader an idea of the many roles in which embodied conversational agents may appear.

## **Face-to-Face Communication with Pedagogical Agents and Virtual Advisors**

An early example of educational software employing an embodied conversational agent includes the pedagogical agent Cosmo which inhabits together with the student, a botanical environment (see Lester, Towns, Callaway, Voerman, & FitzGerald, 2000). Cosmo advises a student on the task of designing plants capable of surviving in certain environments. The pedagogical agent Steve co-habits with the student a virtual environment and instructs him or her to operate technical devices within this environment (see Rickel & Johnson, 1999). Steve monitors the student's actions, provides advice if necessary, or demonstrates actions to the student. AutoTutor is an embodied conversational agent that teaches the student elementary computer literacy (see Rajan, Craig, Gholson, Person, & Graesser, 2001). It converses with the learner via synchronizing synthesized speech, intonation, facial expressions, and gestures. The pedagogical agent Baldi teaches language skills, such as speech articulation, grammar, and vocabulary. It appears as an animated 3D head and makes use of realistic visual speech (see Massaro, 2004). The use of both auditory and visual information makes it, in particular, suitable as a teacher for children with hearing loss and adults learning a new language.

Advisory software as discussed in this chapter typically includes a virtual advisor that engages in a face-to-face dialogue with a human user. Bickmore and Picard (2005) developed an exercise advisor agent that employs relational strategies people use in face-to-face communication to establish a long-term relationship with a human user. A similar application has been developed within the MagiCster project with the Greta agent, which provides advice about eating disorders. The Greta Advisor is a "realistic" 3D embodied agent that is animated in real-time and communicates information through the combination of verbal

and nonverbal signals (see deRosis, Pelachaud, Poggi, Carofiglio, & Carolis, 2003).

## **Educational Role Play with Embodied Conversational Agents**

Another line of research is educational role-play using embodied conversational agents. Among other things, educational role play has been exploited for language learning, the acquisition of social skills, as well as therapeutic purposes.

One of the first applications making use of educational role play is Carmen's Bright IDEAS (see Marsella & Gratch, 2000). It relies on drama-based interventions in order to help mothers of young cancer patients to develop problem solving skills. A more recent example includes FearNot! which was developed within the VICTEC project by Paiva et al. (2004) for education against bullying with emotion-driven virtual characters (see Figure 1). The project investigates how social learning may be enhanced through interactive role play with virtual characters that establish empathetic relationships with the learners. It creates interactive stories in a virtual school with embodied conversational agents in the role of bullies, helpers, victims, and so forth. The children run through various bullying episodes, interact with the virtual characters after each episode, and provide advice to them. The benefit of educational role plays in these scenarios lies in the fact that they promote reflective thinking. By taking on the perspective of the virtual characters, the children are supposed to learn how to cope with bullying in class. For autumn 2007, a longitudinal study with 900 children from Germany and the UK is planned to investigate in how far the software helps to combat bullying and supports effective conflict resolution.

Prendinger and Ishizuka (2001) make use of educational role play to improve the English conversation skills of native speakers of Japanese. In a conversation training session, the user interacts with one or several characters, for example, as

Figure 1. The FearNot! system (Paiva et al., 2004)

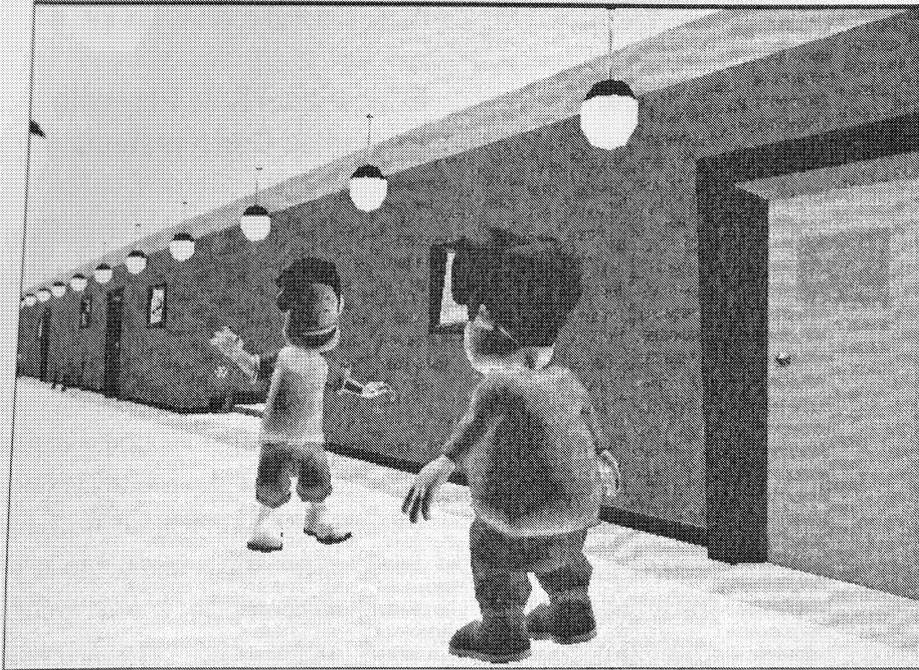
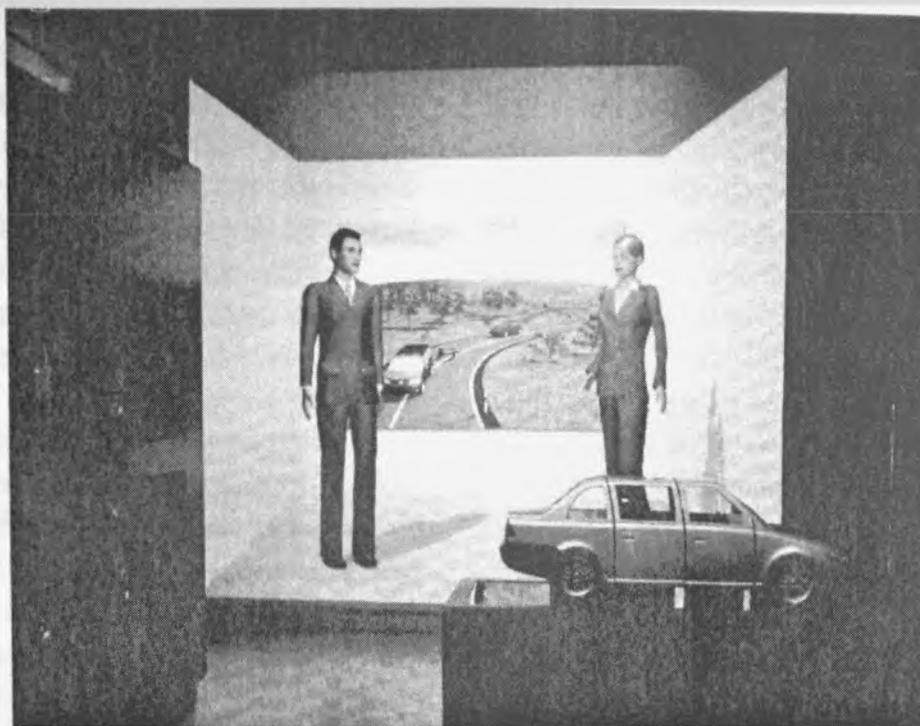


Figure 2. The tactical language system (Johnson et al., 2004)



Used with permission.

