An Analysis of Student Belief and Behavior in Learning by Explaining to a Digital Doppelganger

Ning Wang
Institute for Creative Technologies
University of Southern California
Playa Vista, CA
nwang@ict.usc.edu

Ari Shapiro
Institute for Creative Technologies
University of Southern California
Playa Vista, CA
shapiro@ict.usc.edu

Andrew Feng
Institute for Creative Technologies
University of Southern California
Playa Vista, CA
feng@ict.usc.edu

Cindy Zhuang
Institute for Creative Technologies
University of Southern California
Playa Vista, CA
czhuang@ict.usc.edu

David Schwartz
Department of Psychology
University of Southern California
Los Angeles, CA
davschw@usc.edu

Stephen L. Goldberg
STTC
US Army Research Laboratory
Orlando, FL
stephen.l.goldberg.civ@mail.mil

ABSTRACT

Digital doppelgangers are virtual humans that highly resemble the real self but behave independently. Using a low-cost and high-speed computer graphics and character animation technology, we created digital doppelgangers of students and placed them in a learning-by-explaining task where they interacted with digital doppelgangers of themselves. We investigate the research question of how does increasing the similarity of the physical appearance between the agent and the student impact learning. This paper discusses the design and evaluation of a digital doppelganger as a virtual human listener in a learning-by-explaining paradigm. It presents an analysis of how students' perceptions of the resemblance impact their learning experience and outcomes. The analysis and results offer insight into the promise and limitation of the application of this novel technology to pedagogical agents research.

CCS CONCEPTS

• **Human-centered computing** → *Empirical studies in HCI*; • **Applied computing** → *Interactive learning environments*;

KEYWORDS

pedagogical agent, learning by explaining, rapid avatar capture and simulation

1 INTRODUCTION

The power of intelligent tutoring systems lies in the personalized learning experience tailored to individual student's needs. Much of such personalization focuses on understanding a student's cognitive, affective, and motivational states, and the adaptation of the tutorial feedback to such states. The adaptation is not limited to just the content of the feedback but also its delivery, such as the agent that delivers the feedback. For example, researchers in pedagogical agents—embodied animated virtual characters designed to help students learn [21]—have experimented with many aspects of pedagogical agent design, including animation [24], gesture [13], voice [27], and social intelligence [40], and how such design can adapt to learner needs in order to maximize the efficacy of the personalized experience in facilitating student learning [31].

One aspect of such personalization that is less studied is the agent's appearance. Research on a pedagogical agent's appearance has indicated the impact of such design decisions on learning outcomes (e.g. [5], for review see [31]). However, research questions along the dimension of agent similarity with the learner are largely left unanswered. For example, when a pedagogical agent shares the exact appearance and behavioral characteristics of the learner, will such increased resemblance improve, make no difference on, or hamper learning? In the Teachable Agents [7] paradigm, for example, where students cultivate their protege by teaching a pedagogical agent, would an agent that is a digital doppelganger of a student make him/her more motivated to help the protege learn? To systematically study the impact of the similarity of the pedagogical agents' appearance to students requires the generation of such agents for a large enough population and at sufficient speed to accommodate experiment sessions of limited duration.

An emerging technology, the Rapid Avatar Capture and Simulation (RACAS) system [33], enables low-cost and high-speed scanning of a human user and creation of a fully animatable 3D "digital doppelganger" — a virtual human that highly resembles the real self but behaves independently [3]. This allows researchers in pedagogical agents to explore a previously unexplored research question: how does personalizing pedagogical agent's appearance to share physical similarity with the student impact learning? A previous study has suggested that there is limited impact of such personalization on learning outcomes in a learning-by-explaining paradigm [41]. In this paper, we discuss the design of a digital doppelganger as a virtual human listener and a follow-up analysis of student perception of personalization and the impact of such perception on student behavior and learning.

