FISEVIER

Contents lists available at SciVerse ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh



How do learners respond to pedagogical agents that deliver social-oriented non-task messages? Impact on student learning, perceptions, and experiences

George Veletsianos*

University of Texas at Austin, College of Education, Instructional Technology, United States

ARTICLE INFO

Article history: Available online 19 October 2011

Keywords:
Pedagogical agents
Non-task interactions
Off-task interactions
Learner experience
Human-computer interaction
Uncanny valley effect

ABSTRACT

In this paper, I investigate the impact of non-task pedagogical agent behavior on learning outcomes, perceptions of agents' interaction ability, and learner experiences. Quasi-experimental results indicate that while the addition of non-task comments to an on-task tutorial may increase learning and perceptions of the agent's ability to interact with learners, this increase is not statistically significant. Further addition of non-task comments however, harms learning and perceptions of the agent's ability to interact with learners in statistically significant ways. Qualitative results reveal that on-task interactions are efficient but impersonal, while non-task interactions were memorable, but distracting. Implications include the potential for non-task interactions to create an uncanny valley effect for agent behavior.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Pedagogical agents are virtual characters employed in digital environments for instructional purposes. These characters are most frequently presented as digital teachers, tutors, or learning companions, and exhibit an exclusive focus on the task and the content that is to be learned/taught. While content is crucial to instruction, teaching and learning are also social endeavors (Gehlbach, 2010; Jones & Issroff, 2005), encompassing interactions and activities beyond content delivery. To illustrate this argument, imagine a teacher during a 45-min lesson. Is s/he completely focused on the task without deviating at all from the lesson's goals and objectives? On the other extreme, is s/he entirely focused on matters outside of the lesson, discussing his/her latest home improvement project instead of content relevant to the course? Classroom interactions encompass a mix of on-task and non-task interactions and behaviors. Presumably, when instructors interact with learners on matters that are unrelated to the lesson, they perceive non-task interactions to be beneficial to teaching. Cooper and Baynham (2005, p. 18) for instance note that "talk over cars, beer and football paves the way and lightens the atmosphere in the process of learning about more anxiety-causing topics" enabling instructors and learners to "develop positive relationships not just through interaction over subject matter but over personal and more general issues, where again more positive emotion can be engendered." Why, then, are most agents exclusively

E-mail address: veletsianos@gmail.com

task-oriented? If real-world instructors interject social remarks and comments in their teaching, can agents use non-task contexts to enhance educational endeavors? Does non-task commentary add any benefits to agent-based instruction? Or, does it hinder learning? How do learners respond to socially-sensitive agents that introduce non-task contexts in their instruction?

In this paper, I investigate pedagogical agents that introduce non-task comments to a lesson and examine (a) impacts on learning and student perceptions, and (b) student experiences. I begin by reviewing the literature relevant to pedagogical agents and the impact that non-task contexts may have on learning. Next, I present my research questions, hypotheses, and method for investigating the research questions. I conclude by presenting and discussing the results of this investigation, making recommendations for future research and practice.

2. Review of relevant literature

The pedagogical agent literature suggests that agents can serve numerous instructional functions. For instance, in a review of the literature, Gulz (2004, p. 315) found that researchers claimed that pedagogical agents could enable "increased motivation, increased sense of ease and comfort in a learning environment, stimulation of essential learning behaviors, increased smoothness of information and communication processes, fulfillment of need for personal relationships in learning, and gains in terms of memory, understanding, and problem solving." In extending this investigation of the literature to 2011, Veletsianos and Russell (2011) found that pedagogical agents were also expected to engage learners, provide systematic instruction, and engender realistic instructional

^{*} Address: Instructional Technologies Program, Curriculum and Instruction Department, 244L Sanchez Building, 1 University Station D5700, University of Texas at Austin, Austin, TX 78712, United States. Tel.: +1 512 232 2284.

approaches that aid learning and support both cognitive processing and metacognitive skills. Empirical results supporting these claims however are ambiguous and often mixed (Gulz, 2004; Veletsianos & Russell, 2011), largely due to inconsistent experimental designs (Clark & Choi, 2005), varied agent modalities (Baylor & Ryu, 2003), and a multiplicity of variables (e.g., agent role, image, and voice) interacting in complex ways, thus rendering comparisons difficult (Louwerse, Graesser, Lu, & Mitchell, 2005). Importantly, the pedagogical agent literature centers on on-task contexts, such that much less of the literature is focused on socio-cultural issues surrounding pedagogical agent implementations (Kim & Baylor, 2006). A smaller set of empirical studies seeks to understand agent deployments that encompass non-task contexts.

Non-task social aspects of learning have received limited attention in the education literature overall. Morgan-Fleming, Burley, and Price (2003) argue that this is the result of an implicit assumption that no pedagogical benefits are derived from non-task behavior, hence the reduction of off-task activities in schools such as recess time. The lack of attention received by this topic extends to the technology-enhanced learning literature (Abedin, Daneshgar, & D Ambra, in press). Nevertheless, researchers have argued that non-task interactions (e.g., greetings, small-talk) and contexts (e.g., opportunities to discuss off-task topics) are beneficial to learning and the learning process. For instance, social interactions appear to be positively related to student satisfaction (Aragon, 2003) and enjoyment (Muilenburg & Berge, 2005). Furthermore, Kreijns, Kirshner, and Jochems (2003, p. 336) note that one reason that digital learning environments fail to yield positive outcomes is due to the "tendency to restrict social interaction to educational interventions aimed at cognitive processes while social (psychological) interventions aimed at socio-emotional processes are ignored, neglected or forgotten." Essentially, Kreijns et al. argue that digital environments that focus exclusively on supporting task-related work are problematic because student interaction and participation that is only focused on the learning task ignores social aspects of communication and hinders the development of attitudes and feelings (e.g., belongingness, community, comfort, etc.) that are central to the learning process. For instance, Veletsianos and Doering (2010) heavily capitalized on non-task contexts to draw students to an online learning environment and to the study of climate change and sustainability. The environment centered on a curriculum that followed the adventures of a team of explorers and educators as they traveled through Arctic regions of the world on dog-sledding expeditions. The unfolding narrative of the expedition, students' fascination with Arctic exploration, students' enthusiasm about dogs, and students' personal connection to the explorers helped foster emotional connections, and captured students' attention and imagination drawing them into the lesson (Veletsianos & Doering, 2010).