Figure 3. The Autostadt Park system (Kipp et al., 2006)



*Used with permission.*

a customer in a virtual interactive coffee shop, by uttering one of a set of predefined sentences displayed on a screen. The Tactical Language Training System (see Johnson et al., 2004a) goes a step further and aims not only at helping people acquire a foreign language, but also cultural skills (see Figure 2). Unlike earlier language learning systems, the student may engage in a dialogue with multiple characters employing speech as well as gestures which are chosen from a menu and then exhibited by the player's avatar. Courses are offered in three different languages: Iraqi Arabic, Pashto, and Sahel French. The system has been used by thousands of U.S. military officers over a longer period of time, many of which consider the system better than instructor-based classes.

Kipp, Ndiaye, and Gebhard (2006) rely on educational role play to develop a learning scenario for the Autostadt Entertainment Park (see Figure 3). Two characters observe a user while he is moving around physical building blocks to construct a miniaturized car. The user is not able to engage in a conversation with the characters,

but the characters make comments on the user's performance and give advice. An interesting feature of the system is the combination of a tangible interface with virtual helpers.

The research by Cassell et al. (2000) focuses on the development of virtual playmates that help children to develop literacy skills through interactive storytelling. Their system allows children to play with natural figurines inhabiting a physical castle in collaboration with a virtual character. A later version of the system described in Tataro and Cassell (in press) is used as an educational tool for children with autism spectrum disorder. The authors hope that the system will contribute to a better understanding of autism and help the children to overcome social deficits associated with autism.

## **EMPIRICAL EVALUATION OF PEDAGOGICAL AGENTS**

In the following, a number of empirical studies that have been conducted to investigate the potential pedagogical benefits of embodied conversational agents are discussed. Some general design guidelines that may be derived from such studies have also been come up with.

### **Potential Benefits of Pedagogical Agents**

Most empirical studies emphasize the social role of pedagogical agents. Studies by Lester et al. (1997) have shown that the pure presence of an embodied agent can have a positive effect on the students' perception of the learning experience. The effect was confirmed by Mulken, et al. (1998) who investigate the attitude of students' towards a virtual agent. Subjects who had seen presentations guided by a virtual agent indicated on a questionnaire that they found the presentations themselves and the corresponding tests less difficult than subjects who had seen presentations without the virtual agent. In addition, subjects found these presentations significantly more entertaining.

Studies investigating the effect of animated agents on the user's learning performance did not provide consistent results. Mulken, André, and Müller (1998) investigated the impact of a virtual agent on understanding and recall, but did not detect any effect. Similar results were obtained by Kraemer (2005) who investigated whether the use of an embodied conversational agent may improve the effectiveness of a TV/VCR interface. In line with Mulken et al., she did not observe a positive effect on recall in comparison to pure text and audio. Furthermore, the subjects did not rate the presentations given by the agent more helpful than those conveyed by pure text and audio. Contrariwise, there are studies revealing a positive effect of an embodied agent on learning. For

example, Beun and Witteman (2003) conducted a memory test which supported their hypothesis that the presence of embodied conversational agent may contribute to better recall.

Summing up, it may be said that there is empirical evidence that pedagogical agents lead to an increased sense of ease and comfort and motivate a learner to engage in a learning task for a longer time. In this way, they may indirectly also contribute to a better learning performance. Such an effect has, however, still to be attested by longitudinal studies. Studies that investigate whether or not an embodied conversational agent directly contributes to learning after shorter interactions have led to inconsistent results so far. A lot of research is still required to investigate in which cases an embodied agent improves learning. In any case, creators of agents should make sure that the agent is more than a decorative feature, but has a functional role in an educational setting. Otherwise, there is the danger that that agent just produces an additional cognitive load for the learners and distracts them from the actual contents.

### **Design Guidelines from Empirical Studies**

There are a number of factors that influence the user's attitude towards the agent and the effectiveness, such as the character's degree of realism, its shape, its ethnicity, its gender, and so forth, and it is difficult to generalize empirical results to other user groups and interaction scenarios. Nevertheless, some general findings seem to crystallize out.

Gulz and Haake (2006) emphasize the importance of a character's audio-visual appearance in educational settings. Subjects in their experiment preferred stylistic agents over realistic agents. Based on these findings, the authors recommend that creators of pedagogical agents should not necessarily strive for great realism. Baylor and Kim (2004) found out that more realistic agent

images led to a greater transfer of learning. Furthermore, they observed that agents that were perceived as less competent resulted into better self-efficacy. These studies show that a number of factors need to be considered in order to find a good match between a character's role and its appearance. Depending on whether the emphasis is on motivational goals or learning goals, a different design might be appropriate.

Various studies investigate whether users prefer characters that are similar to themselves and whether such characters are more effective, for example, as virtual trainers or consultants. Nass and Gong (2000) have shown that computer agents representing a user's ethnic group are perceived as socially more attractive and trustworthy. Similar results were obtained by Baylor and Kim (2004). Van Vugt, Konijn, Hoorn, and Veldhuis (2006) observed a preference of users for non-ideal fatter characters as e-health advisors. Especially, users similar in body shape trusted these agents more than slim characters.

Studies conducted by Nass and co-workers reveal the importance of consistencies in an agent's multimodal behavior. Lee and Nass (2003) observed that a user's feeling of social presence is positively affected if the personality that an utterance conveys is reflected by the employed synthetic voice as well. Nass and Gong (2000) claim that maximizing the quality of each modality does not necessarily improve human-computer interaction. Even though recorded human voices are more natural than the output of a text-to-speech synthesizer, an interface may become more appealing when a synthetic face is accompanied by a synthetic voice.

### **Extraction of Multimodal Tutorial Strategies Using a Corpus-Based Approach**

A number of approaches to modeling the behaviors of embodied conversational agents are based on a direct simulation of human behav-

iors. Consequently, it comes as no surprise that the use of data-driven approaches which allow the validation of design choices empirically has become increasingly popular. To get insight into human-human conversation, researchers rely on a large variety of resources including recordings of users in "natural" or staged situations, Wizard of Oz studies, or motion capturing data.