2 RELATED WORK

Pedagogical Agents: There is a large body of work on pedagogical agents across diverse learning domains. Most relevant to this work is the research on agent appearance. A series of studies conducted by Baylor and Kim examined the realism of agent appearance, its gender, and ethnicity. Results showed that agents with a realistic image improved transfer of learning, while the agent's gender and ethnicity contributed to its perceived competency, which in turn impacted student's self-efficacy [5]. Veletsianos and colleagues manipulated the agent's appearance relevant to stereotypes (artists vs. scientists) and found that such manipuations impacted the perceived knowledge level of the agent and students' performance on recall tasks [38]. Domagk studied the appealingness of the agent's appearance and uncovered the impact of this variable on students' transfer of learning [14]. In more recent work, Finkelstein and her colleagues extended prior studies of agent appearance by looking at not only the appearance of the agent relative to the learner (e.g., ethnicity), but also behaviors consistent with such appearance (e.g., the use of dialect) [16]. Results highlighted both the benefit (e.g., improved student learning) and potential drawback (e.g, decreased perceived agent competence) of the similarity between the agent and the learner. Indeed, personalizing the pedagogical agent to share the exact physical appearance with the student may be a multi-faceted issue. For example, interacting with an avatar looking like the real self have shown to induce anxiety in a public speaking task, compared to a dissimilar avatar [2]. In a learning paradigm such as learning-by-teaching, such anxiety may hinder student learning.

Learning by Explaining: We will study the impact of a digital doppelganger in the context of learning-by-explaining. Learning-by-explaining is an effective learning technique used by human tutors. In this technique, students are encouraged to explain a concept either to another or themselves. Decades of research shows that generating such explanations can lead to deep understanding of the learning material [10, 28], and that these learning effects are particularly strong when the explanations are delivered in a social context (i.e., explaining to a peer or tutor), as opposed to explaining to oneself [12, 43]. These effects have been observed when the "other" is a computer-generated character: For example, Cassell found that kids who told stories to a virtual partner developed better language skills [9]. Holmes found that students who explained to a software learning partner spent 50% more time asking questions and making explanations and generated significantly deeper explanations compared to students who worked without a software learning partner [20]. Leelawong and Biswas demonstrated that students benefitted from teaching a virtual peer [22]. We will leverage this effective learning paradigm as our testbed and use pedagogical agents, played by either a virtual human or a digital doppelganger, to create the social context to facilitate learning by explaining.

Digital Doppelgangers: Much effort in computer graphics and animation has been put forth to create digital doppengangers. Creating a virtual character from a particular subject is not a trivial task and usually requires extensive work from a 3D artist to model, rig, and animate the virtual character. To reconstruct a 3D model at low cost, researchers have experimented with low-cost 3D cameras (e.g., Kinect and Primesense) for 3D human shape acquisition. The work by [37] employs three Kinect devices and a turntable. Multiple shots are taken, and all frames are registered using the Embedded Deformation Model [35]. The work done in [45] utilizes two Kinect sensors in front of the self-turning subject. The subject stops at several key poses, and the captured frame is used to update the online model. Instead of multiple 3D sensors, the work in [42] used a fixed 3D sensor to acquire four key poses from a self-turning subject. Additionally, to create an animated virtual character from the scanned 3D human model, the 3D model needs to be rigged with a skeleton hierarchy and appropriate skinning weights. Traditionally, this process needs to be done manually and is time-consuming even for an experienced animator. An automatic skinning method was proposed in [4] that can produce reasonable results while reducing the manual efforts of rigging, but requires a connected and watertight mesh to work. The method proposed by [6] complements the previous work by automatically skinning a multi-component mesh. It works by detecting the boundaries between disconnected components to find potential joints. Such a method is suitable for rigging mechanical characters that consist of many components. Other rigging algorithms can include manual annotation to identify important structures such as the wrists, knees, and neck [26]. In recent years, video-based methods have enabled the capture and reconstruction of human motion as a sequence of 3D models that can preserve both the captured appearance and actor style, without the need of a rigging step [34]. Such methods, which are capable of reproducing surface and appearance details over time, have been used to synthesize animation by combining a set of mesh sequences [8]. However, current approaches are limited to basic locomotion such as walking, jogging, and jumping. Rigging of the video-based 3D models is still needed for complex movements.

