Virtual Substitute Teacher: Introducing the Concept of a Classroom Proxy

Doron FRIEDMAN,[a,1] Oren SALOMON[a] and Béatrice S. HASLER[a]

[a]*Advanced Reality Lab, Sammy Ofer School of Communications, Interdisciplinary Center Herzliya, Israel*

Abstract. We describe the concept and a first prototype of a digital replacement for a teacher in class. While earlier work on intelligent pedagogical agents has focused on the development of a generic idealized teacher agent, our focus lies on modeling the identity of a *specific* teacher. This digital replacement is called his "classroom proxy" and acts on his behalf as a personalized, virtual substitute teacher. We evaluate the concept of the classroom proxy in an exploratory case study and discuss the broader implications of digital replacements for educational purposes and other applications. We conclude with a description of future research directions.

Keywords. Intelligent virtual agents, autonomous agents, digital clones, mixed initiative, mixed-reality teaching

Introduction

Like every other human, teachers occasionally become ill, have emergencies, and cannot go to work. It is often impossible to find a qualified substitute teacher [1]. In higher education, classes are often canceled if no short-time replacement is at hand. In order to maintain quality educational standards, it is always desirable that the regular teacher is teaching a class, or at least that the teacher is carefully instructing his/her replacement. However, teachers' absences are often at short notice and a substitute teacher needs to improvise due to the lack of sufficient preparation time. A substitute teacher takes on a great deal of responsibility in holding things together until the regular teacher returns [2]. This means not only to move the curriculum forward but to ensure a smooth-running classroom [3]. The substitute teacher needs to be informed about what needs to be taught in a particular class and what the expectations are, for example, regarding the students' participation. In the current paper, we explore the question of how technology can help to overcome the issue of teacher absence by providing the best replacement possible.

These days we hold many personal and professional meetings that do not take place face-to-face (FTF) but are mediated through communication technologies, such as mobile phones, video conference systems, social networks, or online virtual worlds. This opens the possibility that our representation in the communication medium would be controlled by autonomous software rather than by ourselves. The most well-known digital replacements are automatic out-of-office email replies and call answer machines.

---

[1] Corresponding Author.

But also automated avatars (i.e., bots) are increasingly populating social virtual worlds, such as *Second Life* (SL), that are able to initiate or continue a social interaction while the human operator behind the avatar is AFK (i.e., "away from keyboard").

We take the concept of autonomous personal agents into an educational setting, and ask how a virtual substitute teacher could fill in while the regular teacher is sick, attending a conference, or taking care of urgent family business. Our virtual substitute teacher is essentially a digital clone of the absent teacher that resembles him regarding his appearance and behavior. We call this personalized digital replacement of a teacher, his "classroom proxy." As the term "proxy" implies, this replacement will not only resemble him, but is authorized to act on his behalf.

The proxy is a specific type of an intelligent virtual agent, sometimes referred to as a virtual human. The idea of virtual humans is a popular science fiction theme. Science fiction movies often portray autonomous artificial humans, as in *Blade Runner*, or remote control of a virtual or physical representation, as seen in the movie *Surrogates*. Some of these entities are reminiscent of our vision of a proxy replacement; for example, in the book series *Safeload* the writer David Weber introduces the concept of a personality-integrated cybernetic avatar (commonly abbreviated to PICA): this is a safe and improved physical representation of yourself in the physical world, and you can also opt to "upload" your consciousness into it. Such concepts are penetrating popular culture; for instance, the appearance of digital Tupac at the Coachella festival in 2012, which provoked much discussion of the social implications and ethical use of digital technology (the original Tupac was a rap musician killed in 1996).

Clarke [4] introduced the concept of a digital persona, but this was a passive entity made of collected data, rather than an autonomous or semi-autonomous agent. Today we see such virtual identities in online social networks, such as *Facebook* profiles, but these are passive representations. Ma et al. [5] presented a manifesto of something they call a cyber-individual, which is "a comprehensive digital description of an individual human in the cyber world" (p. 31). This manifesto is generic and is presented in a high level of abstraction, and it is partially related to our concept of a proxy. The notion that computers in general have become part of our identity in the broadest sense is being discussed by psychologists [6] and philosophers (e.g., as in the extended mind theory) [7]. In this project we take a step in making these ideas concrete, toward merging the identity of the person with his avatar representation; specifically, the teacher's identity with his virtual substitute.