The use of corpora provides a promising approach to the modeling of virtual teachers as well. The basic idea is to record human teachers, to analyze their behaviors, and to replicate these behaviors using an embodied conversational agent. For instance, the tutorial tactics as well as the conversational cues by AutoTutor are based on recordings of naturalistic tutorial sessions between students and unaccomplished tutors (see Person, Graesser, Kreuz, Pomeroy, and The Tutoring Research Group, 2001). Johnson et al. (2004b) videotaped interactions between learners and an expert human tutor while the learners were working with a Web-based learning environment for factory modeling and simulation to extract politeness tactics. Based on an analysis of these tutorial sessions, Johnson et al. identified eight categories of tutorial advice, such as direct commands or socratic hints, depending on the degree of politeness being expressed. For instance, indirect suggestions, such as "They are asking you to go back and maybe change it" were perceived as more polite than direct commands, such as "Click the ENTER button." The eight categories have been used as the basis for the implementation of politeness behaviors in a tutoring system. An evaluation of the system revealed that the incorporation of politeness behaviors may have a positive effect on learning (see the eighth section).

Various annotation schemes have been designed to extract relevant information for multimodal behaviors, such as facial expressions, gestures, postures, and gaze, from corpora. In addition, there has been increasing interest in the design of annotation schemes to capture emotional behaviors in human-human conversation. Rec-

ommendations for the definition of multimodal coding schemes have been set up within the European ISLE project (see Dybkjær and Bernsen, 2004). Progress in the field has been boosted by the availability of new tools that facilitate the acquisition and annotation of corpora.

For the analysis of educational dialogue, the discount markup scheme has been proven useful (see Pilkington, 1999). Discount describes the structure of educational dialogs both in terms of exchanges between the single parties (e.g., initiate, respond, reinitiate) and in terms of rhetorical predicates (e.g., consequence, condition, contrast). Among other things, the discount scheme has been employed to extract politeness tactics in tutorial sessions.

A variety of coding schemes that capture relevant features of multimodal dialogue has been designed. Schemes for annotating gestures usually follow McNeill's (1992) classification of gestures. He defined several communicative gesture types: iconic (refers to some physical/spatial/temporal properties of the speech referents), metaphoric (refers to abstract property of the speech referents), deictic (indicates a concrete or abstract point in the world), beat (follows the rhythmic structure of the speech), and emblem (has a conventionalized meaning). Rehm and André (in press) investigated which kinds of gestures people choose to accompany their speech in face threatening situations. Their coding scheme consists of three main layers: the transliteration, that is, the words spoken, the hand gestures of the speaker visible in the video, and the politeness tactics employed by the speaker.

Within the Humaine Network on Emotion Research, an annotation scheme has been developed which focuses on emotion-specific behaviors in speech, head and torso movements, facial expressions, gaze, and hand gestures. Such a scheme may be useful to extract emotional responses by teachers as well as by students in a tutorial session. Expert teachers try to recognize a student's emotional state and to respond to it in a way

that promotes learning. Recordings of tutorial sessions labeled with the emotional states of the learners and the teacher enable correlations between a student's emotional state and a teacher's subsequent action to be identified. Furthermore, they may help to find out which emotional state teachers deliberately convey to positively influence the student's learning process.

How can corpus data be used to control the behavior of an agent? Basically, a direct and an indirect use of corpus data can be distinguished between. In the first case, animations are generated that directly correspond to behaviors found in the data. In the second case, abstract regularities that govern the generation process are derived from the corpus. An example of the direct use of corpus data is the work by Stone et al. (2004). They generate multimodal behaviors by recombining the speech and motion samples from recorded data. The technique produces more naturalistic behaviors than techniques that synthesize behaviors from scratch. However, to allow for variations in the performance of an agent, data have to be collected for different kinds of situation, personality, emotion, and so forth. These problems may be compared to problems occurring when using a unit selection approach to synthesize speech. An example of the indirect use of corpus data is the work by Foster and Oberlander (2006) who conducted experiments with a majority-choice and a weighted-choice model for the generation of facial displays. In the first case, the facial display that occurs the largest number of times in the corpus is chosen for the agent. In the latter case, a random choice is made where the choice is weighted according to the relative frequency of facial displays in the corpus. An empirical study revealed that humans seem to have a preference for the second option since it reproduces the variations in the corpus in a better way.

Overall, the observation of human-human conversational behavior is a valuable source of information for the design of pedagogical

agents—especially in cases where knowledge on specific teaching guidelines is missing. Despite new annotation tools, the collection and annotation of corpora is still cumbersome and time-consuming. In practice, it is often not possible to rely on existing corpora since they have been recorded in a specific context which might mismatch the context in which a learning application has eventually been used. A great challenge for the future is, therefore, the automated adaptation of context-specific behavior patterns to a new situation.

### **Realization of Tutorial Tactics in Pedagogical Agents**

Embodied conversational agents may take on a diversity of roles in learning scenarios. They may serve as teachers, as mentors, as learning companions, or as actors in educational role plays. Depending on the agents' roles in a learning scenario, different conversational templates and styles are required. In this chapter, various mechanisms to code tutorial strategies so that they can be used to drive the behavior of animated agents are discussed.

A straightforward approach to present teaching content to the student is to rely on pre-authored presentation clips. For instance, Lester et al. (2000) combined a coherence-based behavior sequencing engine to control the behavior of Herman the Bug, the pedagogical agent of *Design a Plant*. This engine dynamically selects and assembles behaviors from a behavior space consisting of animated segments and audio clips. The material has been manually designed by a multidisciplinary team of graphic artists, animators, musicians, and voice specialists. On the one hand, this allows the authoring of high quality presentations as the human author has much control over the material to be presented. On the other hand, enormous effort by the human author is required to produce the basic repertoire of a course. A particular problem with manually authored clips is that the author has to anticipate scripts for all possible situations

and tasks, and that the clips must allow for sufficient variations in order to avoid characters that behave in a monotonous and too predictable way. Creating clips manually is, however, not feasible for many applications since it would require to anticipate the needs of all potential learners and preparing presentations for them.

To automatically determine the behavior of pedagogical agents, planning approaches have been proven useful. André, Rist, van Mulken, Klesen, and Baldes (2000) define multimodal communication tactics as operators of a planning system. A similar approach has been used by Rickel and Johnson (1999) to control the behavior of the pedagogical agent, Steve. Their plan operators include not only speech acts, but also references to tasks to be carried out in the physical world.

Most tutoring systems that use embodied conversational agents rely on settings in which the agent addresses the user directly as if it were a face-to-face conversation between human beings. For example, an agent may serve as the user's personal consultant or tutor. However, other situations exist in which the emulation of direct agent-to-user communication is not necessarily the most effective way to present information. Empirical evidence suggests that, at least in some situations, indirect interaction can have a positive effect on the user's performance. For example, Craig et al. (1999) found that, in tutoring sessions, users who overheard dialogues between virtual tutors and tutees, subsequently asked significantly more questions and also memorized the information significantly better.

With regard to presentation design, a team of presenters enriches the repertoire of possible communication strategies. For example, they allow the conveyance of certain relationships among information units in a more canonical way. Among other things, this benefits decision support systems where the learner has to be informed about different points of view, pairs of

arguments and counterarguments, or alternative conclusions and suggestions. For solving such presentation tasks, it seems natural to structure presentations according to argumentative and rhetorical strategies common in real dialogues with two or more conversational partners. For instance, a debate between two characters representing contrary opinions is an effective means of informing an audience of the pros and cons of an issue. In educational role play, characters may take on different roles, for example, that of a customer and a seller in a language learning scenario (see Prendergast & Ishizuka, 2001).