While the pedagogical agent literature has not purposefully examined non-task contenxts/interactions, the topic surfaces in the literature. For instance, researchers studying interactions between users and agents installed in museums have discovered that non-task social interactions account for a significant proportion of agent-user conversations: Gustafson and Bell (2000) found that about one third of user utterances (n = 10,058) were social in nature (greetings and personal remarks, excluding insults), while Robinson, Traum, Ittycheriah, and Henderer (2008) found that approximately 65% of user-initiated utterances focused on dialogue functions (greetings, polite social phrases, and closing) or non-task information requests (biographical questions, personal preferences). Doering, Veletsianos, and Yerasimou (2008) developed a mixed-initiate agent to support learners in creating a digital portfolio to showcase their work, and found that learners engaged in multifaceted social dialogue with the agent, asking questions about the agent's personality and general interests, as well as posing questions relating to general encyclopedic information. Prior research has also demonstrated that non-task interactions may not always be appropriate, as abusive users may bully and harass virtual agents (DeAngeli, 2009; Veletsianos, Scharber, & Doering, 2008). While this literature demonstrates that agent–learner interactions encompass non-task behaviors, a limited number of studies has sought to specifically examine agent non-task behavior or understand learner experiences with and perceptions of non-task agents.

One of the literature findings suggests that users may not necessarily want or prefer to interact with social-oriented agents that make use of non-task comments. For instance, Gulz (2005) asked ninety adolescents to choose between a task-oriented agent and a task-and relation-oriented agent and to rationalize their choice. A relation-oriented agent was described as one that "attempt[ed] to develop a relationship with the participant: supplying information about him or herself, experiences, interests, etc. engaging in small talk and more personal kinds of discussions, and so on" (Gulz, 2005, pp. 411–412). Thirty-seven students (about 41%) stated that they would prefer a strictly task-oriented agent and argued that such an agent would keep them focused on the task while avoiding unnecessary conversations and meaningless interactions. On the other hand, students who selected the task- and relation-oriented agent argued that such an agent would make the experience more fun and playful while also enabling them to develop a personal relationship with the agent. Similar findings have been presented by Bickmore and Cassell (2005). In their study, introvert users liked the agent more when it only talked about the task, while extroverts liked the agent more when it used social dialogue.

Whether non-task contexts are beneficial to learning and agent-learner interaction is still a matter of debate. On the one hand, Bickmore, Schulman, and Yin (2009) reported that users that interacted with virtual agents that shared autobiographical stories in the first person enjoyed the interactions and completed more conversations with the agent, when compared with users whose agents shared stories in the third person. On the other hand, cognitive load theory suggests that learning might be harmed when learners divide their attention between information sources and process information that is peripheral to their learning. For instance, when agent animation (an off-task context) is presented simultaneously with other visual information (e.g., on-screen text), learning might be hampered (Choi & Clark, 2006; Clark & Choi, 2005). In the context of this paper, non-task comments and contexts might impede learning because they might force learners to divide their attention between information that is relevant and information that is irrelevant. For instance, learners in Veletsianos (2009) noted that they were distracted by non-task contexts peripheral to their learning such as the agent's lack of expressiveness, lack of enthusiasm, and frequent movement. On the other hand, Veletsianos, Miller, and Doering (2009) have hypothesized that agents that maintain an appropriate balance between on- and off-task communications may be able to enhance agent-learner interactions because short-term interruptions might allow learners the chance to refocus and recuperate. Nevertheless, unless there exists an algorithm for deciding when learners need a break, a balance between on- and off-task communications is difficult to achieve prior to learners engaging with the task. Silvervarg, Haake, Pareto, Tärning, and Gulz (2011) refine this line of argument by noting that while the use of off-task conversations in their pedagogical agent research enhanced the learner experience and did not necessarily distract from learning, it is possible that the positive benefits apply only to long-term learning contexts where the purpose of off-task comments is to foster agent-learner relationships.

3. Research questions

For the purpose of this study, I posed the following research questions:

- What is the impact of a pedagogical agent interjecting non-task comments on:
 - Learning?
 - Student perceptions of the agent's interaction ability?
- How do learners respond to agents that introduce non-task information during instruction?

4. Hypothesis

I hypothesize that adding non-task comments to a completely on-task lesson will *initially* improve (a) learning outcomes, and (b) learners' perceptions of the agent's interactional ability. The continuous addition of non-task comments to a lesson however, will yield incrementally smaller benefits, eventually reaching a point after which additional non-task commentary will harm learning and agent–learner interaction. While non-task contexts may assist in the development of a social and relaxed atmosphere in which learning can happen, continuously adding non-task comments will distract learners.

5. Method

5.1. Participants

Participants were enrolled in four elementary/special education technology courses. The courses were content- and cohort-specific and part of a 15-month post-baccalaureate masters program in education. One hundred and nine students were invited to participate. Out of those, 88 chose to participate. Of the 88 students who participated, 80 were females and 8 were males, and 82 reported their age (Mean = 23.46 years, SD = 6.26).