We can think of many other cases where we would have loved to have a proxy to replace us in an annoying, boring, or unpleasant event. However, in what contexts will such digital replacements of our selves be socially acceptable? Hence, the current paper not only evaluates the performance of a first prototype of a proxy used in a classroom context, but also evaluates the social acceptance of the proxy concept in general.

1. Related Work

1.1. Intelligent virtual agents

Intelligent virtual agents have been studied widely, either as autonomous or semi-autonomous entities [8, 9]. There has been a lot of research on believability [10], expressiveness and multi-modal communication [11], and the role of nonverbal behavior in coordinating and carrying out communication [12].

Semi-autonomous avatars were first introduced by Cassell et al. [13, 14]. Their system enables users to communicate via text while their avatar automatically animates attention, salutations, turn taking, back-channel feedback, and facial expression. Following their seminal work, there have been many attempts at automating nonverbal communication (see [12] for a review). Penny et al. [15] describe a system that incorporates avatars with varying levels of autonomy: *Traces*, a Virtual Reality (VR) system in which a user's body movements spawn avatars that gradually become more autonomous.

Shared control has been discussed in various fields of intelligent systems under titles like "adjustable autonomy" and "mixed initiative." According to Bradshaw et al. [16], adjustable autonomy is when the "degree of autonomy is continuously and transparently adjusted in order to meet whatever performance expectations have been imposed by the system designer and the humans and agents with which the system interacts. (…) Thus, a primary purpose of adjustable autonomy is to maintain the system being governed at a sweet spot between convenience (i.e., being able to delegate every bit of an actor's work to the system) and comfort (i.e., the desire to not delegate to the system what it can't be trusted to perform adequately)" (p. 240). Mixed initiative was defined by Allen [17] as "a flexible interaction strategy, where each agent can contribute to the task what it does best" (p. 14). Both, automating nonverbal behavior and shared control play an important role in our realization of the classroom proxy concept.

1.2. Pedagogical agents

IVAs have also been built for educational purposes, mainly in the area of intelligent tutoring systems, and are typically referred to as "intelligent pedagogical agents" (IPAs) (see [18] for a review). Graesser et al. [19] present an IPA, called "AutoTutor," which simulates an idealized version of a human tutor capable of holding a natural-language dialogue with a learner. They implemented the pedagogical strategy of "deep reasoning" into the agent, which engages in a mixed-initiative question-answer interaction with the learner, and shows context-sensitive synthetic speech, facial expressions and gestures. Others have implemented advanced emotional capabilities into IPAs. For example, Xuejing et al. [20] present an IPA that considers the learner's emotions by adapting to his psychological state throughout the interactive learning and teaching process. Besides such individualized learning scenarios, IPAs can also be used for supporting collaborative learning in collaborative online learning communities. It is generally possible to have multiple students and agents interacting in a shared virtual environment. Dowling [21] describes the complex roles of such a socially interactive IPA.

While common IPA research aims at simulating an ideal teacher, their efforts typically remain within imitating natural learner-instructor interactions. Bailenson, Blascovich and colleagues [22, 23] took the idea of an idealized virtual teacher into an immersive VR environment where both the teacher and the students are represented by avatars and/or automated agents. Following their Transformed Social Interaction (TSI) approach [24], they focus on the unique properties of immersive VR technology in order to create interactions that are not possible in FTF settings. In one of their experiments, they used VR technology in order to design a teacher to appear engaging in mutual eye contact with each of the virtual students simultaneously. Receiving maximum teacher attention through such augmented eye gaze has been found to result

in more persuasive instruction. They also used gaze feedback information for teacher training purposes. Participant instructors learned to distribute their attention more evenly to their virtual audience when they received visual feedback about their gaze directions.

In summary, these related lines of research attempt to simulate a human teacher (or more precisely, an idealized, digital version of a teacher), and alter teachers' and students' online representations in order to facilitate learning. However, none of them have attempted to model a *specific* teacher, including his identity, expertise and preferences. This is the main goal of our research into designing a personalized classroom proxy.

