Learning by Explaining to a Digital Doppelganger

Ning Wang[1], Ari Shapiro[1], Andrew Feng[1], Cindy Zhuang[2], Chirag Merchant[1], David Schwartz[2], Stephen L. Goldberg[3]

[1]Institute for Creative Technologies, University of Southern California
[2]Department of Psychology, University of Southern California
[3]U.S. Army Research Laboratory
[1]nwang, shapiro, feng, merchant@ict.usc.edu
[2]cindyi.0324@gmail.com, davschw@dornsife.usc.edu
[3]stephen.l.goldberg.civ@mail.mil,

Abstract. Digital doppelgangers are virtual humans that highly resemble the real self but behave independently. An emerging computer animation technology makes the creation of digital doppelgangers an accessible reality. This allows researchers in pedagogical agents to explore previously unexplorable research questions, such as how does increasing the similarity in appearance between the agent and the student impact learning. This paper discusses the design and evaluation of a digital doppelganger as a virtual listener in a learning-by-explaining paradigm. Results offer insight into the promise and limitation of this novel technology.

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

1 Introduction

Pedagogical agents are embodied animated virtual characters designed to help students learn [1]. Over the past two decades, since Herman the Bug [2] and Steve [3], researchers have studied many aspects of pedagogical agents, including animation [4], gesture [5], voice [6], and social intelligence [7], and role [8], to facilitate student learning across a great number of domains. As embodied virtual characters, one of the first decisions pedagogical agent designers have to make is what the agent looks like. Research on a pedagogical agent's appearance has indicated the impact of such design decisions on learning outcome [9], including recall [10] and transfer of learning [11] (for review see [12]). Research into appearance similarity between the agent and the learner mainly focused on ethnicity and behaviors consistent with such appearance (e.g., the use of dialect) [13]. Research questions further along the dimension of agent similarity with the learner have been left largely unanswered because of the need to generate 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, enables low-cost and high-speed scanning of a human user and creation of a fully animatable virtual 3D "digital double" of the user. This allows researchers to explore a previously unexplored research question: how does increasing the similarity in appearance between the agent and student impact learning. In this paper, we discuss the design of a digital doppelganger as a virtual listener and the evaluation of such an agent in a learning-by-explaining paradigm.

2 Explaining to a Digital Doppelganger

Digital doppelgangers are virtual humans that highly resemble the real self but behave independently [14]. The RACAS system, described in detail in [15], makes the digital doppelganger a more accessible reality. We designed a virtual listener and incorporated digital doppelgangers created by RACAS to embody the listener. A human speaker can converse with the agent and the agent can respond with conversational backchannel feedback [16]. The feedback is generated based on analysis of the speaker's nonverbal behavior, such as head nods, prosody, etc. [16]. Previous research has shown the value of such feedback in creating rapport with the human speaker [16]. The current work focuses specifically on examining the impact of agent appearance on measures related to student learning. We hypothesize that teaching a virtual listener who looks just like oneself can impact a learner's motivation and self-regulation in learning (e.g. persisting in a learning task), and ultimately improve learning outcomes. Specifically, we hypothesize that, in a learning-by-explaining paradigm:

H1: A virtual listener that shares the appearance of the learner can improve learner motivation to teach the agent.
H2: A virtual listener that shares the appearance of the learner can improve student learning of domain knowledge through teaching the virtual agent.
H3: A virtual listener that shares the appearance of the learner can improve student self-efficacy through teaching the virtual agent.

3 Evaluation

Design We conducted a study with the digital doppelganger serving as a virtual listener in the task of learning-by-explaining. In this task, a student first reads a passage on the human circulatory system, then verbally explains the topic to the virtual listener. The study is a between-subject design with two experiment conditions: the Digital Doppelganger condition and the virtual human condition.