5.2. Materials

The materials used in this study consisted of three tutorials, one pedagogical agent, a post-test survey, a post-test exam, and an open-ended interview protocol.

5.2.1. Tutorial

Three tutorials were developed for this study. Each tutorial consisted of a lecture and answers to four questions. The tutorial was presented in an informal tone, and its content introduced participants to the use of technology in the classroom and raised issues that teachers need to consider when integrating technology in their classrooms. For example, the agent noted that "researchers claim that video games are fun, motivating, challenging, and relevant to children's life out of schools" but "not all video games are appropriate" for educational settings. The three versions of the tutorial differed in terms of non-task comments. In particular, the first lesson consisted only of on-task material (hereafter *on-task* group); the second lesson was composed of the first lesson plus non-task commentary (hereafter *on/off task* group); and the third lesson was composed of the second lesson plus additional

non-task commentary (hereafter *off-task* group). Appendix A includes sample tutorial content, and Table 1 shows the duration of the three tutorials.

The tutorial's content was expected to be unfamiliar to the students as the study was conducted on the first time each class met and this was the only educational technology requirement for students enrolled in this degree. Nevertheless, to ensure that prior knowledge was uniform in the three groups (and therefore not a confounding variable), student knowledge regarding technology use in education was evaluated and used as a covariate in the statistical analysis.

5.2.2. Pedagogical agent

One female pedagogical agent was used in this study (Fig. 1). The same pedagogical agent delivered the three versions of the tutorial lesson described above. Each tutorial followed the same delivery pattern: at first, the pedagogical agent delivered a lecture verbally using Text-To-Speech technology. At the end of the lecture, students were provided with four buttons, each consisting of one question. At the click of each button, the pedagogical agent provided a verbal response to the question.

5.2.3. Post-task survey

A survey was used to collect (a) demographic information (gender, age, and grade point average), (b) information on computer knowledge and skills, (c) information regarding knowledge of technology use in education, and (d) student perceptions of the agent's interaction ability. Survey responses were combined to form an index measuring computer knowledge and skills (i.e. the extent of participants' technical skills), an index measuring knowledge of technology use in education (i.e. the extent of participants' knowledge of using technology for pedagogical purposes), and an index measuring agent's perceived interaction ability. Cronbach's alpha a coefficient of reliability - was used to measure how well the survey responses measured the internal consistency of the aforementioned indices. The coefficient is used to justify combination of a set of items in an index. Cronbach's alpha for the computer knowledge and skills index was assessed at .79, for the knowledge of technology use in education index was measured at .94, and for the perceived interaction ability index was assessed at .80. Values above .70 are considered satisfactory.

5.2.4. Post-task exam

Participants completed an exam consisting of fifteen questions focusing on the on-task content presented to them. Twelve of these questions asked participants to recall information, and three asked them to apply information that they recalled (7 were multiple-choice, 5 were fill-in-the-blanks, and 3 were true-false). To minimize threats to the exam's validity, participants were encouraged to leave questions unanswered and avoid guessing if they did not know or remember the answer to a question.

5.2.5. Focus group protocol

Students were invited to participate in four focus group sessions (one per course). In these sessions, participants were asked to discuss their experiences of interacting with the pedagogical agent. The focus group protocol (Appendix B) allowed for open-ended discussion sessions.

Table 1
Tutorial duration (in min).

Agent on/off task commentary	Lecture	Answer #1	Answer #2	Answer #3	Answer #4	Total time
On-task	2:06	0:40	0:42	0:48	0:35	4:51
On/off task	2:31	0:46	0:49	0:52	0:42	5:40
Off-task	3:03	0:55	0:53	0:57	0:53	6:41



Fig. 1. The pedagogical agent used in this study.

Table 2 Study design.

Agent on/off task commentary	Participants		
On-task	N = 30		
On/off task	N = 26		
Off-task	N = 32		

5.3. Experimental design and treatments

A between subjects factorial design with three independent samples was employed. The experimental factor was pedagogical agent on/off task commentary with courses randomly assigned to one of three groups: The on-task, on/off task, or off-task group (Table 2). Each group consisted of students from a single class, with the exception of the third experimental group, which consisted of students from two classes.

5.4. Dependent measures

5.4.1. Perceived interaction ability

The agent's ability to interact with learners was evaluated as a composite measure of three survey items. Specifically, participants were asked to rate their communication with the agent in terms of smoothness, naturalness, and effectiveness and these variables were combined to form the *perceived interaction ability* index. The conceptual and theoretical basis for defining agents' interaction ability in terms of their communication being smooth, natural, and effective was originally proposed in Veletsianos, Miller, and Doering (2009) and was grounded in literature pertaining to group theory (Johnson & Johnson, 2006), conflict theory (Deutsch, 1973), and cooperative learning (Johnson, Johnson, & Holubec, 1993). The implication is that agents interaction ability is heightened when their communications is perceived to flow, occur naturally, and able to achieve its intended purpose without being abruptly interrupted or imposed upon communicators.

5.4.2. Learning

In this study, learning was operationalized as information recall. Outcomes were assessed via the post-task exam described above. Answers were graded to form a total score for each participant. Because there was only one correct answer for each question, one researcher graded all responses. Responses to the five fill-in-the-blanks questions that were considered to be ambiguous were discussed with a second researcher to limit researcher bias and reach a grading consensus.

5.5. Data sources

The data informing this study are both quantitative and qualitative. The quantitative data were collected via the post-task survey and exam. The qualitative data were collected via four focus groups sessions, after students completed the post-task survey and exam. Oualitative data were audio-recorded and transcribed verbatim.