3 DIGITAL DOPPELGANGERS WITH RACAS

The RACAS system is a virtual avatar generation system based on a 3D body scan. The RACAS system has two main capabilities: automatic rigging transfer and interactive avatar reshaping [15, 33]. The RACAS takes scans of a user from the front, back, left, and right sides using an RGB-D sensor. These scans are "stitched together" to create a 3D model. The 3D model is then enhanced by inferring a skeletal and muscular structure, as well as generating a model for the deformation of the skin and clothes. To do so, RACAS first utilizes SCAPE [1] to build a morphable human model from a 3D human model database. In order to allow pose deformations via linear blend skinning, RACAS researchers also manually rigged a template mesh from the database. Therefore given a 3D human body scan, RACAS can fit the morphable human model produced by SCAPE onto the input scan and establish mesh correspondences between them. Once RACAS establishes such correspondences, they can be used to transfer both skeleton and skin binding weights from the template mesh onto the input scan to generate a 3D virtual avatar. The users of RACAS can also interactively adjust semantic body attributes of the fitted model by exploring the body shape space generated from the database. Such body shape deformations can then be transferred to the aforementioned 3D scan to further create various virtual avatars with different body sizes and proportions. The resulting virtual avatars can then be animated in a simulation environment to execute various behaviors using animation retargeting. SmartBody, a character animation system, drives the animation of the 3D virtual character [32]. Using RACAS, researchers can easily create a digital doppelganger that serves as an ideal model for maximizing feelings of similarity, and customizing the virtual self's behavior to portray an optimal performance that the physical self cannot yet achieve.

4 EXPLAINING TO A DIGITAL DOPPELGANGER

We designed a virtual human listener and incorporated the digital doppelgangers created by RACAS to embody the virtual human listener. A human speaker, such as a student, can interact with the virtual human listener by talking to the it—for example, explaining what the student has learned. The virtual human listener provides conversational backchannel feedback as the student explains the concepts [19]. The feedback is generated based on analysis of the student's nonverbal behavior, such as head nods, prosody, etc. [19]. Previous research on virtual human listeners has shown the value of such feedback in creating rapport with the human speaker [19]. In the work discussed here, we focus on examining the impact of personalizing the agent's appearance on measures related to student learning. We hypothesize that teaching a virtual agent who looks just like oneself can impact a learner's motivation and self-regulation in learning (e.g. coping with challenges and persisting in a learning task). As a result of the increased the motivation, students may spend more effort to explain the learning concepts to the virtual human listener, resulting in better learning outcomes and higher self-efficacy. Specifically, we hypothesize that a virtual human listener that shares the appearance of the learner can improve learner motivation to teach the virtual human listener in a learning-by-explaining paradigm (**Hypothesis 1**). Additionally, such virtual human listeners can improve student learning of domain knowledge (**Hypothesis 2**) and improve student self-efficacy (**Hypothesis 3**).

5 EVALUATION

5.1 Design

In order to test our hypotheses on a pedagogical agent's appearance's impact on student learning, we conducted a study with the digital doppelganger serving as a virtual human listener in the task of learning-by-explaining. In this task, a student first studies a topic by himself/herself, such as reading a passage on the human circulatory system, then interacts with the agent by verbally explaining the topic to the agent. Regardless of appearance, the virtual human listener responded to the participants with the same behaviors (e.g., backchannel behaviors) described in Section 4. We designed a virtual human listener with two different appearances:

- **Digital Doppelganger condition:** A virtual human listener constructed at the beginning of each experiment session using RACAS, thus sharing the appearance of the participants, was used in this condition.
- **Virtual Human condition:** A virtual human with photo-realistic appearance, not based on any resemblance to the participant, was used in this condition. To control the realism of the virtual human listener used in both experiment conditions, the virtual human listener in this condition was generated using captures of non-participants obtained with RACAS through the same process used in the Digital Doppelganger condition. During the study, the virtual human listener was gender-matched to the participant, e.g., male participants interacted with a male virtual human. This is to match the obvious gender matching in the Digital Doppelganger condition. Figure 1 shows the female and male virtual human listeners used in this condition.

Figure 1: Virtual Human listeners, captured using RACAS, from the control condition.

5.2 Population

We recruited 41 participants from the student population of a university located in North America. Participants were recruited either from the Psychology Department subject pool or via fliers posted on campus. Participants recruited from the subject pool received course credit for their participation, while participants recruited through the fliers received $10 for their participation.