- **Digital Doppelganger** In this condition, a virtual listener was constructed at the beginning of each experiment session using RACAS, thus sharing the appearance of the participants.
- **Virtual Human** In this condition, a virtual listener with a photo-realistic appearance not based on the participant was used. To control the realism of the virtual listener used in both conditions, the agent in this condition was

generated using captures of non-participants obtained with RACAS through the same process used in the other condition (Figure 1). The virtual listener was gender-matched to the participant, e.g., male participants interacted with a male virtual human. Aside from the difference in appearance, both virtual listeners responded to the participants with the same behaviors, described in Section 2.

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

Population and Procedure We recruited 41 student either from the Psychology Department subject pool (received course credit) or via fliers posted on campus (received $10) at the University of Southern California. Participants first read an informed consent. Then the experimenter completed face and body scans of the participants, in both conditions. The full-body scan was captured with an iPad equipped with a specialized structure sensor. The face scan was captured using an Intel webcam with depth sensors. Next, the participants filled out a Background Survey and Pre-Test, then read a tutorial on the human circulatory system (adopted from [17]) on a web browser. The participants were told that they would later have to teach the material to a virtual student. Then, the participants sat in front of a 30-inch computer monitor with the display of the virtual student, and were told that the virtual student would represent him/her in a competition against other virtual students in a quiz on the same subject. Two cameras were fitted on top of the monitor: one recorded the participants' face, and the other served as input to the virtual listener. Participants then verbally explained what they had learned from the tutorial to the virtual listener. Finally, the participants filled out a Post-Interaction Survey and Post-Test. Each session was designed to last one hour.

Measures The Background Survey included measures of demographic, education, Rosenberg Self-Esteem Scale [18], Adolescent Body Image Satisfaction Scale (ABISS) [19], Anxiety scale [20], and Self-Efficacy in domain knowledge (designed by the research team). The Self-Efficacy scale included items such as "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 [21] and [22]), 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 described 10 concepts on the human circulatory system and the path of blood through the body. The Post-Test included the Pre-Test questions and questions adopted from previous studies on human tutoring [17].

4 Results

Data from all 41 participants (26 female, 15 male, $M_{age} = 21.5$, age range: 19.7-29.7 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 an experiment condition. While a balanced assignment was desired, in the end, 17 participants were assigned to the Digital Doppelganger condition and 24 to the Virtual Human.

Learning Domain Knowledge An expert on the human circulatory system from the research team graded the Pre- and Post-Tests. On the Post-Test, we separated the score on questions that were repeated from the Pre-Test (Post-Test-Repeat) and scores on the rest of the questions (Post-Test-NonRepeat). We conducted an ANOVA with scores on Pre-Test and Post-Test-Repeat as a repeated measure and the experiment conditions as the Between-Subject factor. The result shows that there was a significant within-subject effect between Pre- and Post-Tests ($p < .001, F = 91.404$), while the between-subject effect due to the experiment manipulation was not statistically significant ($p = .308, F = 1.069$, see Fig. 2 for means). Although there is a noticeable difference on Pre-Test scores between the two experiment conditions, the difference is not statistically significant ($p = .308$). We also conducted an Independent Sample T-Test on the scores on Post-Test-NonRepeat and found no significant difference ($p = .821, M_{VH} = 32.33, M_{DD} = 33.12$, 62 total points available). This suggests that Hypothesis 2 regarding agent appearance and learning of domain knowledge is not supported.

Self-Efficacy We conducted an ANOVA with self-efficacy before and after the study as the repeated measure and experiment condition as the Between-Subject factor. The result shows that there was a significant within-subject effect before and after the study ($p = .003, F = 9.78$), while the between-subject effect due to the experiment manipulation was not statistically significant ($p = .891, F = .019$, see Fig. 2 for means). Additionally, we analyzed the participants' self-efficacy in