5.6. Data analysis

5.6.1. Quantitative data

Experimental data were analyzed using the between subjects Multivariate Analysis of CoVAriance (MANCOVA) procedure. Specifically, MANCOVA assisted in examining the extent to which non-task commentary influenced (a) learning outcomes, and (b) perceptions of pedagogical agent interaction abilities. Significant MANCOVA effects were further examined with univariate ANOVA procedures. For all quantitative analyses, alpha was set at .05.

5.6.2. Qualitative data

Qualitative data were analyzed using the constant comparative method (Glaser & Strauss, 1967). Data were open-coded to arrive at salient categories and data patterns. First, data from each experimental group were read and analyzed independently to note emerging patterns and to gain a broad understanding of the learner experience. Next, data across all groups was analyzed in search of common themes and meanings. Once all transcripts were analyzed in the manner described above, patterns were compiled and reanalyzed in order to derive themes across all qualitative data. Analysis continued until no more patterns could be identified and it was felt that the data had been completely represented by the final themes (i.e. saturation had ben reached). In order to ensure that the researcher's pre-understanding did not influence the analysis of data, the researcher engaged in a process known as bracketing (Giorgi, 1997) which is a disciplined and systematic effort to consciously contain previous understandings of studied phenomena.

5.7. Procedure

A researcher visited four educational technology course sections on their first class session. At that time, the students were informed of the research and the tasks involved. Informed consent forms were distributed and students were given time to read the forms and ask any questions they may have had about the process or the research. Specifically, participants were told that a virtual character would present to them information regarding the use of technology in the classroom. To avoid confounding the results, participants were *not* informed that unrelated commentary was the main variable examined in this research. Participation in the investigation was strictly voluntary and students who chose not to participate were permitted to work on course assignments. Students who chose to participate returned a signed copy of the consent form and were given an extra copy for their own records.

Participants were then directed to view the tutorial lesson presented by the pedagogical agents. Each participant was seated in front of a desktop computer, equipped with headphones and the pedagogical agent software. Prior to commencing the task, participants were directed to pay attention to the lesson and to refrain from taking notes or engaging with any other computer task.

Participants wore the headphones, launched the software, tested the audio equipment, and, if all worked fine, began viewing the tutorial lesson. If any issues arose prior to the commencement of the presentation, participants were directed by the software to raise their hand and the researcher would provide any necessary assistance. At the end of the lesson, participants were redirected to a website where they could enter their answers to the post-task survey and post-task exam. On average, this process lasted for approximately 40 min. At the end of this task, participants were invited to share and discuss their experiences in a focus group format. The researcher first explained the focus group setting and its intention and then engaged participants in a semi-structured discussion regarding the experience of interacting with pedagogical agents. On average, the focus group sessions lasted 20 min each. Eight individuals from each group participated in the focus groups.

6. Results

The study's findings are presented under a quantitative and a qualitative section.

6.1. Quantitative results

A MANCOVA indicated a significant main effect for the treatment factor (*Wilks'* Λ = .71, F(4,160) = 7.35, p < .001, *partial* η^2 = .16). Results indicated no significant effects for the gender, grade point average, computer skills, and knowledge of classroom integration practices covariates. These variables do not explain any variations across the three groups. Dependent variable means and standard deviations are displayed in Table 3.

6.1.1. Interaction ability

Follow-up univariate ANOVA tests indicated that the three pedagogical agents (on-task, on/off task, off-task) differed significantly in terms of their interaction ability [F = (2.81) = 5.42, p = .006]. Pairwise comparisons between the treatment groups using the Bonferroni correction indicated two significant differences between the groups at the .05 level: (a) a significant difference (p = .015) between the on-task group (M = 9.00, SD = 2.69) and the off-task group (M = 7.28, SD = 2.13), and (b) a significant difference (p = .028) between the on/off task group (M = 9.15, SD = 1.85)and the off-task group (M = 7.28, SD = 2.13). The standardized effect size for the first difference was medium-large (Cohen's d = 0.71) and for the second difference was large (Cohen's d = 0.94). Simply stated, the on-task and the on/off task groups each rated the pedagogical agent as having better interaction ability than the off-task group. No significant differences were found (p = .99) between the on-task (M = 9.00, SD = 2.69) and on/off task (M = 9.15, SD = 1.85) groups.

Table 3Sample size, mean, and standard deviations.

Dependent variable	Agent	n	Mean	SD
Learning	On-task On/off task Off-task Total	30 26 32 88	7.90 9.46 6.81 7.97	1.61 1.68 2.01 2.07
Interaction Ability	On-task On/off task Off-task Total	30 26 32 88	9.00 9.15 7.28 8.42	2.69 1.85 2.13 2.40

6.1.2. Learning

Univariate ANOVA tests indicated that learner outcomes differed significantly between the three treatment groups [F = (2,81) = 11.32, p < .001]. Pairwise comparisons between the treatment groups indicated two significant differences between the groups at the .05 level: (a) a significant difference (p = .043) between the on-task group (M = 7.90, SD = 1.61) and the off-task group (M = 6.81, SD = 2.01), and (b) a significant difference (p = .001) between the on/off task group (M = 9.46, SD = 1.68) and the off-task group (M = 6.81, SD = 2.01). The standardized effect size for the first difference was medium-large (Cohen's d = 0.60) and for the second difference was large (Cohen's d = 1.43). In other words, the on-task and the on/off task groups each scored higher than the off-task group in the post-task test. No significant differences were found between the on-task and on/off task groups (p = .058), even though the on/off task group scored higher (M = 9.46, SD = 1.68) than the on-task group (M = 7.90, SD = 1.61).