5.3 Procedure

Participants first read and signed an informed consent about the study. Then the experimenter completed two scans of the participants from both conditions. The first was a full-body scan using an iPad equipped with a specialized structure sensor. A second scan was a face-only scan, captured using an Intel webcam with depth sensors. Both scans were conducted in a lab fitted with wall-to-ceiling diffused LED lights. Next, the participants filled out a Background Survey and took a Pre-Test about the domain knowledge, while the RACAS system completed the generation of the 3D model of the participants. Then the participants read an essay on the human circulatory system, adopted from previous studies on expert tutoring [11], on a web browser. The participants were told that they would later have to teach the material to a virtual student. After that, the participants sat in front of a 30-inch computer monitor with the display of the virtual human listener. Two cameras were fitted on top of the monitor: one recorded the participants' face, and the other served as input to the virtual human listener. The participants were told that the virtual student (i.e., the virtual human listener) would represent him/her in a competition against other virtual students in a quiz on the same subject. Participants then explained what they had learned from the tutorial to the virtual human listener. Finally, the participants filled out a Post-Interaction Survey and took a Post-Test on the human circulatory system. Each study session was designed to last one hour.

5.4 Measures

The Background Survey included measures of demographic information, education, Rosenberg Self-Esteem Scale [30], Adolescent Body Image Satisfaction Scale (ABISS) [23], Anxiety scale [29], and Self-Efficacy in domain knowledge (designed by the research team). The Self-Efficacy scale included items such as "I am confident that I know the path blood takes through the circulatory system of the human body" and "If there is a quiz on human circulatory system, I expect to do well on the quiz."

The Post-Interaction Survey included measures of Presence (constructed using items from [44] and [17]), Avatar Similarity ("To what extend do you feel that the virtual avatar resembled you?"), Desired Avatar Similarity ("If you had to design your own avatar for this task, how similar to your real appearance would you make your avatar?"), a repeated measure of Self-Efficacy in domain knowledge, and Self-Efficacy in the virtual student ("I think the avatar I just taught will do well in the competition.").

In the Pre-Test, participants were asked to describe 10 concepts on the human circulatory system (e.g., the aorta and capillaries) and the path of blood through the body. The Post-Test included the same questions from the Pre-Test. Scores on these questions were termed Post-Test-Repeat scores. In addition, the Post-Test included questions adopted from tests used in previous studies on human tutoring [11]. Scores on these questions were termed Post-Test-NonRepeat scores.

In addition to the surveys and tests, the agent's behaviors in response to the participants were logged. The explanations generated by the participants were recorded.

6 RESULTS

Data from one participant was excluded due to extremely short participation time (in filling out surveys and tests, reading the essay, and interacting with the agent). As a result, data from 40 participants (25 female and 15 male, $M_{age} = 21.5$, age range: 19.7-25.7 years) are included in the analysis. The participants came from a variety of majors, ranging from psychology to fine arts, to biology, and many more. One participant had a graduate degree, while all other participants had some college education. Participants were randomly assigned to one of the conditions in the study. While a balanced assignment was desired, the result was that 16 participants were assigned to the Digital Doppelganger condition and the remaining 24 participants were assigned to the Virtual Human condition.

6.1 Hypothesis Testing

We conducted an analysis of variance test with scores on Pre-Test and Post-Test-Repeat as a repeated measure and with the experiment condition (Virtual Human vs. Digital Doppelganger) as the Between-Subject factor. We also conducted an analysis of variance test with self-efficacy reported before and after the study as the repeated measure and experiment condition as the Between-Subject factor. Finally, we conducted an Independent Sample T-Test on the duration of participants' explanations between experiment conditions. Duration of the explanation is used in the analysis as an indication of the participants' motivation to teach the virtual human listener. Results show that although the within-subject effect before and after the study on test scores and self-efficacy are statistically significant (improved after compared to before), we did not find any significant differences between experiment conditions on these two measures. We also did not find any significant difference on the duration of explanation between the two experiment conditions. Details of the analysis on hypothesis testing are discussed in [41]. This result suggests that Hypotheses 1, 2, and 3 are not supported.

6.2 Alternative Hypotheses

Because the results of the analysis suggest that there is no statistically significant difference between the two experiment conditions, we conducted further analysis of the self-report and behavioral measures to examine why that was the case.