1.3. Our previous work

Our first-generation proxy inhabited the virtual world *Second Life*. It is based on our SL bot platform, which enables bots to perform useful tasks, such as automatically conducting large-scale survey interviews [25]. The functionality of the SL bots has later been extended to learn from other avatars' behavior within SL, and imitate them (e.g., social navigation patterns) [26]. An actual proxy in the sense of a virtual clone of a specific person has been prepared for one of the co-authors of this paper to give a talk inside SL as part of a real-world conference (Figure 1): a workshop on Teaching in Immersive Worlds, Ulster, Northern Ireland, in 2010. The appearance of the proxy was canceled on the day of the event due to audio problems in the conference venue, but a video illustrates the vision and concept.[2]

In this paper we take the concept of the proxy to the offline world, though we take advantage of the fact that the offline world is also penetrated with digital sensors (cameras, microphones, mobile phones) and digital effectors (in this case a standard audio-visual display and loudspeakers) that enable a virtual proxy to take part in it.
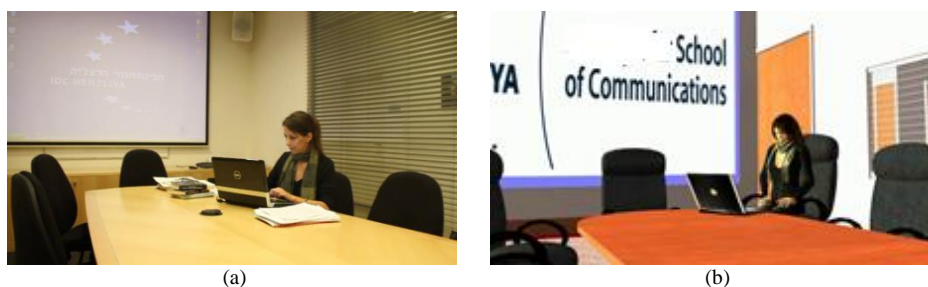

(a)	(b)

Figure 1. (a) The second co-author of this paper at work. (b) The virtual-world avatar, used as her proxy.

2. Proxy Concept

The proxy is part of an EU research project, called BEAMING.[3] The project deals with the science and technology intended to give people a real sense of physically being in a

---

[2] http://www.youtube.com/watch?v=7RByLajho-w
[3] http://beaming-eu.org/

remote location with other people, and vice versa – without actually traveling [27]. As part of this project, we explore the concept of the proxy in the context of several applications: remotely attending a business meeting, remote rehearsal of a theater play, remote medical examination, and remote teaching. In each of these applications it is useful to have a digital extension of yourself that is able to assist you while you are distracted with other tasks or replace you completely, either for short periods of a few seconds, due to some interruption, or even for a complete event.

In the BEAMING scenarios, we assume a mixed reality rather than a virtual environment. There is a main physical space, termed the destination, and its inhabitants are called locals. Remote users "beam" into this physical space using various tele-presence technologies, and the proxy is expected to operate in the same setup. In general, the proxy can be embodied as a virtual agent or as a physical robot. As part of the BEAMING project, we are exploring both, but in this paper we focus on a 3d virtual agent projected on a large screen.

The long-term goal of the proxy is to capture several aspects of a specific person, so that the proxy would be able to replace that person in various real-life contexts. There are many aspects of a person that we may wish to capture in the proxy: appearance, verbal and nonverbal communication styles, personality, preferences, professional knowledge, and more. In order to fulfill its purpose as a "proxy" (i.e., an authorized entity that acts on behalf of its owner), the proxy does not only need to pass as a recognizable representation of its owner, it also has to represent the interests of its owner. To that purpose, the proxy needs to be aware of its owner's goals and preferences. This is clearly a very ambitious goal, which would require integration of many fields for many years, and in this paper we only present a first prototype of a proxy used as a substitute teacher in a mixed-reality classroom setting.

The main concern in the implementation of our first prototype was the proxy's modes of operation. Other factors, such as physical appearance that resembles its owner and an accurate replication of the owner's body language, were treated as secondary. The proxy has three modes of operation: (a) background mode, (b) foreground mode, and (c) mixed-mode operation.