6.2. Qualitative results

Qualitative analysis generated four themes (Table 4). I present and describe each theme using representative participant quotes to illustrate the findings. Inferences drawn from these findings are discussed in the implications section of this paper.

6.2.1. On-task interactions are efficient, but impersonal; non-task interactions are memorable but distracting and "fake"

While the conscious decision was made to avoid introducing the topic of non-task behavior, participants chose to share their feelings and perceptions regarding the variable of interest, indicating that (at the very least) they took note of the agent traversing the on-task/off-task space.

Participants in the on-task group noted that when the agent only focuses on the task, this makes for an efficient but impersonal and problematic means of agent-learner interaction. If learners lose attention, the agent may not be able to redirect their attention back to the task. Jenny for instance noted, "She [the agent] was very straightforward. There was no way you can go off the subject versus when you have a teacher and if you have any questions you can go on a tangent and talk about other things and get side-tracked." Brenda reiterated the efficiency argument, but also argued that that strict adherence to on-task issues might prevent the agent to redirecting students back to the lesson if they lose focus, "I think it could be efficient, but if you lose the attention of the students then they are gone."

Students participating in the on/off task group reported that, while the agent's non-task comments were distracting at times, they were also memorable. Penny for example, noted that her attention was directed towards extraneous comments rather than on-task information, "I would just, like, hear the main five points and then I found myself focusing more on irrelevant stuff... that she was saying, than the actual presentation." Furthermore, students also noted that they could relate to some of the non-task comments and this served to make non-task comments more memorable. Grace for example, noted, "I remember her comments about the dog because it related to something that made sense."

Table 4 Qualitative Themes.

Theme	
On-task interactions are efficient, but impersonal; non-task interactions are	
memorable but distracting and "fake"	
Dissatisfaction with computer-generated voice	
Agent-learner relationships are complex	
Student expectations of "naturalness" and "normality"	

Finally, in addition to reporting that non-task comments were distracting and memorable, the off-task group also noted that such comments seemed planned and fake, as if the agent was attempting to present itself as a human. Susan for instance, noted that once the task was finished she, "remembered those [the non-task comments] more than the [on-task] points she was making. I remember the sunglasses comment but I don't remember what she said about technologies." In explaining why this could be the case, students noted that information that was not directly related to the task was "interesting," "different," and "fun." Yet, participants noted that the agent's comments that were non-task sounded "rehearsed," and were unconvincing and, at times, distracting. Fred captured such feelings quite well when he stated that, "comments that were irrelevant sounded rehearsed...so I was thinking 'Oh, you were planning on saying that the whole time!' I feel like when you make comments like that and they are scripted, they don't capture it."

6.2.2. Dissatisfaction with computer-generated voice

One theme was overwhelmingly uniform and consistent across all three groups: Participants were dissatisfied with the agent's computer-generated voice, noting that it was "robotic," "monotonous," "annoying," "obnoxious," "non-enthusiastic," and "distracting." In general, participants found it "difficult to listen to" and "easy to stop listening to it." Janet summarized such feelings by expressing her ideas about what was problematic about the computer-generated voice and what can be done to improve it, "The pitch and the tone didn't really change so I started to doze off a little bit...[It could be improved] if the volume would have fluctuated more with a potential raising or lowering of the voice to insinuate a more of a depth of point." Tina further explained, "I thought it was sort of difficult to listen to because she put the wrong emphasis on certain parts of words. Certain words wouldn't sound right. It was the computer's interpretation. I found this sort of hard to listen to at times."

6.2.3. Agent-learner relationships are complex

Another theme that quickly arose from the focus group sessions is the intricate nature of the student-agent relationship. Learner comments indicated that when learners interact with agents, the experience and the interaction between the two appears to be multidimensional and mediated by social, cultural, and interpersonal issues. Interacting with a pedagogical agent is not a simple act of information exchange and goes beyond strict notions of computer-assisted instruction. For example, when asked to describe the agent, Austin commented, "She is petite," essentially ascribing human-like qualities to the agent, and suspending belief that he was interacting with a computer program. At other times, participants stated that this is merely a computer-based tool, "When I started thinking about it, I imagine that four or five programmers are thinking about these questions and topics." Others, delved even deeper into the human vs. tool dichotomy, describing the agent as a tool with intentions, "Every time she did that [moved her head/ made a non-task remark] I just thought, Oh, it's trying to have human interaction, have a personality." Even when the agent appeared to have intentions (through designers' purposeful attempts at infusing human-like elements in the design of the pedagogical agent), learners resisted treating the agent as a real human, "I think that personal stories and stuff are interesting when it's someone you got to know like your teachers and stuff, but it's hard to be relatable when it's something virtual like this digital teacher."

Another level of complexity arose when participants were asked to discuss agent shortcomings or present their ideas on how the agent can be improved. In these cases, suggestions drew parallels to human-to-human interactions. For instance, Joan noted the importance of trust and credibility,

I think as kind of a general rule, in order to pay attention and learn from somebody you first have to assess whether you feel them as valuable. And I didn't see a pre-history or anything, I just saw a person kind of come up and teach me. When we came to this room, we met our instructor, our instructor actually has some, we know of a credential behind [him/her], so we have a preconceived notion that we will learn something versus just seeing a picture on the screen and not having a prior value or prior credential to judge it.

Jack extended the notion of credibility to respect and the fact that the pedagogical agent lacked reactivity, advanced understanding, and the human element,

I don't think she is very demanding of respect or attention or anything, because if we all look at the back of the room, she would just keep on going and not have any idea that we weren't paying attention at all. You know? She doesn't respond. If we weren't understanding something she was saying, a teacher could see people's puzzled looks on their faces that even a robotic thing with voice inflection won't be able to read the looks on people's faces and all the other cues that students give.