6.2.1 Number of Words in the Explanations. Upon close examination of the audio recordings of the participants' explanations, we noticed that some participants spoke few words before quickly giving up teaching the virtual human listener, while others spoke at greater speed (and with greater confidence) so that they managed to cover all the material they learned in the same duration. Thus the duration of the explanation make not paint the whole picture about the participants' motivation. To further compare the motivation to teach the virtual human listener, we analyzed the length of the participants' explanation in total number of words spoken, instead of the duration. We transcribed the participants' explanations using the Temi automated online transcription service [36]. A member of the research team then manually inspected the transcription and corrected any errors in the automatically generated transcription. Independent Sample T-Tests show that there is no significant difference between the two experiment conditions for the number of words ($p = .090, M_{VH} = 504, M_{DD} = 360$). This result again suggests that Hypothesis 1 regarding the similarity of agent appearance and motivation to learn (and to explain and teach) is not supported.

6.2.2 Novelty Effect. When comparing the duration of the participants' explanations during the hypothesis testing, we noticed that participants, by and large, interacted with the agent for a very short period of time (from just under a minute to over 10 minutes). RACAS is a novel technology. Participants, especially the ones in the Digital Doppelganger condition who had never seen themselves transformed into a digital character before, may have directed much of their attention to visually inspecting their own avatar. Such activity may have distracted the participants from the learning activity, both recalling and explaining. Behaviorally, these participants may experience speech disfluencies [19], such as the frequent use of filler words. In our previous studies with a virtual human listener in a story-telling task, when the virtual human listener's behavior interrupted the "flow" of the human participant's story-telling, such as non-stop staring or ill-timed back-channel feedback, the human participants spoke with more disfluencies [19, 39]. We annotated the explanations generated by the participants, particularly speech disfluencies, such as "um", "uh". We did not find any significant difference between the two conditions on both the number of disfluencies ($p = .394, M_{VH} = 18.6, M_{DD} = 15.2$) and frequency of disfluencies ($p = .362, M_{VH} = 4.0, M_{DD} = 4.5$, per minute). This may be due to the fact that participants from both conditions spent an equal amount of mental effort to examine the virtual human listener. Because participants from both conditions were scanned, even the participants from the non-digital doppelganger condition may have initially devoted much mental effort to examining the visual similarities between the virtual human listener and themselves.

6.2.3 Believed Similarity with the Agent. Because the result on the main hypothesis testing indicates that the two experiment groups were very similar on the dependent measures, we wondered how the manipulation of the independent variable was perceived by the participants. We compared the perceived resemblance of the virtual human listener to the participants. We expect the Avatar Similarity to be much lower in the Virtual Human condition, compared to the Digital Doppelganger condition. Independent-sample T-Test shows that it is indeed the case ($p = .001, M_{VH} = 1.67, M_{DD} = 3.94$). Participants from the Virtual Human condition did not perceive the virtual human listener's appearance to be similar to themselves. However, participants from the Digital Doppelganger condition were also somewhat unsure whether the virtual human listener looked like them (rated 3.94 on a 7-point Likert scale), even though the virtual human listeners were created using photo-realistic captures of these participants.

We examined the participants' free-form answers to a question regarding their reaction to seeing the virtual representation of themselves on the Post-Survey. While comments from participants in the Digital Doppelganger condition are consistent with the experiment manipulation, comments from participants from the Virtual Human condition reveal that some of the participants thought the virtual human listener looked like them and was a digital representation of themselves. It is plausible that those participants saw enough resemblance between the gender-matched generic virtual human's face (see Fig. 1) and their own, because participants from all conditions received a face and full-body scan using the RACAS system. In addition, these participants may have thought that the body scan only captured the shape of one's body, and not the appearance (e.g., clothing), and saw enough resemblance between the gender-matched generic virtual human's body shape and their own. Regardless, these participants from the Virtual Human condition believed that they saw, explained to, and taught their own avatar. Based on this finding, we regrouped the participants based on the condition they *believed* they were assigned to. Four participants from the Virtual Human condition are now part of the Digital Doppelganger condition.