While the proxy is in background mode, its owner has full control over the actions. The owner can talk "through" the proxy. In this case he is represented by his avatar. Alternatively, the proxy can be idle, and simply record its owner's behavior in the background. The data collected while the proxy is in background mode provides the main source for behavior models for the proxy. Typically, we record the proxy owner's nonverbal behavior; the skeleton tracking data is tagged with meta-data regarding the context. This data is later being used in foreground and mixed mode in order to allow the proxy to have its owner's body language. During regular lectures, the proxy could always be in background mode and learn not only *what* the teacher is lecturing about, but also *how* he is teaching (e.g., movement patterns in the room, eye contact with students, length of turns between lecturing and addressing questions, etc.).

When the proxy is in foreground mode, it takes full control over the actions. In the role of a substitute teacher, the proxy needs to represent the teacher in the best way possible. In the simplest case, this means delivering a lecture and answering students' questions. In some situations, the owner may not wish to give the proxy full control, but decides to control part of the communication himself. In this case of shared control the proxy operates in mixed mode. The spectrum between user control and agent control is illustrated in Figure 2.
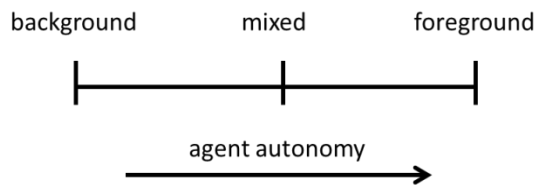
Figure 2. The spectrum from full agent autonomy (foreground mode) on the one side to full human autonomy (background mode) on the other, with mixed mode in the middle.

Our approach aims at a seamless transition between modes of operation throughout a communication session. The decision about switching between modes during a session can be made by the proxy automatically. The proxy should know when to take control of the remote representation (i.e., switch to foreground mode), and when to release control back to the user (i.e., switch back to background mode).

3. Evaluation Study

We evaluated a prototype of the classroom proxy as a proof of concept in an exploratory case study. As pilot tests we sent the proxy to replace the first author of the current paper to give short presentations in two international work meetings of the BEAMING project and in one appearance in front of a small group of students who were aware of the research goals. These were relatively informal pilot sessions, which were used to refine the study. In the evaluation study presented here, a classroom proxy was built for the first author, and was sent to replace him in one of his academic lectures. The case study is documented in a video clip.[4]

3.1. Participants

The evaluation study was conducted in a class of 3rd-year communication students, from an international track at IDC Herzliya. The students were expecting a regular class and did not have previous knowledge of the proxy. We deliberately selected social science students rather than students with a technical background. The class, entitled "Topics in Digital Culture," emphasizes the tension between digital utopias and dystopias. Thus, we expect the students to be aware of the social opportunities and risks of advanced technologies. The study received confirmation of our institutional review board. Twenty-one students of 11 nationalities filled in the questionnaires as part of the evaluation study (ages 21-29, $M = 23.9$, $SD = 2.47$, 7 men and 14 women).

3.2. Classroom Proxy Scenario

The prototype scenario included a classroom as the physical destination. The proxy's owner teleported to class from his research lab located in another building on campus (see Figures 3, 4). We tested three modes of operation of the proxy: (a) foreground mode, in which the teacher is in control of the avatar; (b) background mode, in which

---

[4] http://www.youtube.com/watch?v=43l739kKFfk

the teacher is away and the proxy "covers up" for him, and (c) mixed-mode operation, in which the teacher and the proxy each have partial control.

The classroom was prepared for the session. A machine running the proxy representation system was connected to the projector and loud speakers, and a webcam captured audio and video streams (with standard quality provided by the campus Wi-Fi network) of the classroom, and sent the streams to the lab (for the teacher, when the proxy was in background mode). The camera was positioned on the lecturer podium in front of the class. The lab interaction space includes a "power-wall" setup: a back-projection screen and a few standard sensors intended to capture the teacher: a high-quality camera, a depth camera (Kinect device), and a wearable wireless microphone. The teacher was speaking in front of the large screen, on which he could see the video feed from the classroom, and also a smaller display with the avatar that represents him, as seen by the class. Thus, unlike desktop-based interactions, the body language during this part of the experiment was appropriate and is expected to have resembled the teacher's body language in front of a class.