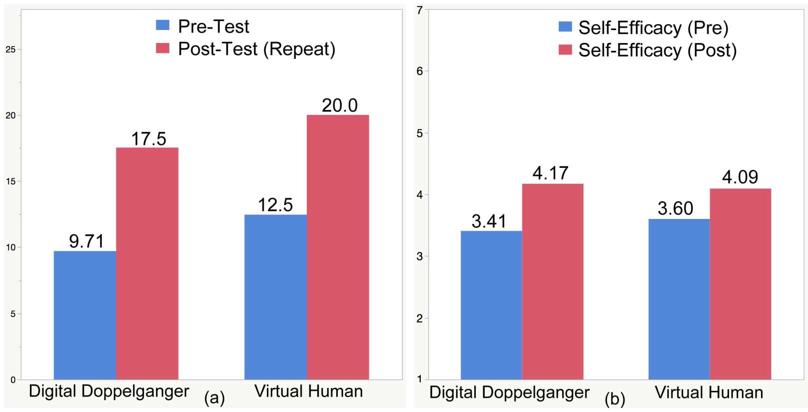

Fig. 2. (a) Comparison of Pre-Test and Post-Test-Repeat scores (maximum score was 28) (b) Comparison of Self-Efficacy (7-point Likert scale) before and after study between experiment conditions.

the virtual listener, whom they taught and thought would represent them to compete with other agents. Again, we did not find any significant difference between the two experiment conditions ($p = .561, M_{VH} = 3.17, M_{DD} = 3.53$). This result suggests that Hypothesis 3, regarding the similarity of agent appearance and learner's self-efficacy, is not supported.

Motivation to Teach the Virtual Listener We analyzed the time participants spent explaining the material to the virtual listener. An Independent Sample T-Test shows that there is no significant difference between the two conditions ($p = .105, M_{VH} = 277.88, M_{DD} = 208.63, Min = 55, Max = 645$, in seconds). This result suggests that Hypothesis 1 regarding the similarity of agent appearance and motivation to learn (and to explain and teach) is not supported.

Further Analysis Because the results suggest that there is no statistically significant difference between the two experiment conditions, we conducted further analyses to examine why that was the case. We first performed a "manipulation check" on the Avatar Similarity scale. We expected 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 ($p = .004$). Fig. 3 shows that participants from the Virtual Human condition did not perceive the agent's appearance to be similar to themselves. However, participants from the Digital Doppelganger condition did not think the virtual listener looked like them either (rated 3.76 on a 7-point Likert scale). Furthermore, we compared the Desired Avatar Similarity. The difference, as shown in Fig. 3, is marginally significant ($p = .077$). In particular, participants in the Virtual Human condition, who did not see their digital doppelganger, wished the virtual listener would look like them. Conversely, participants in the Digital

Doppelganger condition, after seeing their own image manifested as an animated character, reported that they would rather the virtual listener *not* look like them.

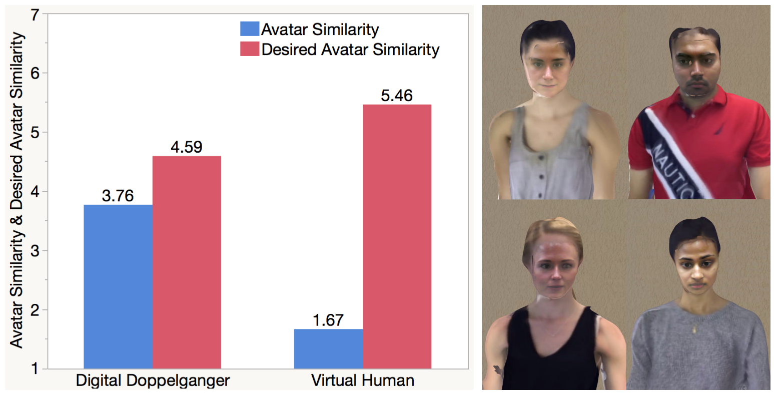

Fig. 3. Left: Comparison of the perceived similarity of the virtual listener's appearance to the participant's and the participant's desired level of such resemblance (7-point Likert scale). **Right:** Digital Doppelgangers that had pronounced imperfections, such as lighting, face mis-alignment, and missing pixels.