Once again, we conducted the tests for the main hypothesis and found no significant main effect of the *believed condition* on changes of pre- and post- test scores (i.e., scores on the questions repeated from pre- and post test) and the self-efficacy. We saw marginally significant effects on the Post-Test-NonRepeat scores (i.e., scores on questions that appeared only on the post-test, $p = .065, M_{VH} = 29.75, M_{DD} = 36.15$) and self-efficacy in the virtual student ("I think the avatar I just taught will do well in the competition", $p = .071, M_{VH} = 2.80, M_{DD} = 3.9$). This suggests a trend that, when the virtual human listener shares enough resemblance with the participant for them to *believe* that it is a virtual representation of themselves, they performed better on the near transfer Post-test, compared to those who did not have such belief. Additionally, they had higher confidence that their own avatar would perform well in a quiz competition with other avatars.

7 DISCUSSION

In this paper, we discussed the design of a pedagogical agent for the learning-by-explaining paradigm. We applied a novel character animation technology, RACAS, to create virtual human listeners that share the physical appearance of a human learner. Evaluation of such virtual human listener showed that the digital doppelganger did not significantly impact student learning of domain knowledge, their motivation to teach the agent, or their own self-efficacy [41]. We subsequently hypothesized that a novelty effect may be at play, e.g., participants be distracted by the virtual human listener in during the initial interactions. Post Hoc analysis of speech disfluency of the participants' explanations, a possible behavioral indication of such distraction, shows that there were no significant differences in the disfluencies during the explanations. Because participants from all conditions were scanned by the RACAS system, they might have all expected to see their own avatar as the virtual human listener and thus devoted equal amount of attention to examining the virtual human listener during the explanation process.

This also led to our second alternative hypothesis that if the participants perceive enough physical resemblance between the virtual human listener, they may *believe* that they were interacting with their own digital doppelganger. This is the case for several of our participants who interacted with a non-doppelganger virtual human. Analysis of the main hypothesis based on this *perceived* experiment manipulation suggests that when the participants explained to a virtual human listener that they believed looked like them, they performed better on the near-transfer post-test. They also had higher confidence that their perceived digital doppelganger would perform well in a future quiz competition with other avatars. However, both effects only approach statistical significance ($p = .65$ and .071 respectively). So the results suggest a possible trend that personalizing a pedagogical agent's appearance to be similar to the student's physical appearance may play a role in the efficacy of pedagogical agents. One participant commented that "It (the digital doppelganger) made me want to teach him the material so that he could score well on the test. " Such comments resonate with our initial hypothesis and the results here.

The RACAS system is a new technology that still requires time to mature. The pedagogical agents created through such a process are less than perfect. Even very slight glitches in the virtual agent's appearance, such as misalignment of face and body, or animation, such as a slight shift of the face when the eyes open/close, can greatly distract the learner and interfere with the engagement of the learning task. Comments by the participants indicate that they certainly took notice of such imperfections. Other studies on digital doppelgangers created using RACAS also found that having an avatar that looks like the participants improves subjective experience, but made no difference on performance measures (e.g., running in a virtual maze with mines) [25]. On the other hand, earlier studies on non-animated, hand-craftied, and high-fidility digital doppelgangers have shown great promise on attitude and behavioral change (e.g., [18]). Further investigations on the potential of the personalization pf the appearance of pedagogical agents can be carried out once the computer graphics and animations technologies, such as RACAS, mature.

ACKNOWLEDGMENTS

This project is funded by the U.S. Army Research Laboratory. Statements and opinions expressed do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

REFERENCES

[1] Dragomir Anguelov, Praveen Srinivasan, Daphne Koller, Sebastian Thrun, Jim Rodgers, and James Davis. 2005. SCAPE: shape completion and animation of people. In *ACM transactions on graphics*, Vol. 24. 408–416.

[2] Laura Aymerich-Franch, René F Kizilcec, and Jeremy N Bailenson. 2014. The relationship between virtual self similarity and social anxiety. *Frontiers in human neuroscience* 8 (2014), 944.

[3] Jeremy N Bailenson. 2012. Doppelgangers-a new form of self? *The Psychologist* 25, 1 (2012), 36–39.

[4] Ilya Baran and Jovan Popović. 2007. Automatic rigging and animation of 3d characters. *ACM Transactions on Graphics (TOG)* 26, 3 (2007), 72.

[5] Amy L Baylor and Yanghee Kim. 2004. Pedagogical agent design: The impact of agent realism, gender, ethnicity, and instructional role. In *Proceedings of the 7th International Conference on Intelligent Tutoring Systems*. Springer, 592–603.