When realizing learning scenarios with multiple characters, the question arises of whether to follow a plot-centered approach where a centralized component assigns actions of the single characters or whether to realize the agents as autonomous actors. A plot-centered approach seems useful for educational presentations that are not interrupted continuously by interaction of the user. It may be realized by a centralized planner which decomposes a complex presentation goal into elementary dialogue and presentation acts that are allocated to the individual agents. In educational role play, agents are usually realized as autonomous agents. In this case, each agent has individual goals that it tries to achieve. Actions to be executed by the agents are no longer determined by a centralized planning component, but instead planned on the fly by each agent individually.

Planning approaches also bear the benefit that they may be combined with an affective appraisal system. First, emotions can arise in response to a deliberative planning process (when relevant risks are noticed, progress assessed, and success detected). For example, several approaches derive an emotion's intensity from the importance of a goal and its probability of achievement (see Marsella & Gratch, 2000). Second, emotions can influence decision-making by allocating cognitive resources to specific goals or threats. Plan-based approaches support the implementation of decision and action selection mechanisms that are

guided by an agent's emotional state. Examples of educational role play applications that integrate artificial intelligence planning with an affective appraisal system include Carmen's Bright Idea's (see Marsella & Gratch, 2000) and FearNot! (see Aylett, Dias, & Paiva, 2006).

## **COMMUNICATIVE SKILLS OF EMBODIED CONVERSATIONAL AGENTS IN LEARNING ENVIRONMENTS**

Teaching requires versatile communicative competence. For instance, a teacher needs to explain concepts, ask questions, demonstrate methods, correct mistakes, encourage students, challenge them, and so forth.

In most cases, dialogue systems employed for educational purposes are confined to the analysis of written text input and make use of shallow parsing. An example of a virtual tutor that allows for relatively unconstrained typed natural language input from the user includes AutoTutor (see Graesser, Person, Harter, & The Tutoring Research Group, 2001). After posing a question, AutoTutor initiates a dialogue of 10 to 30 turns with the student, whereby the learner types in his responses. The user is less constrained when providing input than in most tutoring systems. Nevertheless, the authors of the system decided not to analyze the input in depth. Instead the quality of the student's answers is assessed by relying on a knowledge-based statistical pattern matcher that compares the student's actual responses with good and bad answers stored in a library along with the questions the tutor may ask. The authors argue that the analysis of detailed student responses provides more reliable information on the students' depth of knowledge.

Despite attempts to equip synthetic tutors with manifold communication skills and multiple communication channels, most approaches are still characterized by a strong asymmetry in

communication channels. AutoTutor coordinates speech with facial expressions and body gestures and is able to display typed text, but the student is confined to typed input. In the Tactical Language system by Johnson et al. (2004a), students may communicate via speech using a microphone and augment their speech with gestures they consider as culturally appropriate. However, the gestures to go along with the speech have to be selected with a mouse wheel instead of being directly performed by the learner.

First approaches are being made to analyze non-verbal communicative cues from the user with the aim to improve the robustness of multimodal analysis, to recognize the user's level of attention, and to regulate the flow of the interaction. For example, the kiosk Mack agent uses gaze as a deictic device as well as a feedback and turn-taking mechanism (see Nakano, Reinstein, Stocky, & Cassell, 2003). Based on an analysis of human-human conversation, Sidner, Lee, Kidd, Lesh, and Rich (2005) implemented a conversational robot that is able to track the face of the conversational partner and adjusts its gaze towards him or her. Empirical studies by Sidner et al. and Nakano et al. indicate that gaze is an excellent predictor of conversational attention in multiparty conversations. While the listener employs gaze to indicate that s/he is paying attention to the speaker, the speaker monitors the listener's gaze to find out whether s/he is still interested in continuing the conversation. Robust technology to track the user's non-verbal cues offers great promise to educational learning environments since they provide additional information on the user's level of attention and motivational state. In intelligent learning environments, eye tracking technology is usually used to assess a learner's attentional state (see Qu & Johnson, 2005) or a learner's meta-cognitive skills, such as the ability to learn from free exploration and the ability to self-explain instructional material (see Merten & Conati, 2006).

## **Enhancing Pedagogical Agents by Emotional Behaviors**

There is still an ongoing debate whether it is necessary to equip computers with emotional sensitivity. When implementing pedagogical agents, the representation of affective state together with a simulation of emotion triggering seems, however, indispensable. For example, Aist, Kort, Reilly, Mostow, and Picard (2002) showed that human-provided emotional scaffolding to an automated tutoring system resulted in increased student persistence, that is, students were willing to spend more time on a task, which should lead to improved learning.

The availability of robust methods for emotion recognition is an important step in the development of pedagogical agents that are sensitive to a learner's emotional state. Recent research has concentrated on a large variety of verbal and non-verbal communicative cues that may provide information on a learner's affect state. Cues that have been investigated include postures and facial expressions (see Kapoor & Picard, 2005), acoustic and prosodic features of speech (see Ai et al., 2006), as well as physiological signals (see Bosma & André, 2004). Others focused on features of the interaction history. D'Mello and Graesser (2005) mine log files of the interaction history to obtain information about the student's affective state in AutoTutor. McQuiggan and Lester (2006) make use of a combination of log files of the interaction history and bio sensors to assess the learner's confidence to perform well in a learning situation. Ai et al. (2006) consider features extracted from the dialogue between the tutor and the student, such as the prosody of speech, as well as features relating to user and system performance, to improve the emotion recognition process in the ITSpoke tutoring system. Conati (2002) presents a probabilistic framework to derive a learner's motivational state both from their bodily reactions, for example, biometric

signals, as well as the state of the interaction, for example, successful completion of a task.

The question arises of how a teacher should respond to a learner's emotional state. As a first promising step towards the development of a pedagogical agent with emotional sensitivity, Kort, Reilly, and Picard (2001) present a model of a learner's cognitive-emotive state that explicitly relates learning phases with emotional states. The aim of their research is to build a computerized learning companion that makes use of a variety of sensors to track the learner's cognitive-emotive state during learning and intervenes if necessary.

Other researchers focus on the question of which emotions a virtual tutor should express. Obviously, virtual tutors should not always reveal their real emotions. Instead they should follow a communication-driven approach to the expression of emotions. That is, emotions are intentionally expressed with the goal to increase the learning effect. An early example of an affective pedagogical agent is the Cosmo System (see Lester et al., 2000) where the agent's pedagogical goals drive the selection and sequencing of emotive behaviors. For instance, a congratulatory act triggers a motivational goal to express admiration that is conveyed with applause. To convey appropriate emotive behaviors, agents such as Cosmo need to appraise events not only from their own perspective, but also from the perspective of others.