This result indicates that learner-agent relationships are intricate, not just because learners may treat agents as human counterparts, but also because learners might actively resist and reject the agents' human-like features. This finding is important for the design and research of such tools and I expand further upon this issue in the implications section of the paper.

6.2.4. Student expectations of "naturalness" and "normality"

Throughout the focus groups, learners made multiple references to the agent's voice, movement, and manner of interaction as being "unnatural." In addition, her non-task comments and her "personality" were described as "not normal." In the following conversation with the researcher, Danny captures the essence of such comments,

Danny: It seemed like: "I [the agent] am trying to seem like I am not a computer person. I am natural"...and it's sort of...I don't know.

Researcher: So, what are your thoughts about this?

Danny: Well, I was like, 'No! You are a computer! Stay there!' Yeah, that's what it was. I'd rather have a normal person...It wasn't convincing cause I could tell she was a computer person and not a real person.

It appears that learners expected the agent to be "natural" and "normal." Even though participants may have had preconceived notions of the ideal computer-based digital teacher (e.g., anthropomorphous robots popularized by science fiction films), the references to "natural" and "normal" imply that (a) agents were compared to humans, and (b) learners did not feel that the agents were acting as if they were humans or humanlike. In this context, when learners refer to the agent as being unnatural, they mean that the agent was not close to human life - essentially, the agent did not have a human demeanor. In short, "being natural" means being lifelike, and the agent was not lifelike. With regards to the agent "being normal," learners seem to mean that the agent did not abide by the expected social norms that govern human behavior. As shown above, participants were often skeptical of agents' lifelike behaviors, and were unconvinced by their attempts at small talk and non-task interactions. This issue reveals a paradox: even though participants suggested human-like features in order to improve the agents such that they appear more normal and natural, they resisted their lifelike behaviors and attempts at indicating human-like traits and behaviors.

7. Discussion

This investigation commenced with the following questions: Does non-task commentary add any benefits to agent-based instruction? Or, does it hinder learning? How do learners respond to socially-sensitive agents that introduce non-task contexts in their instruction? Students' responses to agents and to their non-task comments, raise implications for the design of technology-mediated learning in general, and pedagogical agents in particular.

Empirical results show that incremental amounts of non-task comments harm learning and perceptions of the agent's ability to interact with learners in statistically significant ways. What this research has not investigated however is (a) the point at which agent non-task behavior begins to harm agent-learner interaction, and (b) the type of non-task behavior that is most beneficial to agent-learner interactions and learning. With regards to the former, researchers need to investigate more refined granulations of non-task behavior. Regarding the latter, the agent employed in this study utilized multiple non-task comments (e.g., greetings, humor). Even though it's the totality of those comments that was investigated in this study, it is possible that only a set of these contribute to enhanced learning and improved communication when learners interact with agents. Investigations focusing on both of these issues may yield valuable insights for research and practice. One of our challenges is to discover how to infuse on-task and nontask interactions in agent-based environments to develop learner experiences that enable smooth transitions between topics of interest regardless of whether such topics are on-task or non-task. The complex nature of agent-learner relationships also means that we need to investigate and consider pedagogical agents from multiple perspectives. Researchers who are interested in exploring what may be possible with digital companions should not only consider whether agents contribute to learning gains, but should also investigate learner experiences and the meaning and importance of such experiences.

The findings of this research also indicate that learners both humanized the agents and expected them to abide by social norms, but also identified the agents as programmed tools, resisting and rejecting their lifelike behaviors. While non-task interactions may motivate human counterparts for instance, in this research it was found that some students found such comments to be "rehearsed" and "scripted." This finding is in contrast to the media equation literature which notes that humans treat media as human counterparts (Reeves & Nass, 1996). Additionally, this finding poses a design conundrum: learners may identify human behaviors/traits as contributing to enhancing the agents, but designers' actions to create more life-like agents may damage the learning experience. Designing agents that resemble humans more closely runs the risks of inducing inflated psychosocial expectations of agent abilities, resulting partly due to the fact that the more human-like an agent appears, the more learners will ascribe human qualities to it (Krach et al., 2008). Designing pedagogical agents that resemble human appearance and behavior may be desirable from the learners' perspective, but undesirable from a social and educational standpoint. For instance, it is possible that non-task behavior pushed pedagogical agents into a territory akin to the "uncanny valley," the theory which posits that as virtual characters approximate human appearance they become distracting and unsettling (MacDorman, Green, Ho, & Koch, 2009; Mori, 1970). While the uncanny valley effect refers to agent's visual appearance, results form this research hint at an uncanny valley effect in terms of character behavior. Thus, humans may not just find virtual characters with human faces unsettling, but they might also find virtual characters with human behaviors disturbing. This effect may help explain some of the results of this study: Incremental amounts of

non-task comments, may have made pedagogical agents more disconcerting, influencing learning and learner perceptions in an adverse manner. This hypothesis could be further examined by future research.

Finally, this paper clearly illustrated learners' discontent with Text-To-Speech (TTS) software and the distracting qualities of computer-generated audio. At present, the dynamic generation of natural and crisp human voices is not attainable, so designers may need to devise alternate methods to compensate for this deficiency. For instance, designers can use TTS voices sporadically so as to avoid distractions. Even though humans are accustomed to oral communication, we also depend on complex non-verbal channels of communication to interact with others (Argyle, 1988). Capitalizing on humans' ability to communicate equally well in non-verbal terms, agents could employ differential patterns of interaction to avoid being distracting (e.g., employing deictic gestures) when directing learner attention to an item of interest).