[6] Gaurav Bharaj, Thorsten Thormählen, Hans-Peter Seidel, and Christian Theobalt. 2012. Automatically rigging multi-component characters. In *Computer Graphics Forum*, Vol. 31. Wiley Online Library, 755–764.

[7] Gautam Biswas, Hogyeong Jeong, John S Kinnebrew, Brian Sulcer, and ROD ROSCOE. 2010. Measuring self-regulated learning skills through social interactions in a teachable agent environment. *Research and Practice in Technology Enhanced Learning* 5, 02 (2010), 123–152.

[8] Dan Casas, Marco Volino, John Collomosse, and Adrian Hilton. 2014. 4D video textures for interactive character appearance. In *Computer Graphics Forum*, Vol. 33. Wiley Online Library, 371–380.

[9] Justine Cassell. 2004. Towards a model of technology and literacy development: Story listening systems. *Journal of Applied Developmental Psychology* 25, 1 (2004), 75–105.

[10] Michelene TH Chi and Miriam Bassok. 1989. Learning from examples via self-explanations. *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (1989), 251–282.

[11] Michelene TH Chi, Stephanie A Siler, Heisawn Jeong, Takashi Yamauchi, and Robert G Hausmann. 2001. Learning from human tutoring. *Cognitive Science* 25, 4 (2001), 471–533.

[12] National Research Council et al. 2000. *How people learn: Brain, mind, experience, and school: Expanded edition*. National Academies Press.

[13] Scotty D Craig, Barry Gholson, and David M Driscoll. 2002. Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features and redundancy. *Journal of educational psychology* 94, 2 (2002), 428.

[14] Steffi Domagk. 2010. Do pedagogical agents facilitate learner motivation and learning outcomes? *Journal of media Psychology* (2010).

[15] Andrew Feng, Dan Casas, and Ari Shapiro. 2015. Avatar reshaping and automatic rigging using a deformable model. In *Motion in Games*. ACM, 57–64.

[16] Samantha Finkelstein, Evelyn Yarzebinski, Callie Vaughn, Amy Ogan, and Justine Cassell. 2013. The effects of culturally congruent educational technologies on student achievement. In *Proceedings of the 16th International Conference on Artificial Intelligence in Education*. Springer, 493–502.

[17] Jesse Fox and Jeremy N Bailenson. 2009. Virtual self-modeling: The effects of vicarious reinforcement and identification on exercise behaviors. *Media Psychology* 12, 1 (2009), 1–25.

[18] Jesse Fox and Jeremy N Bailenson. 2010. The use of doppelgängers to promote health behavior change. *CyberTherapy & Rehabilitation* 3, 2 (2010), 16–17.

[19] Jonathan Gratch, Ning Wang, Anna Okhmatovskaia, Francois Lamothe, Mathieu Morales, Rick J van der Werf, and Louis-Philippe Morency. 2007. Can virtual humans be more engaging than real ones?. In *Proceedings of the International Conference on Human-Computer Interaction*. Springer, 286–297.

[20] Jeffrey Thomas Grant Holmes. 2003. *Learning by explaining: the effects of software agents as learning partners.* Ph.D. Dissertation. University of California, Berkeley.

[21] W Lewis Johnson, Jeff W Rickel, James C Lester, et al. 2000. Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial intelligence in education* 11, 1 (2000), 47–78.

[22] Krittaya Leelawong and Gautam Biswas. 2008. Designing learning by teaching agents: The Betty's Brain system. *International Journal of Artificial Intelligence in Education* 18, 3 (2008), 181–208.

[23] James E Leone, Elizabeth M Mullin, Suanne S Maurer-Starks, and Michael J Rovito. 2014. The adolescent body image satisfaction scale for males: exploratory factor analysis and implications for strength and conditioning professionals. *The Journal of Strength & Conditioning Research* 28, 9 (2014), 2657–2668.

[24] James C Lester and Brian A Stone. 1997. Increasing believability in animated pedagogical agents. In *Proceedings of the first international conference on Autonomous agents*. ACM, 16–21.