Figure 3. Screenshots from the evaluation study. Top left: the classroom. Top right: the proxy representation as projected in class. Bottom: the teacher receiving a Skype call from his proxy.
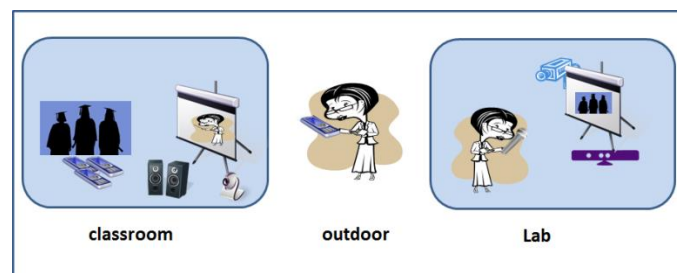

Figure 4. A schematic diagram of the study setup, showing the teacher in the lab presented as an avatar on the screen in the classroom. The students communicate with the proxy using mobile devices. During the experiment the teacher was also outdoors and communicated with the class in mixed mode.

The setup was prepared before the students arrived. At the first stage the teacher teleported into the class. He talked to the class with his own voice and controlled the avatar's body language (as captured by the Kinect device). After a few seconds the teacher left the lab and the proxy automatically took over and switched to the second phase of the study, in which it was in foreground mode. The proxy started reading a lecture from a predefined text. The proxy's body language was based on pre-recorded body movements of the teacher, and the voice was automatically generated with a standard text-to-speech engine (from Microsoft). The proxy has a list of words that should be accentuated when lecturing (prepared manually), and a few slides were displayed in the background in sync with the presentation. The topic of the lecture fragment was post-humanism vs. trans-humanism, with some details provided about the scholar Francis Fukuyama and his views on that matter.

Early during the talk the proxy encouraged the students to ask questions. Since speech-recognition technologies today cannot handle classroom conditions, the proxy instructed the students to send questions by Twitter, using a specific hash tag. Since these were new-media students they were all familiar with Twitter, and all had digital devices capable of running Twitter during class. Five questions were sent during the proxy's talk. The first 4 questions were answered by the proxy automatically after reading the questions out loud. The chat bot was, in general, able to provide reasonable answers to the questions based on the pre-configured templates (e.g., Q: where are you now? A: I am in IDC Herzliya).

At some point toward the end of the talk, the proxy decided that a specific question should be addressed by the teacher himself. In the current case study, this decision was made arbitrarily (based on the duration of the presentation fragment). Generally, a certainty criterion can be obtained from the speech recognition and natural-language understanding components, to allow for the proxy to automatically attempt to transfer control back to its owner. The proxy explained to the audience that it was asked a question that it cannot answer, and that it has decided to consult its owner. At this stage the proxy called the teacher who was now on the way from the lab to the classroom. The class could not see this call take place, but they could hear Skype ringing, the proxy asking the question again, and the teacher answering (using his smartphone). A few seconds later the teacher entered the class and explained that this was a scientific study and that the experiment is over. The whole session lasted approximately 15 minutes, after which research assistants handed out a questionnaire, and three students were randomly selected to provide more details in semi-structured interviews. The interviews were videotaped.

3.3. Measures

The students were first asked to rate the believability (i.e., perceived realism) of 9 variables related to the classroom proxy's performance: "How realistic was…?" (voice, facial expression, speaking style, arm movements, body postures and gestures, timing between speech and gestures, content of the talk, responses in the Q&A session, use of language). Two questions were asked in order to measure the perceived similarity regarding (1) the body language and (2) appearance between the proxy and its owner. Each of these items were rated on a scale from 1=not at all, to 5=extremely.

The following questions included measures of social acceptance regarding the proxy concept in general. We asked the students: "Do you think it is generally acceptable that proxies will replace real humans in the future?" This question was rated

on a scale from 1=not at all, to 5=absolutely. We further asked two open-ended questions: (1) "For what kind of social situations would you like to have a proxy that replaces you or wouldn't mind to interact with someone else's proxy?" and (2) "Knowing that you are interacting with someone's proxy, do you think you would behave differently toward a proxy compared to interacting with an avatar controlled by a real person?"

3.4. Results and Discussion

Perceived realism: A median comparison of the ratings of the classroom proxy's believability revealed the following order (from more realistic to less realistic): (1) use of language ($Md = 4$), (2) content of the talk, voice, speaking style, timing between gestures and speech, and Q&A ability ($Md = 3$), (3) facial expressions, body posture, and arm movements ($Md = 2$). The low ratings for the proxy's body language can be explained by the fact that we have used Kinect as a motion capture tool for both the live and pre-recorded animations. Kinect is intended for gesture-based interaction rather than as a motion-capture device, and apparently the resulting body language was unrealistic due to noise.