The combination of educational role play and affective computing has also led to promising results in therapeutic applications or training applications teaching social skills. Examples include FearNot! (see Paiva et al., 2004), an interactive system for education against bullying, Carmen's Bright IDEAS (see Marsella & Gratch, 2000), a drama-based learning environment aiming at helping mothers of young cancer patients to develop problem solving skills, or the Mission Rehearsal Exercise (MRE) which focuses on the creation of emotionally charged training scenarios (see Rickel, Marsella, Gratch, Hill, Traum, &

Swartout, 2002). Many of these applications are based on the observation that it is easier to empathize with how another person might feel if one has experienced a similar situation even if it is just a simulation.

As discussed, there is a strong asymmetry in communication channels which also affects an agent's emotional sensitivity. While the behaviors of some agents are highly expressive, they do not have sophisticated perceptive skills to perceive the user's emotional state. In many cases, the agents' emotive state just depends on the user's performance during a learning task, but not on his or her actual emotional state. For example, it may happen that an agent expresses empathy because the student continuously fails to provide a correct answer even though the student does not care about his or her learning progress.

### Enhancing Pedagogical Agents by Social Competence

There is no doubt that the social and psychological skills of a human teacher have an important impact on learning. It seems therefore inevitable to equip embodied conversational agents with such skills as well.

Johnson et al. (2004b) investigated the potential benefits of politeness in a tutoring system. When humans interact with each other, they risk continuously threatening the face of their conversational partners, for example, by showing disapproval or by putting the other person under pressure. To avoid this, humans usually rely on various face threat mitigation strategies. Strategies identified by Brown and Levinson (1987) include negative politeness (e.g., showing approval for the addressee), positive politeness (e.g., emphasizing the addressee's freedom), and off-record statements (e.g., vague indications that an action is needed). For instance, instead of formulating a direct request "Solve the equation," a teacher might rely on a mixture of positive and negative politeness and say "Why don't we solve the equation together?"

Examining the interactions between a real tutor and his students, Johnson et al. came up with a set of templates each of which is annotated according to the amount of redress that tactic gives to the learner's face. The templates have been employed in a tutorial tactic planner that selects different tutorial strategies automatically depending on the type of expected face threat. Wang, Johnson, Mayer, Rizzo, and Shaw (2005) investigated the politeness effect in this system by confronting subjects with polite tutorial feedback that mitigates face threats and direct feedback that ignored face threats. The polite version yielded better learning performances and the effect was even amplified for students who expressed a preference for indirect feedback. Experiments by Prendinger, Mayer, Mori, and Ishizuka (2003) revealed that a polite and empathetic computer agent can contribute to a more positive perception of the interaction. For example, people feel less stressed if a computer agent apologizes for delays.

While most work concentrates on verbal aspects of politeness behavior, Rehm and André (in press) conducted a corpus study with human speakers to shed light on the question of how face threats are mitigated by non-verbal means. Their study revealed that gestures are used to strengthen the effect of verbal acts of politeness. In particular, vagueness as a means of politeness is not only reflected by verbal utterances, but also by gestures. Iconic and deictic gestures were predominantly used in more direct criticism while metaphoric gestures frequently occurred in off-record strategies. The results of the corpus analysis were employed to inform the gesture selection mechanism of an embodied conversational agent.

Bickmore and Cassell (2001) investigated the use of small talk in a health advice system. As a mentor, their agent should demonstrate competence, but at the same time build up a social relationship with the learner. Dynamically changing socio-emotional relationships between a human and an agent are represented by a relational model that includes three dimensions: solidarity,

familiarity, and affect. In order to interleave the realization of task-based and social goals, they developed a computational model of mixed task and social dialogue by adapting an activation network-based approach to discourse planning. The objective of the system is to find conversational moves that pursue task goals as efficiently as possible, but at the same time minimize the expected face threat to the user and maximize trust (as a function of solidarity, familiarity, and affect). Depending on the expected face threat and the state of the relational model, the agent decides whether or not to engage in small talk and what kind of small talk to choose. The work is grounded in a theory of social talk and, in addition, informed by a number of pilot experiments. Unlike most existing work, Bickmore and Cassell focus on the establishment of long-term relationships between a user and a virtual agent. A long-term study comparing a social agent with a pure task-oriented agent revealed that the social agent led to a more positive user attitude towards the agent and also increased the user's wish to continue working with the agent after the termination of the study.

Against this background, it seems worthwhile to equip a pedagogical agent with social skills. Overall, empirical evidence indicates that agents with social skills improve the users' perception of the interaction experience compared to agents not having such skills. Nevertheless, a lot more studies are necessary to define concrete guidelines for an agent's social behavior. One difficulty lies in the fact that the acceptance of a social agent depends, to a large extent, on the user's personality. For instance, Bickmore and Cassell (2005) observed that social talk may help to build up trust between a user and a virtual real estate agent. A positive effect could, however, not be observed across all user groups (introverts vs. extroverts) and interaction styles (face-to-face vs. phone conversations). For instance, introverts perceived the agent as significantly more trustworthy in a pure task-based face-to-face conversation than

in conversations over the phone or conversations including social talk while extroverts trusted the agent the least in this condition.

## **FUTURE DIRECTIONS**

A lot of challenges remain to be solved in order to create really interactive virtual humans for learning and advisory software that are able to engage learners over a longer period of time. In the following, some future research trends are sketched.

### **Novel Interfaces with Pedagogical Agents**

Despite significant progress in the area of multimodal dialogue, interaction with pedagogical agents is still strongly limited. On the one hand, agents need to be equipped with more sophisticated perceptive skills. On the other hand, tutorial tactics should exploit the full bandwidth of multimodality. This also includes the conduction of empirical studies that investigate the effect of an agent's gestures and mimics on the user's perception of the learning task and his or her performance.

A promising direction is the development of educational software that goes beyond interactions via keyboard and mouse and includes the use of tangible objects learners may explore and manipulate. In the second section, first examples of interfaces that made use of pedagogical agents in combination with a tangible interface were presented. The benefit of tangibles interfaces lies in the fact that they engage learners physically in a task. One could even go a step further and have the agents leave the desktop to enter the learner's physical world. First attempts are being made to populate real environments with embodied conversational agents resulting in so-called mixed realities that integrate agents as digital overlays (see Dorfmueller-Ulhaas & André, 2006). Mixed

realities allow learners to perform real-world tasks in space which may not only contribute to a better understanding of spatial concepts, but also increase their engagement. Learners and agents do no longer inhabit separated spaces, but share a physical reality. To assist the user in such environments, the agent needs not only be aware of the user's virtual world, but also of his or her physical surroundings. As a consequence, such interfaces put high demands to the agent's perceptive skills.