[25] Gale Lucas, Evan Szablowski, Jonathan Gratch, Andrew Feng, Tiffany Huang, Jill Boberg, and Ari Shapiro. 2016. The effect of operating a virtual doppleganger in a 3D simulation. In *Proceedings of the 9th International Conference on Motion in Games*. ACM Press, Burlingame, CA, 167–174.

[26] Mixamo. 2013. *Mixamo auto-rigger.* http://www.mixamo.com/c/auto-rigger

[27] Natalie K Person. 2003. AutoTutor improves deep learning of computer literacy: Is it the dialog or the talking head? *Artificial intelligence in education: Shaping the future of learning through intelligent technologies* 97 (2003), 47.

[28] Peter Pirolli and Margaret Recker. 1994. Learning strategies and transfer in the domain of programming. *Cognition and instruction* 12, 3 (1994), 235–275.

[29] International Personality Item Pool. 2015. *The Items in the Preliminary IPIP Scales Measuring Constructs Similar to Those Included in Lee and Ashton's HEXACO Personality Inventory.* http://ipip.ori.org/newHEXACO_PI_key.htm#Anxiety

[30] Morris Rosenberg. 2015. *Society and the adolescent self-image.* Princeton university press.

[31] Noah L Schroeder, Olusola O Adesope, and Rachel Barouch Gilbert. 2013. How effective are pedagogical agents for learning? A meta-analytic review. *Journal of Educational Computing Research* 49, 1 (2013), 1–39.

[32] Ari Shapiro. 2011. Building a character animation system. *Motion in Games* (2011), 98–109.

[33] Ari Shapiro, Andrew Feng, Ruizhe Wang, Hao Li, Mark Bolas, Gerard Medioni, and Evan Suma. 2014. Rapid avatar capture and simulation using commodity depth sensors. *Computer Animation and Virtual Worlds* 25, 3-4 (2014), 201–211.

[34] Jonathan Starck and Adrian Hilton. 2007. Surface capture for performance-based animation. *IEEE computer graphics and applications* 27, 3 (2007).

[35] Robert W Sumner, Johannes Schmid, and Mark Pauly. 2007. Embedded deformation for shape manipulation. In *ACM Transactions on Graphics*, Vol. 26. 80.

[36] Temi. [n. d.]. Audio to Text Transcription Service: Temi.

[37] Jing Tong, Jin Zhou, Ligang Liu, Zhigeng Pan, and Hao Yan. 2012. Scanning 3d full human bodies using kinects. *IEEE transactions on visualization and computer graphics* 18, 4 (2012), 643–650.

[38] George Veletsianos. 2010. Contextually relevant pedagogical agents: Visual appearance, stereotypes, and first impressions and their impact on learning. *Computers & Education* 55, 2 (2010), 576–585.

[39] Ning Wang and Jonathan Gratch. 2010. Don't just stare at me!. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1241–1250.

[40] Ning Wang, W Lewis Johnson, Richard E Mayer, Paola Rizzo, Erin Shaw, and Heather Collins. 2008. The politeness effect: Pedagogical agents and learning outcomes. *International Journal of Human-Computer Studies* 66, 2 (2008), 98–112.

[41] Ning Wang, Ari Shapiro, Andrew Feng, Cindy Zhuang, Chirag Merchant, David Schwartz, and Stephen L Goldberg. 2018. Learning by Explaining to a Digital Doppelganger. In *Proceedings of the 14th International Conference on Intelligent Tutoring Systems*. Springer.

[42] Ruizhe Wang, Jongmoo Choi, and Gerard Medioni. 2012. Accurate full body scanning from a single fixed 3d camera. In *3D Imaging, Modeling, Processing, Visualization and Transmission*. IEEE, 432–439.

[43] Noreen M Webb. 1989. Peer interaction and learning in small groups. *International journal of Educational research* 13, 1 (1989), 21–39.

[44] Bob G Witmer, Christian J Jerome, and Michael J Singer. 2005. The factor structure of the presence questionnaire. *Presence: Teleoperators and Virtual Environments* 14, 3 (2005), 298–312.

[45] Ming Zeng, Jiaxiang Zheng, Xuan Cheng, and Xinguo Liu. 2013. Templateless quasi-rigid shape modeling with implicit loop-closure. In *Computer Vision and Pattern Recognition*. IEEE, 145–152.