A principal component analysis of all 9 variables revealed three distinct factors with the following factor loadings: Factor 1: speaking style (.71), arm movements (.92), body postures (.82), and language (.68), Factor 2: voice (.60), facial expression (.68), and content (.76), and Factor 3: timing (.80) and Q&A (.60). This points to a trend where the use of language (selection of words, grammar) and speaking style are attached to the body movements and postures while the content of the conversation is more affected by the voice and facial expression (associated with the head or face area). This possibly implies that although we must match a person's speaking style to his own body movements, the content of what a person is saying may be manipulated independently, without affecting believability. This would be in contradiction to the classic literature regarding the high correspondence between speech and gesture in humans [28]. Three students further elaborated on their experience during the post-study interviews. These reports confirmed that the main obstacles were the body language of the proxy that was not realistic enough, and the audio quality of the live Skype talks.

Perceived similarity: The perceived similarity between the proxy and its owner regarding appearance ($M = 2.33$, $SD = 1.35$), and body language ($M = 2.38$, $SD = .87$) was about average. The proxy was neither perceived as totally dissimilar nor very similar to its owner. Although a high level of similarity would eventually be expected, a moderate level of perceived similarity can be considered as satisfying for a first prototype. As expected, the perceived similarity ratings regarding appearance and body language were positively correlated, $r = .40$, $p = .04$ (one-tailed). This indicates that body language is perceived as an integral part of one's appearance. Interestingly, perceived similarity regarding body language was positively correlated with perceived realism of the proxy's *verbal* behavior, $r = .41$, $p = .03$ (one-tailed). This provides further support for the strong connection between verbal and nonverbal communication styles in creating a coherent, authentic impression.

Social acceptance: The general acceptance of the concept of a proxy was found to be moderate ($M = 2.71$, $SD = 1.38$); indicating some ambivalence regarding its social acceptance. Interestingly, men ($M = 4.00$, $SD = 1.16$) accepted the proxy concept significantly more than women ($M = 2.07$, $SD = 1.0$), $t(19) = 3.97$, $p = .001$.

When asked to specify whether they would use a proxy and for what purposes only two students specified that they would rather not use a proxy at all. The rest of the students suggested that they would like to use proxies for: attending classes (7 mentions), daily errands such as visiting the bank or grocery shopping (4 mentions), work-related activities (3 mentions), or socially-stressful tasks such as public speaking (1 mention) or ending a relationship (1 mention). These answers testify that at least some of the students did not fully comprehend the difference between the proxy and an assistant robot (e.g., shopping), but others clearly grasped the social implications (e.g., ending a relationship).

When asked whether their behavior would be different if they knew that they are interacting with someone's proxy as opposed to a human-controlled avatar, 50% indicated yes, 15% said no, and 35% were uncertain. Those who were not sure about whether or not their behavior would change toward a fully autonomous proxy indicated that it depends on how realistic the proxy was. This finding has important implications in our case of a virtual substitute teacher. We would not only want the classroom proxy to be perceived as realistic and similar to the teacher it is replacing. We would also want the students to behave just as if their real teacher was teaching the class (i.e., treat the proxy with respect, pay attention to what it is saying, etc.).

4. Outlook and Conclusions

We presented an exploratory case study that aimed at evaluating the concept of a classroom proxy used as a virtual replacement of a teacher in class. We deliberately decided to run this case study in a real-world setting rather than in the lab. It is crucial to confront a live audience with the concept of a proxy; in our case the proxy of the students' regular teacher. Only in such real-world scenarios, we can evaluate the social acceptance of the proxy concept as first-hand experience is more informative than being confronted with a merely hypothetical scenario.

For many of us it would be convenient to use proxy replacements, as indicated by the majority of our student audience. However, would we want to live in a society where others are often represented by proxies? There was a significantly higher acceptance of the proxy concept by male participants than female participants, but the majority of both genders would be happy to use a proxy. Based on these preliminary findings we conclude that we are likely to see various kinds of proxies gradually deployed, and envision a future where proxies inhabit our society.