### **Long-Term Interaction with Pedagogical Agents**

Most pedagogical agents have only been designed for short-term interactions with a user. In the best case, the agents adapt their behavior to the user's current knowledge and the dialogue state. However, the systems do not emulate the social dynamics of long-term relationships between user and agent. A great challenge for the future is the design of virtual teachers and advisors that are able to create affective bonds with a user over a longer period of time. A first approach has been presented by Bickmore and Cassell (2005) who maintain an explicit representation of the dynamically changing social relationships between user and agent over a series of conversations. A great challenge is the avoidance of repetitions in interaction in order to ensure that the users are still interested in conversing with agents after a longer period of time. Characters that accompany a user over a longer period of time should be able to remember earlier conversations. Within the FearNot! System (see the second section), a first attempt has been made by Ho and Watson (2006) to equip characters with autobiographic knowledge which contains the agents' individual history and is constructed dynamically during their lifetime. The autobiographical knowledge influences the characters' behaviors in subsequent episodes resulting into a higher believability of the agent. Furthermore, it enables the characters

to tell children about past experiences which may help to establish empathetic relationships.

## CONCLUSION

Research has brought about a great variety of characters in educational contexts that differ widely with respect to audio-visual embodiment, observable behavior and communication skills, and internal cognitive complexity. Such variants may have a great impact on subjective factors, such as the perception of the interaction with the agent, and objective factors, such as learning progress. In this chapter, approaches that facilitate the development of agents for educational environments including the extraction of multimodal tutorial strategies from human-human teaching dialogues as well as the simulation and evaluation of such strategies in computer-mediated learning environments have been discussed. A lot of challenges remain to be solved in order to create truly interactive virtual humans for learning and advisory software that appear to possess at least some qualities of human teachers in a restricted domain. To achieve this goal, the asymmetry in communication channels needs to be diminished. Since humans generally expect human-like verbal and non-verbal conversational behaviors of such agent, it is inevitable to enrich them with social competencies. Many more empirical studies are required to shed light on the question of when to use a pedagogical agent and how it should behave when interacting with a specific user in a specific situation. There is a large variability in learners which influences how the agent is perceived and which impact it has on learning. To derive reliable guidelines for the design of pedagogical agents, a systematic investigation of the crucial variables that influence the user's attitude towards the agent and the effectiveness of the agent are inevitable.

## ACKNOWLEDGMENT

The research presented here was partially sponsored by the EU FP6 Network of Excellence Humaine and by the EU FP6 Project E-Circus (Education through Characters with Emotional-Intelligence and Roleplaying Capabilities that Understand Social Interactions). Special thanks to Ana Paiva, Patrick Gebhard, and Lewis Johnson for giving the permission to use figures 1, 2 and 3. The copyright is to their organizations.

## REFERENCES

- Ai, H., Litman, D. J., Forbes-Riley, K., Rotaru, M., Tetreault, J., & Purandare, A. (2006). Using system and user performance features to improve emotion detection in spoken tutoring dialogs. In *Proceedings of INTERSPEECH-2006*. Pittsburgh, PA.
- Aist, G., Kort, B., Reilly, R., Mostow, J., & Picard, R. W. (2002). Adding human-provided emotional scaffolding to an automated reading tutor that listens increases student persistence. In *Proceedings of the 6th International Conference on Intelligent Tutoring Systems ITS '02*. London: Springer.
- André, E., Rist, T., van Mulken, S., Klesen, M., & Baldes, S. (2000). The automated design of believable dialogues for animated presentation teams. In J. Cassell, S. Prevost, J. Sullivan, & E. Churchill (Eds.), *Embodied conversational agents* (pp. 220-255). The MIT Press.
- Aylett, R., Dias, J., & Paiva, A. (2006). An affectively-driven planner for synthetic characters. In D. Long, S. Smith, D. Borrajo, & L. McCluskey (Eds.), *Proceedings of the Sixteenth International Conference on Automated Planning and Scheduling (ICAPS 2006)*. AAAI Press.
- Bates, J. (1994). The role of emotion in believable agents. *Commun. ACM*, 37(7), 122-125.

- Baylor, A. L., & Kim, Y. (2004). Pedagogical agent design: The impact of agent realism, gender, ethnicity, and instructional role. *Intelligent Tutoring Systems*, 592-603.
- Beun, R. J., & Witteman, C. L. M. (2003). Embodied conversational agents: Effects on memory performance and anthropomorphisation. In T. Rist, R. Aylett, D. Ballin, & J. Rickel (Eds.), *Proceedings of the 4<sup>th</sup> International Workshop on Intelligent Virtual Agents* (pp. 315-319). Berlin: Springer.
- Bickmore, T., & Cassell, J. (2001). Relational agents: A model and implementation of building user trust. In *Proceedings of CHI '01* (pp. 396-403). ACM Press.
- Bickmore, T., & Cassell, J. (2005). Social dialogue with embodied conversational agents. In J. van Kuppevelt, L. Dybkjaer, & N. O. Bernsen (Eds.), *Advances in natural, multimodal dialogue systems* (pp. 23-54). Springer.
- Bickmore, T. W., & Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Trans. Computer-Human Interaction*, 12, 293-327.
- Bosma, W., & André, E. (2004). Exploiting emotions to disambiguate dialogue acts. *Proceedings of the 9<sup>th</sup> International Conference on Intelligent User Interfaces IUI '04* (pp. 85-92). New York: ACM Press.
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.
- Cassell, J., Ananny, M., Basu, A., Bickmore, T., Chong, P., Mellis, D., et al. (2000). Shared reality: physical collaboration with a virtual peer. In *Proceedings of CHI '00 Extended Abstracts on Human Factors in Computing Systems* (pp. 259-260). New York: ACM Press.
- Conati, C. (2002). Probabilistic assessment of user's emotions during the interaction with educational games. *Applied Artificial Intelligence*, 16, 555-575.
- Craig, S., Gholson, B., Garzon, M., Hu, X., Marks, W., Wiemer-Hastings, P., et al. (1999). *Auto tutor and otto tudor*. Paper presented at the AIED-Workshop on Animated and Personified Pedagogical Agents, Le Mans, France.
- deRosis, F., Pelachaud, C., Poggi, I., Carofiglio, V., & Carolis, B. D. (2003). From Greta's mind to her face: Modelling the dynamics of affective states in a conversational embodied agent [Special Issue on Applications of Affective Computing in HCI]. *International Journal of Human-Computer Studies*, 59, 81-118.
- D'Mello, S., & Graesser, A. (2006). Affect detection from human-computer dialogue with an intelligent tutoring system. In J. Gratch, M. Young, R. Aylett, D. Ballin, & P. Olivier (Eds.), *Proceedings of the 6<sup>th</sup> International Conference on Intelligent Virtual Agents (IVA 2006)* (pp. 54-67).
- Dorfmueller-Ulhaas, K., & André, E. (2005). The synthetic character Ritchie: First steps towards a virtual companion for mixed reality. In *Proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR'05)*.
- Dybkjær, L., & Bernsen, O. (2004). Recommendations for natural interactivity and multimodal annotation schemes. In *Proceedings of the LREC 2004 Workshop on Multimodal Corpora* (pp. 5-8).
- Foster, M. E., & Oberlander, J. (2006). Data-driven generation of emphatic facial displays. In *Proceedings of the 11<sup>th</sup> Conference of the European Chapter of the Association for Computational Linguistics (EACL)*.
- Graesser, A. C., Person, N. K., Harter, D., & The Tutoring Research Group (2001). Teaching tactics and dialog in AutoTutor. *International Journal of Artificial Intelligence in Education*, 12, 257-279.