8. Limitations

While this study informs designers and researchers of the use of non-task interactions in pedagogical agent use, the following limitations should be considered:

- Approximately 91% of the participants in this study were female students. Results might have been influenced by the selected sample, and it is advised that future studies investigate more diverse populations. For instance, those individuals who selected a task-oriented agent in the study conducted by Gulz (2005) might have performed better with the on-task agent than any of the other two agents.
- Results might have been impacted by the agent's voice. Future studies are advised to account for this variable, ensuring that treatments are not disproportionately influenced by the agent's voice (e.g., off-task group participants might have responded more positively if they were not distracted by the agent's voice).
- The design of non-task comments. In this evaluation, off-task comments appeared in pre-defined intervals determined by the researcher. This design can be improved in future interventions in the following ways:
 - Agent-controlled: In this case, the agent decides the appropriate time to interject non-task comments (e.g., by monitoring students' cognitive load and intervening when it is deemed that cognitive resources are overloaded). This approach requires the development of software capable of making decisions on the on-task/non-task mix needed to satisfy current learning needs.
 - User-controlled: In this case, the decision to move from on-task to non-task interactions is left to the learner (e.g., the learner decides that s/he feels tired and needs a break, so s/he informs the agent that s/he would like to have a non-task conversation). This approach requires the learners to be self-directed and able to decipher whether non-task interactions would be helpful or distracting to their learning. This is the approach used by Silvervarg et al. (2011).

9. Conclusion

In this paper, I examined the impact of non-task agent-learner interactions on learning outcomes, perceived interaction ability, and learner experiences. With regards to learning, results revealed that (a) learners in the on-task group and learners in the on/off task group each recalled significantly more information than learners in the off-task group, and (b) learners in the

on/off task group and learners in the on-task group did not differ significantly in the amount of information recalled. With regards to perceived interaction ability, results revealed that (1) the ontask and the on/off task groups each rated the pedagogical agent as having significantly better interaction ability than the off-task group, and (2) learners in the on/off task group and learners in the on-task group did not differ significantly in their ratings of agent interaction ability. Qualitative results illuminated the reasons for these findings, indicating the (a) efficient but impersonal nature of on-task interactions, (b) memorable, but distracting nature of non-task interactions, (c) participant dissatisfaction with the agent's computer-generated voice, (d) complexity of agent-learner relationships, and (e) student expectations for the agent to appear "natural" and "normal." While the results presented are not conclusive, they point to the potential of exploiting non-task interactions for enhancing learning and interaction in pedagogical agent contexts, while also highlighting potential pitfalls, such as the potential for an uncanny valley effect for agent behavior. Designers are encouraged to further explore the inclusion of non-task commentary in pedagogical agent work, and refine the results arising from this emerging area of study. Although task-focused agents may be an efficient way to deliver instruction, the increasing use of pedagogical agents within digital learning environments (vis-avis stand-alone agents) (Veletsianos and Russell, 2011) necessitates our continuous evaluation of agent-learner interactions, agent-learner relationships, and agent roles.

Appendix A. Sample content from the three tutorials used in this study

The off-task agent presented the narrative below. The on-off task agent presented the narrative below, excluding the bold text. The on-task agent presented the narrative below excluding the bold and italicized text.

Sample content #1: "Your technology staff can do anything with technology, but they also need your help. Since you are a beginning teacher, I will try to help you learn about the use of technology in your classroom. I have been helping beginning teachers for two years already! Below is a list of four questions that you can ask me. I know that you really want to ask about my hair, but you can do that at some other time. You can only ask me each question once. So pay attention. When you are done, I will quiz you on this material, just like you will quiz your students when you become a teacher!".

Sample content #2: "Yet, not all video games are appropriate for education. Video games used in classrooms should have instructional value, be appealing, allow ALL children to use them, and refrain from depicting violence or aggression. My cousin spends way too much time on video games. I am not familiar with most of the games she plays."

Appendix B. Focus group questions

- What did you think of the virtual character?
- What was difficult about your interaction with the virtual character?
- What was easy about your interaction with the virtual character?
- What did you like the most about the virtual character?
- What did you like the least about the virtual character?
- What are some adjectives that come to mind when asked to describe the virtual character?

B.1. If and when participants mentioned on/off task comments

- Tell me more about these [off-task] comments.
- The virtual character gave a presentation that was totally focused on the task. Do you have any comments about this?
- The virtual character gave a presentation that was off -task quite a bit of the time. Do you have any comments about this?
- The virtual character gave a presentation that included some comments that were not related to the task. Do you have any comments about this?

References

- Abedin, B., Daneshgar, F., & D Ambra, J. (in press). Do nontask interactions matter? The relationship between nontask sociability of computer supported collaborative learning and learning outcomes. British Journal of Educational Technology
- Aragon, S. (2003). Creating social presence in online learning environments. New Directions for Adult and Continuing Education, 100, 57–68.

Argyle, M. (1988). Bodily communication. London: Methuen.