The case study demonstrates that a useful proxy can already be implemented using current state of the art technology. The main drawbacks were in the level of production: the quality of live audio and the animation. The students' ratings are, of course, specific to our implementation of the classroom proxy, and may not be generalized to more advanced technical implementations. However, the evaluation results are useful for ongoing refinement steps of the proxy's development to be used in classroom settings and other types of applications. We encourage the developers of such autonomous intelligent agents to be aware of the social, ethical, and legal implications while pushing the technology further.

One of the main issues for further research, which we have not fully addressed in this paper, is the potential of applying intelligent transformations to the teacher's representation, as proposed in Bailenson et al.'s [24] TSI approach. Since we are operating at the border between the digital and the physical space, an interesting

opportunity presents itself: Instead of replicating one's actual physical appearance and behavior, enhancements can be implemented through the use of digital transformations. The proxy can be used to represent the owner better than the owner would represent him- or herself. For example, you may consider a proxy that is based on your appearance with a beautifier transformation applied [29]. Similarly, in our case study the teacher has opted to use a good-looking male avatar, even though a look-alike avatar was available, and text-to-speech was selected to have an impeccable British accent.

Elsewhere we have demonstrated the possibility of a proxy that extends your vocabulary in foreign gestures [30]. For example, assume that the proxy owner communicates with people of a different culture. In mixed-mode operation, the proxy can be configured to recognize that the owner has performed a culture-specific gesture that may be misinterpreted by his collocutor(s). The proxy would then replace this gesture by the equivalent gesture in the target culture's vocabulary or provide an annotation of the gesture's meaning if no equivalent gesture exists. This scenario is particularly useful in an educational setting. Imagine a teacher is invited to give a lecture in a foreign country. Instead of traveling, he may want to send his proxy as a culturally-adapted version of himself. Besides adjusting the proxy's body language to the cultural norms of the audience, it is able to give a talk in any foreign language. The lecture script can be easily translated and presented using text-to-speech software. Considering the constant improvement in translation systems and synthetic speech, this scenario will not be so uncommon in the near future.

In the case of a virtual substitute teacher, it is of particular importance to increase the proxy's awareness of what is going on in the classroom. In its current state of development, the proxy would not even notice if all students left the room. It is certainly one of our next steps to increase the proxy's awareness of the students' (mis)behavior in class, and implement a set of actions that the proxy can take in certain situations. For example, the classroom proxy may want to make sure that the students are taking notes and that no one sleeps while it is talking or showing a movie. "Smart classroom" technologies will help to enhance the virtual proxy's ability to take actions in the physical classroom. Such ubiquitous classroom technologies have been prototyped for automatic capture of class events and experience, as well as for context awareness and proactive services [31].

If in the future many teachers have a classroom proxy, then these personally trained proxies could be synchronized through a network. When a proxy is called to give a lecture on a certain topic, it may search for other proxies' experiences in teaching a similar class. Social tagging could be used in order to identify the best-rated lectures within this knowledge base. This scenario is not so far-fetched since teaching evaluations are already common practice these days. The proxy could then integrate parts of other teachers' scripts and resources into its own lecture. Such a joint knowledge base would also be useful as a resource for questions on the topic that might be asked by the students during the lecture and a selection of quality-rated answers for the proxy.

While substitute teachers generally have the unfavorable image of being "not the real" teacher and may be considered as less qualified than the regular teacher, proxies have the potential to become even better than the regular teachers themselves. However, it has yet to be evaluated whether the proxy holds its promise to be the ideal replacement for an absent teacher. Future evaluations will need to include comparisons between the performance and acceptance of a proxy versus a human substitute teacher.

References

[1] H. Augustin (1987). Substitute Teachers: An Endangered Species. *The Clearing House* **60** (1987), 393-396.

[2] K.C. Benedict, Student Expectations and the Substitute Teacher. *The Clearing House* **61** (1987), 27-28.

[3] C. Bruner, Technology and the Substitute Teacher. In: *Cases on Educational Technology Integration in Urban Schools*, IGI Global, Hershey PA (2012), 61-63.

[4] R. Clarke, The Digital Persona and its Application to Data Surveillance. *The Information Society* **10** (1994), 77-92.

[5] J. Ma, J. Wen et al., Cyber-Individual Meets Brain Informatics. *IEEE Intelligent Systems* **26** (2011), 30-37.

[6] S. Turkle, *The Second Self*, MIT Press, Cambridge MA, 1984.