We then conducted pair-wise correlation tests of these two variables and the dependent variables we tested for the main hypothesis. The Desired Avatar Similarity is positively correlated with post-interaction Self-Efficacy ($r = .334, p = .033$), but not with the other dependent variables. This indicates that participants who were more confident in their domain knowledge had a higher desire for the virtual student to share their appearance. This resonates with the results on general self-confidence and confidence in one's appearance: the Desired Avatar Similarity is positively correlated with the Rosenburg Self-esteem measure ($r = .399, p = .01$) and the self-image measure — ABISS ($r = .436, p = .004$). The perceived Avatar Similarity, on the other hand, is positively correlated with the post-interaction Self-Efficacy in the agent ($r = .359, p = .021$), but not with the other dependent variables. This indicates that the more the participants perceived the agent to resemble themselves, the more confident they felt about how well the agent, whom they taught, would do in competitions and quizzes.

5 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 agents that share the physical appearance of a human learner. Evaluation of such agents showed that such resemblance did not significantly impact student learning of domain knowledge, their motivation to teach the agent, or their own self-efficacy. Further analysis indicates that when students are confident about their knowledge, they would like the agent to look like them. And the more the agent shared their appearance, the more confident they felt about the agent's future performance, as a result of their teaching. While the investigation did not yield a statistically significant result, it is worth noting that this is the first investigation of its kind. The process to scan, reconstruct, and animate a virtual agent, particularly one with an animatable face, in such rapid fashion has rarely been attempted before. The pedagogical agents created through such process are understandably less than perfect (see Fig. 3). Even very slight glitches in the virtual agent's appearance (e.g., misalignment of face and body) or animation (e.g., slight shift of the face when the eyes open/close) can distract the learner and interfere with engagement in the learning task.

The interaction with the digital doppelganger is short. Thus a novelty effect may have played a role in the study. 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, again, may have distracted the participants from the learning activity, both the recalling and the explaining. The distractions may have ultimately impacted the learning outcome. Future studies can allow learners to interact with their own avatar for longer periods of time, beyond the initial influence of the novelty effect. Additionally, previous studies on virtual listener agents have identified behavioral indications of when participants were distracted by the agent's behavior, e.g., speech disfluencies and gaze aversions. Linguistic and video analyses can be carried out on the participants' explanations and videos of their face to test this hypothesis on distraction. Since the study concluded, great improvements have already been made to RACAS that allow even faster capture of higher fidelity and more accurate 3D scans [23], all of which provide great promise for future studies on the appearance of pedagogical agents.

Acknowledgment This work is supported by the US Army Research Laboratory. Statements expressed do not necessarily reflect the position or the policy of the United States Government. No official endorsement should be inferred.

References

1. W. L. Johnson, J. W. Rickel, J. C. Lester *et al.*, "Animated pedagogical agents: Face-to-face interaction in interactive learning environments," *International Journal of Artificial intelligence in education*, vol. 11, no. 1, pp. 47–78, 2000.
2. J. C. Lester, S. A. Converse, S. E. Kahler, S. T. Barlow, B. A. Stone, and R. S. Bhogal, "The persona effect: affective impact of animated pedagogical agents," in *Proceedings of Human factors in computing systems.* ACM, 1997, pp. 359–366.
3. W. L. Johnson and J. Rickel, "Steve: An animated pedagogical agent for procedural training in virtual environments," *SIGART*, vol. 8, no. 1-4, pp. 16–21, 1997.
4. J. C. Lester and B. A. Stone, "Increasing believability in animated pedagogical agents," in *Proceedings of Autonomous agents.* ACM, 1997, pp. 16–21.
5. S. D. Craig, B. Gholson, and D. M. Driscoll, "Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features and redundancy." *Journal of educational psychology*, vol. 94, no. 2, p. 428, 2002.