- Baylor, A. L., & Ryu, J. (2003). Does the presence of image and animation enhance pedagogical agent persona? *Journal of Educational Computing Research*, 28(4), 373–395.
- Bickmore, T., & Cassell, J. (2005). Social dialogue with embodied conversational agents. In J. van Kuppevelt, L. Dybkjaer, & N. Bernsen (Eds.), Advances in natural multimodal dialogue systems (pp. 23–54). New York, NY: Kluwer.
- Bickmore, T., Schulman, D., & Yin, L. X. (2009). Engagement vs. deceit: Virtual humans with human autobiographies. In Z. Ruttkay, M. Kipp, A. Nijholt, & H. H. Vilhjalmsson (Eds.), Proceedings of the 9th international conference on intelligent virtual agents (pp. 6–19). Amsterdam: Springer.
- Choi, S., & Clark, R. (2006). Cognitive and affective benefits of an animated pedagogical agent for learning English as a second language. *Journal of Educational Computing Research*, 34(4), 441–466.
- Clark, R. E., & Choi, S. (2005). Five design principles for experiments on the effects of animated pedagogical agents. *Journal of Educational Computing Research*, 32(3), 209–225.
- Cooper, B., & Baynham, M. (2005). Rites of passage: Embedding meaningful language, literacy and numeracy skills in skilled trades courses through significant and transforming relationships. In National Research and Development Centre for Adult Literacy and Numeracy. http://www.nrdc.org.uk/uploads/documents/doc_2962.pdf Accessed 29.06.11.
- DeAngeli, A. (2009). Ethical implications of verbal disinhibition with conversational agents. *PsychNology Journal*, 7(1), 49–57.
- Deutsch, M. (1973). The resolution of conflict: Constructive and destructive processes. New Haven, CT: Yale University Press.
- Doering, A., Veletsianos, G., & Yerasimou, T. (2008). Conversational agents and their longitudinal affordances on communication and interaction. *Journal of Interactive Learning Research*, 19(2), 251–270.
- Gehlbach, H. (2010). The social side of school: Why teachers need social psychology. *Educational Psychology Review*, 22(3), 349–362.
- Giorgi, A. (1997). The theory, practice, and evaluation of the phenomenological method as a qualitative research procedure. *Journal of Phenomenological Psychology*, 28(2), 235–260.
- Glaser, B. G., & Strauss, A. L. (1967). The discovery of grounded theory. Chicago: Aldine Publishing.
- Gulz, A. (2004). Benefits of virtual characters in computer based learning environments: Claims and evidence. *International Journal of Artificial Intelligence in Education*, 14, 313–334.
- Gulz, A. (2005). Social enrichment by virtual characters Differential benefits. Journal of Computer Assisted Learning, 21(6), 405–418.
- Gustafson, J., & Bell, L. (2000). Speech technology on trial: Experiences from the August system. *Natural Language Engineering*, 1(1), 1–15.
- Johnson, D., & Johnson, F. (2006). Joining together: Group theory and group skills. Boston: Allyn & Bacon.
- Johnson, D., Johnson, R., & Holubec, E. (1993). *Cooperation in the classroom* (6th ed.). Edina, MN: Interaction Book Company.
- Jones, A., & Issroff, K. (2005). Learning technologies: Affective and social issues in computer supported collaborative learning. Computers and Education, 44(4), 395–408.
- Kim, Y., & Baylor, A. (2006). A social cognitive framework for designing pedagogical agents as learning companions. *Educational Technology Research and Development*, 54(6), 569–596.
- Krach, S., Hegel, F., Wrede, B., Sagerer, G., Binkofski, F., & Kircher, T. (2008). Can machines think? Interaction and perspective taking with robots investigated via fMRI. PLoS ONE, 3(7), e2597.
- Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: A review of the research. Computers in Human Behavior, 19(3), 335–353.
- Louwerse, M. M., Graesser, A. C., Lu, S., & Mitchell, H. H. (2005). Social cues in animated conversational agents. *Applied Cognitive Psychology*, 19(6), 693–704.

- MacDorman, K. F., Green, R. D., Ho, C.-C., & Koch, C. T. (2009). Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior*, 25(3), 695–710.
- Morgan-Fleming, B., Burley, H., & Price, M. (2003). The benefits of being off-task. *International Journal of Social Education*, 18(1), 1–8.
- Mori, M. (1970). Bukimi no tani [The uncanny valley]. Energy, 7(4), 33-35.
- Muilenburg, L. Y., & Berge, Z. L. (2005). Student barriers to online learning: A factor analytic study. *Distance Education*, 26(1), 29–48.
- Reeves, B., & Nass, C. (1996). The media equation: How people treat computers, television, and new media as real people and places. Cambridge: Cambridge University Press/CSLI.
- Robinson, S., Traum, D., Ittycheriah, M. & Henderer, J. (2008). What would you ask a Conversational Agent? Observations of human-agent dialogues in a museum setting. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odjik, S. Piperidis, et al. (Eds.), Proceedings of the 6th international language resources and evaluation association conference (pp. 1125–1131). Marrakech, Morocco.
- Silvervarg, A., Haake, M., Pareto, L., Tärning, B., & Gulz, A. (2011). Pedagogical agents: Pedagogical interventions via integration of task-oriented and socially

- oriented conversation. In Paper presented at the AERA 2011 conference, New Orleans, LA.
- Veletsianos, G. (2009). The impact and implications of virtual character expressiveness on learning and agent-learner interactions. *Journal of Computer Assisted Learning*, 25(4), 345–357.
- Veletsianos, G., & Doering, A. (2010). Long-term student experiences in a hybrid, open-ended and problem based Adventure Learning program. Australasian Journal of Educational Technology, 26(2), 280-296.
- Veletsianos, G., Miller, C., & Doering, A. (2009). EnALI: A research and design framework for virtual characters and pedagogical agents. *Journal of Educational Computing Research*, 41(2), 171–194.
- Veletsianos, G., & Russell, G. (2011). Pedagogical Agents. Manuscript under review. Veletsianos, G., Scharber, C., & Doering, A. (2008). When sex, drugs, and violence enter the classroom: Conversations between adolescent social studies students and a female pedagogical agent. *Interacting with Computers*, 20(3), 292–301.