- Gulz, A., & Haake, M. (2006). Design of animated pedagogical agents: A look at their look. *Int. Journal of Human-Computer Studies*, 64, 281-394.
- Ho, W. C., & Watson, S. (2006). Autobiographic knowledge for believable virtual characters. In J. Gratch, M. Young, R. Aylett, D. Ballin, & P. Olivier (Eds.), *Proceedings of the 6<sup>th</sup> International Conference on Intelligent Virtual Agents, IVA 2006* (pp. 383-394).
- Johnson, W. L., Beal, C. R., Fowles-Winkler, A., Lauper, U., Marsella, S., Narayanan, S., Papachristou, D., et al. (2004a). Tactical language training system: An interim report. *Intelligent Tutoring Systems*, 336-345.
- Johnson, W. L., Rizzo, P., Bosma, W., Kole, S., Ghijsen, M., & van Welbergen, H. (2004b). Generating socially appropriate tutorial dialog. In E. André, L. Dybkjær, W. Minker, & P. Heisterkamp (Eds.), *Affective dialogue systems (ADS)* (pp. 254-264).
- Kapoor, A., & Picard, R. W. (2005). Multimodal affect recognition in learning environments. *Proceedings of the 13<sup>th</sup> Annual ACM International Conference on Multimedia MULTIMEDIA '05* (pp. 677-682). New York: ACM Press.
- Kipp, M., Kipp, K. H., Ndiaye, A., & Gebhard, P. (2006). Evaluating the tangible interface and virtual characters in the interactive exhibit. In J. Gratch, M. Young, R. Aylett, D. Ballin, & P. Olivier (Eds.), *Proceedings of the 6<sup>th</sup> International Conference on Intelligent Virtual Agents (IVA 2006)* (pp. 434-444).
- Kort, B., Reilly, R., & Picard, R. W. (2001). An affective model of interplay between emotions and learning: Reengineering educational pedagogy—building a learning companion. *ICALT*, 43-48.
- Krämer, N. C. (2005). Social communicative effects of a virtual program guide. In *Proceedings of the 5<sup>th</sup> International Conference on Intelligent Virtual Agents (IVA 2005)* (pp. 442-453).
- Lee, K. M., & Nass, C. (2003). Designing social presence of social actors in human computer interaction. In *Proceedings of the SIGCHI Conference on Human factors in computing systems CHI '03* (pp. 289-296). New York: ACM Press.
- Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., & Bhogal, R. S. (1997). The persona effect: Affective impact of animated pedagogical agents. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI '97* (pp. 359-366). New York: ACM Press.
- Lester, J. C., Towns, S. G., Callaway, C. B., Voerman, J. L., & FitzGerald, P. J. (2000). Deictic and emotive communication in animated pedagogical agents. In J. Cassell, S. Prevost, J. Sullivan, & E. Churchill (Eds.), *Embodied conversational agents* (pp. 123-154). Cambridge, MA: MIT Press.
- Marsella, S., & Gratch, J. (2000). Modeling the interplay of emotions and plans in multi-agent simulations. In *Proceedings of the 23<sup>rd</sup> Annual Conference on Cognitive Science Society* (pp. 294-599). Mahwah, NY: Lawrence Erlbaum.
- Massaro, D. W. (2004). Symbiotic value of an embodied agent in language learning. In *Proceedings of the 37<sup>th</sup> Annual Hawaii International Conference on System Sciences (HICSS'04)*. Washington, DC: IEEE Computer Society.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago.
- McQuiggan, S. W., & Lester, J. C. (2006). Diagnosing self-efficacy in intelligent tutoring systems: An empirical study. *Intelligent Tutoring Systems*, 565-574.
- Merten, C., & Conati, C. (2006). Eye-tracking to model and adapt to user meta-cognition in intelligent learning environments. In *Proceedings of*

- the 11<sup>th</sup> International Conference on Intelligent User Interfaces IUI '06 (pp. 39-46). New York: ACM Press.
- Mulken, S. V., André, E., & Müller, J. (1998). The persona effect: How substantial is it? In *Proceedings of HCI on People and Computers XIII HCI '98* (pp. 53-66). London: Springer-Verlag.
- Nakano, Y. I., Reinstein, G., Stocky, T., & Cassell, J. (2003). Towards a model of face-to-face grounding. *ACL*, 553-561.
- Nass, C., & Gong, L. (2000). Speech interfaces from an evolutionary perspective. *Commun. ACM*, 43, 36-43.
- Paiva, A., Dias, J., Sobral, D., Aylett, R., Sobreperez, P., Woods, S., et al. (2004). Caring for agents and agents that care: Building empathic relations with synthetic agents. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems AAMAS '04* (pp. 194-201). Washington, DC: IEEE Computer Society
- Person, N. K., Graesser, A., Kreuz, R., Pomeroy, V., & TRG. (2001). Simulating human tutor dialog moves in autotutor. *International Journal of Artificial Intelligence in Education*, 12, 23-39.
- Pilkington, R. (1999). *Analysing educational discourse: The discount scheme*. Technical Report 99/2, Computer Based Learning Unit, University of Leeds.
- Prendinger, H., & Ishizuka, M. (2001). Social role awareness in animated agents. In *Proceedings of the Fifth International Conference on Autonomous Agents AGENTS '01* (pp. 270-277). New York: ACM Press.
- Prendinger, H., Mayer, S., Mori, J., & Ishizuka, M. (2003). Persona effect revisited. Using bio-signals to measure and reflect the impact of character-based interfaces. In T. Rist, R. Aylett, D. Ballin, & J. Rickel (Eds.), *Intelligent virtual agents* (pp. 283-291). Berlin, Heidelberg: Springer.
- Qu, L., & Johnson, L. (2005). Detecting the learner's motivational states in an interactive learning environment. In C. K. Looi, G. McCalla, B. Bredeweg, & J. Breuker, (Eds.), *Proceedings of the 12<sup>th</sup> Int. Conf. on Artificial Intelligence in Education (AIED)* (pp. 547-554). IOS Press.
- Rajan, S., Craig, S., Gholson, B., Person, N., & Graesser, A. (2001). Autotutor: Incorporating backchannel feedback and other human-like conversational behaviors into an intelligent tutoring system. *International Journal of Speech Technology*, 4, 117-126.
- Rehm, M., & André, E. (2007). Informing the design of embodied conversational agents by analysing multimodal politeness behaviours in human-human communication. In T. Nishida (Ed.), *Engineering approaches to conversational informatics* (pp. 69-84). Chichester, UK: John Wiley & Sons.
- Rickel, J., & Johnson, W. L. (1999). Animated agents for procedural training in virtual reality: Perception, cognition, and motor control. *Applied Artificial Intelligence*, 13, 343-382.
- Rickel, J., Marcella, S., Gratch, J., Hill, R., Traum, D., & Swartout, W. (2002). Toward a new generation of virtual humans for interactive experiences. *IEEE Intelligent Systems*, 17, 32-38.
- Sidner, C. L., Lee, C., Kidd, C. D., Lesh, N., & Rich, C. (2005). Explorations in engagement for humans and robots. *Artif. Intell.*, 166, 140-164.
- Stone, M., DeCarlo, D., Oh, I., Rodriguez, C., Stere, A., Lees, A., et al. (2004). Speaking with hands: Creating animated conversational characters from recordings of human performance. *ACM Transactions on Graphics*, 23(3), 506-513.
- Tartaro, A., & Cassell, J. (2007). Using virtual peer technology as an intervention for children with autism. In J. Lazar (Ed.), *Towards universal usability: Designing computer interfaces for diverse user populations* Chichester, UK: John Wiley & Sons.