[7] A. Clark, D. Chalmers, *The Extended Mind*, Analysis **58** (1998), 7-19.

[8] H. Prendinger, M. Ishizuka, Life-like Characters: Tools, Affective Functions, and Applications, Springer, Berlin, 2004.

[9] W.R. Swartout, J. Gratch et al., Toward Virtual Humans, *AI Magazine* **27** (2006), 96-108.

[10] J. Bates, The Role of Emotion in Believable Agents. *Communications of the ACM* **37** (1994), 122-125.

[11] C. Pelachaud, Multimodal Expressive Embodied Conversational Agents. *Proceedings of the 13th Annual ACM International Conference on Multimedia*, 2005.

[12] V. Vinayagamoorthy, M. Gillies et al., Building Expression into Virtual Characters. *Eurographics Conference State of the Art Reports*, 2006.

[13] J. Cassell, H. Vilhjálmsson, Fully Embodied Conversational Avatars: Making Communicative Behaviors Autonomous. *Autonomous Agents and Multi-Agent Systems* **2** (1999), 45-64.

[14] J. Cassell, J. Sullivan et al., *Embodied Conversational Agents*. MIT Press, Cambridge MA, 2000.

[15] S. Penny, J. Smith et al., Traces: Embodied Immersive Interaction with Semi-Autonomous Avatars. *Convergence* **7** (2001), 47-65.

[16] J.M. Bradshaw, P.J. Feltovich et al., Dimensions of Adjustable Autonomy and Mixed-Initiative Interaction. In: *Agents and Computational Autonomy*. Springer, Berlin (2004), 235-258.

[17] J. Allen, Mixed-initiative interaction, *IEEE Intelligent Systems* **14** (1999), 14-16.

[18] M. Soliman, C. Guetl, Intelligent Pedagogical Agents in Immersive Virtual Learning Environments: A Review. *MIPRO*, 2010.

[19] A.C. Graesser, P. Chipman et al., AutoTutor: An Intelligent Tutoring System with Mixed-initiative Dialogue, *IEEE Transactions on Education* **48** (2005), 612-618.

[20] G. Xuejing, W. Zhiliang et al., Design and Implementation of the Emotional Pedagogical Agent. *International Conference on Artificial Intelligence and Computational Intelligence*, 2009.

[21] C. Dowling, The Socially Interactive Pedagogical Agent within Online Learning Communities. *Proceedings of the International Conference on Computers in Education,* 2002.

[22] J.N. Bailenson, N. Yee et al., The Use of Immersive Virtual Reality in the Learning Sciences: Digital Transformations of Teachers, Students, and Social Context. *Journal of the Learning Sciences* **17** (2008), 102-141.

[23] J. Blascovich, A.C. Beall, Digital Immersive Virtual Environments and Instructional Computing. *Educational Psychology Review* **22** (2010), 57-69.

[24] J.N. Bailenson, A.C. Beall et al., Transformed Social Interaction, Augmented Gaze, and Social Influence in Immersive Virtual Environments. *Human Communication Research* **31** (2005), 511-537.

[25] B.S. Hasler, P. Tuchman, D.A. Friedman, Virtual Research Assistants: Replacing Human Interviewers by Automated Avatars in Virtual Worlds. *Computers in Human Behavior* **29** (2013), 1608-1616.

[26] D. Friedman, P. Tuchman, Virtual Clones: Data-driven Social Navigation. *Proceedings of the 11th International Conference on Intelligent Virtual Agents*, 2011.

[27] A. Steed, W. Steptoe et al., Beaming: An Asymmetric Telepresence System. *IEEE Computer Graphics and Applications* **32** (2012), 10-17.

[28] A. Kendon, Do Gestures Communicate? A Review. *Research on Language & Social Interaction* **27** (1994): 175-200.

[29] M. Looks, B. Goertzel et al., Novamente: An Integrative Architecture for General Intelligence. *AAAI Fall Symposium, Achieving Human-level Intelligence*, 2004.

[30] B.S. Hasler, O. Salomon et al., Real-time Gesture Translation in Intercultural Communication. *AI & Society* (in press).

[31] Y. Shi, W. Xie et al., The Smart Classroom: Merging Technologies for Seamless Tele-education. *IEEE Pervasive Computing* **2** (2003), 47-55.