6. N. K. Person, “Autotutor improves deep learning of computer literacy: Is it the dialog or the talking head?” *AI in education*, vol. 97, p. 47, 2003.
7. N. Wang, W. L. Johnson, R. E. Mayer, P. Rizzo, E. Shaw, and H. Collins, “The politeness effect: Pedagogical agents and learning outcomes,” *International Journal of Human-Computer Studies*, vol. 66, no. 2, pp. 98–112, 2008.
8. G. Biswas, H. Jeong, J. S. Kinnebrew, B. Sulcer, and R. ROSCOE, “Measuring self-regulated learning skills through social interactions in a teachable agent environment,” *Research & Practice in Tech Enhanced Learning*, vol. 5, no. 02, pp. 123–152, 2010.
9. A. L. Baylor and Y. Kim, “Pedagogical agent design: The impact of agent realism, gender, ethnicity, and instructional role,” in *International Conference on Intelligent Tutoring Systems*. Springer, 2004, pp. 592–603.
10. G. Veletsianos, “Contextually relevant pedagogical agents: Visual appearance, stereotypes, and first impressions and their impact on learning,” *Computers & Education*, vol. 55, no. 2, pp. 576–585, 2010.
11. S. Domagk, “Do pedagogical agents facilitate learner motivation and learning outcomes?” *Journal of media Psychology*, vol. 22, no. 2, pp. 84–97, 2010.
12. N. L. Schroeder, O. O. Adesope, and R. B. Gilbert, “How effective are pedagogical agents for learning? a meta-analytic review,” *Journal of Educational Computing Research*, vol. 49, no. 1, pp. 1–39, 2013.
13. S. Finkelstein, E. Yarzebinski, C. Vaughn, A. Ogan, and J. Cassell, “The effects of culturally congruent educational technologies on student achievement,” in *AI in Education*. Springer, 2013, pp. 493–502.
14. J. N. Bailenson, “Doppelgangers-a new form of self?” *Psychologist*, vol. 25, no. 1, pp. 36–38, 2012.
15. A. Shapiro, A. Feng, R. Wang, H. Li, M. Bolas, G. Medioni, and E. Suma, “Rapid avatar capture and simulation using commodity depth sensors,” *Computer Animation and Virtual Worlds*, vol. 25, no. 3-4, pp. 201–211, 2014.
16. J. Gratch, N. Wang, A. Okhmatovskaia, F. Lamothe, M. Morales, R. J. van der Werf, and L.-P. Morency, “Can virtual humans be more engaging than real ones?” in *Human-Computer Interaction*. Springer, 2007, pp. 286–297.
17. M. T. Chi, S. A. Siler, H. Jeong, T. Yamauchi, and R. G. Hausmann, “Learning from human tutoring,” *Cognitive Science*, vol. 25, no. 4, pp. 471–533, 2001.
18. M. Rosenberg, *Society & the adolescent self-image*. Princeton university press, 2015.
19. J. E. Leone, E. M. Mullin, S. S. Maurer-Starks, and M. J. Rovito, “The adolescent body image satisfaction scale for males: exploratory factor analysis and implications for strength and conditioning professionals,” *The Journal of Strength & Conditioning Research*, vol. 28, no. 9, pp. 2657–2668, 2014.
20. IPIP, *Preliminary IPIP Scales Measuring Constructs Similar to Those Included in Lee and Ashton's HEXACO Personality Inventory*, 2015 (accessed 2015). [Online]. Available: http://ipip.ori.org/newHEXACO_PI_key.htm#Anxiety
21. B. G. Witmer, C. J. Jerome, and M. J. Singer, “The factor structure of the presence questionnaire,” *Presence*, vol. 14, no. 3, pp. 298–312, 2005.
22. J. Fox and J. N. Bailenson, “Virtual self-modeling: The effects of vicarious reinforcement and identification on exercise behaviors,” *Media Psychology*, vol. 12, no. 1, pp. 1–25, 2009.
23. A. Feng, E. S. Rosenberg, and A. Shapiro, “Just-in-time, viable, 3-d avatars from scans,” *Computer Animation and Virtual Worlds*, vol. 28, no. 3-4, 2017.