van Vugt, H. C., Konijn, E. A., Hoorn, J. F., & Veldhuis, J. (2006). Why fat interface characters are better e-health advisors. In J. Gratch, M. Young, R. Aylett, D. Ballin, & P. Olivier (Eds.), *Proceedings of the 6<sup>th</sup> International Conference on Intelligent Virtual Agents (IVA 2006)* (pp. 1-13).

Wang, N., Johnson, L., Mayer, R., Rizzo, P., & Shaw, E. (2005). The politeness effect: Pedagogical agents and learning gains. In C. K. Looi, G. McCalla, B. Bredeweg, & J. Breuker (Eds.), *Proceedings of the 12th Int. Conf. on Artificial Intelligence in Education (AIED)* (pp. 686-693). IOS Press.

## ADDITIONAL READING

Apart from the special issues of journals, surveys, and books listed, we recommend to regularly check the proceedings of:

Proceedings of the International Conferences on Intelligent Virtual Agents.

Proceedings of the International Conference on Artificial Intelligence in Education.

Proceedings of the International Conferences on Intelligent Tutoring Systems.

André, E. (Ed.). (1999). Animated interface agents [Special issue]. *Journal of Applied Artificial Intelligence*, 13(4-5).

André, E. (Ed.). (2000). Behavior planning for lifelike characters and avatars [Special issue]. *Artificial Intelligence Communications Journal*

André, E., Dybkjær, L., Minker, W., & Heisterkamp, P. (Eds.). *Affective dialogue systems (ADS)*. Springer.

André, E. and Pelachaud, E. (in press). Interacting with embodied conversational agents. In F. Chen

& K. Jokinen (Eds.), *New trends in speech based interactive systems*. Springer.

Baylor, A. (Ed.) (2007). Pedagogical agents [Special issue]. *Educational technology: The magazine for managers of change in education*, 47(1),

Cole, R. A., Van Vuuren, S., Pellom, B., Hacıoglu, K., Ma, J., Movellan, J., et al. (2003). Perceptive animated interfaces: First steps toward a new paradigm for human-computer interaction [Special issue]. *IEEE Human-Computer Multimodal Interface*, 91(9), 1391-1405.

Dautenhahn, K., Bond, A. H., Canamero, L., & Edmonds, B. (2002). *Creating relationships with computers and robots*. Springer.

Dehn, D. M., & van Mulken, S. (2000). The impact of animated interface agents: A review of empirical research. *International Journal of Human-Computer Stud.*, 52(1), 1-22.

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

Gratch, J., & Marsella, S. (2005). Some lessons for emotion psychology for the design of lifelike characters [Special issue]. *Journal of Applied Artificial Intelligence*, 19(3-4), 215-233.

Gratch, J., Rickel, J., André, E., Badler, N., Cassell, J., & Petajan, E. (2002). Creating interactive virtual humans: Some assembly required. *IEEE Intelligent Systems*.

Gulz, A. (2004). Benefits of virtual characters in computer based learning environments. *Claims and Evidence*, 14(3-4), 313-334.

Johnson, L. W., Rickel, J. W., & Lester, J. C. (2000). Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *The International Journal of Artificial Intelligence in Education*, 11, 47-78.

Lester, J., Branting, K., & Mott, B. (2004). Conversational agents. In M. Singh (Ed.), *Practical*

- handbook of Internet computing* Baton Rouge: Chapman Hall & CRC Press.
- Magenat-Thalmann, N., & Thalmann, D. (2001). *Deformable avatars*. Kluwer.
- Magenat-Thalmann, N., & Thalmann, D. (2004). *Handbook of virtual humans*. Wiley.
- Mates, M., & Sengers, P. (Eds.). (2003). *Narrative intelligence*. John Benjamins.
- Nishida, T. (2007). *Engineering approaches to conversational informatics*. Chichester, UK: John Wiley & Sons.
- Paiva, A. (2000). *Affective interactions, towards a new generation of computer interfaces*. Springer.
- Payr, S. (Ed.). (2005). Educational agents and (e-)learning [Special issue]. *Journal of Applied Artificial Intelligence*, 19(3-4).
- Payr, S., & Trappl, R. (Eds.). (2004). *Agent culture: Human-agent interaction in a multicultural world*. London: Lawrence Erlbaum Associates.
- Pelachaud, C., & Cañamero, L. (Eds.). (2006). Achieving human-like qualities in interactive virtual and physical humanoids [Special issue]. *International Journal of Humanoid Robotics*.
- Person, N. P., & Graesser, A. C. (2006). Pedagogical agents and tutors. In J. W. Guthrie (Ed.), *Encyclopedia of education* (pp. 1169-1172). New York: Macmillan.
- Prendinger, H., & Ishizuka, M. (Eds.). (2004). *Life-like characters. Tools, affective functions, and applications cognitive technologies series*. Berlin-Heidelberg: Springer.
- Reeves, B., & Nass, C. (2003). *The media equation: How people treat computers, television, and new media like real people and places*. Chicago: University of Chicago Press.
- Rist, T., André, E., Baldes, S., Gebhard, P., Kleisen, M., Kipp, M., et al. (2003). A review of the development of embodied presentation agents and their application fields. In H. Prendinger & M. Ishizuka (Eds.) *Life-like characters: Tools, affective functions, and applications*. (pp. 377-404). London: Springer.
- Ruttkay, Z., & Pelachaud, C. (Eds.). (2004). From brows to trust: Evaluating embodied conversational agents. In *Human-Computer Interaction Series (Vol. 7)*. Dordrecht, The Netherlands: Kluwer.
- Trappl, R., & Petta, P. (Eds.). (1997). *Creating personalities for synthetic actors*. Berlin/Heidelberg/New York/Tokyo: Springer,
- Trappl, R., Petta, P., & Payr, S. (Eds.). (2003). *Emotions in humans and artifacts*. Cambridge, MA: MIT Press.
- Waern, A., & Höök, K. Interface Agents. (2001). A new interaction metaphor and its application to universal accessibility. In C. Stephanidis (Ed.), *User interfaces for all: Concepts, methods, and tools* [A volume in the Human Factors and Ergonomics Series].