Beyond robotic wastelands of time: abandoned pedagogical agents and new pedalled pedagogies

Savin-Baden, M. , Tombs, G. and Bhakta, R.

Author post-print (accepted) deposited in CURVE January 2015*

This document is the author's post-print version, incorporating any revisions agreed during the peer-review process. Some differences between the published version and this version may remain and you are advised to consult the published version if you wish to cite from it.

*cover sheet updated October 2015.

Beyond Robotic Wastelands of Time: abandoned pedagogical agents and *new* pedalled pedagogies

Maggi Savin-Baden, Gemma Tombs and Roy Bhakta
Disruptive Media Learning Lab
Coventry University, UK

Maggi Savin-Baden is Director of Research and Evaluation for the Disruptive Media Learning Lab at Coventry University. Her research interests and publications include digital fluency, pedagogical agents, qualitative research synthesis arts-based inquiry, conservation pedagogy and problem-based learning.

Gemma Tombs is a Senior Research Assistant for the Disruptive Media Learning Lab, with a particular focus on digital technologies. Her research interests include pedagogical design for digital technologies, educators' experiences of digital technologies, and the socio-cultural impact of new educational technologies.

Roy Bhakta is a Senior Research Assistant for the Disruptive Media Learning Lab. His research interests include conservation pedagogy, pedagogical agents, student experiences, digital fluency and improving learning in STEM subjects.

Abstract

Chatbots, known as pedagogical agents in educational settings, have a long history of use, beginning with Turing's (1950) work; and since then online chatbots have become embedded into the fabric of technology. Yet understandings of these technologies are inchoate and often untheorised. Integration of chatbots into educational settings over the past 5 years suggests an increase in interest in the ways in which chatbots might be adopted and adapted for teaching and learning. This paper draws on historical literature and theories that to date have largely been ignored in order to (re)contextualise two studies that used responsive evaluation to examine the use of pedagogical agents in education. Findings suggest that emotional interactions with pedagogical agents were intrinsic to a user's sense of trust, and that truthfulness, personalisation and emotional engagement are vital when using pedagogical agents to enhance online learning. Such findings need to be considered in the light of the ways in which notions of learning are being redefined in the academy, and the extent to which new literacies and new technologies are being pedalled as pedagogies in ways that undermine what higher education is, is for, and what learning means.

Introduction

In many ways it would seem that emerging communication technologies are disrupting and changing societal norms and conventions (Turkle, 2011). Whitty and Joinson (2009) have suggested that central to making sense of the unique qualities of cyberspace are understandings of such social networks and veracity, and studies by Yee (2006) and Bailenson, Yee, Blascovich and Guadagno (2008) also indicate that issues of online and offline behaviour bear further exploration. We propose that as pedagogical agents are seen to help support and even improve the level of interactive learning on a programme or

course (Kim & Wei, 2011), it is essential that these societal norms and behaviours are considered within pedagogical agent learning situations.

Chatbots are characters on the computer screen with embodied life-like behaviours such as speech, emotions, locomotion, gestures, and movements of the head, the eye, or other parts of the body (Dehn & van Mulken, 2000). We refer here to pedagogical agents, which, by our definition are chatbots used for educational purposes as opposed to commercial or business agents, for example. These technologies have been increasingly adopted and tested in educational settings, yet little is known about the ways in which they can be used effectively, and indeed whether they can provide additional value to learning experiences. Further, the research that has been undertaken has not yet drawn clear distinctions between application across disciplines and in difficult and sensitive settings (Heidig & Clarebout, 2011).

The central argument of the paper is that truthfulness, personalisation and emotional engagement are all vital components in using pedagogical agents to enhance online learning. The paper begins by reviewing the current and historical literature that has largely been ignored in other recent research. It examines knowledge of pedagogical agent use in blended learning contexts, focusing on issues of trust and truth telling as well as emotional engagement. The method and context of the study is then provided. The findings are presented and situated within the relevant body of literature and the paper concludes by offering several recommendations for practice.

Literature review: Secrets and Lies?

Although pedagogical agents have only received sustained attention from educational researchers in the past decade (Heidig & Clarebout, 2011), they have roots in research from eighty years ago. From 1936 onwards, Alan Turing began developing a theory that would culminate in the publication *Computing Machinery and Intelligence* (Turing, 1950). In this paper, Turing originally posited the question, "can machines think?", before discarding this in favour of a more easily measurable (and, to his mind, valuable) question which asked whether a computer might be able to fool a human being into believing they were conversing with another human being. From this point onwards, chatbot research and development focused upon one key goal: to develop a chatbot to which 'human' tendencies such as attitudes, language, reactions and mannerisms could be ascribed. Although multiple attempts have been made, most notably through the *Loebner Prize Contest*, Turing's 1950 challenge has yet to be met, and debates continue on whether the Turing test is either a stumbling block or a productive mechanism for improved chatbot development (Berrar & Schuster, 2014). Recently, developers have focused upon creating chatbots for specific purposes, thus limiting their necessary interactions to one specific context and increasing the likelihood of a successful Turing test. The success of the Eugene chatbot exemplifies this trend in that Eugene convinced 33% of the judges at the Royal Society in London that it was human (http://www.bbc.co.uk/news/technology-27762088), seemingly passing the Turing test. However, this experiment has been highly criticised by academics working in the field of artificial intelligence, on the basis that developers described the Eugene chatbot as a 13-year-old non-native English speaker. Consequently, any inconsistencies or failings in Eugene's interactions with the judges were attributed to his age and English language abilities, rather than ineffective chatbot development.

Thus it seems that the hopes and suggestions of authors such as Frude (1983) remain unrealised. Frude suggested that humans would form beneficial relationships with artificial

companions, based on the idea of animism, in which humans attribute life and human characteristics to inanimate objects. His later book (Frude, 1984) explored the relationship between humans and artificial intelligence in science fiction. In this Special Edition of E-learning and Digital Media, (Frude & Jandrić, 2015) he still remains convinced of his main hypothesis of 1983; that animism, artistry and artificial intelligence, if combined effectively, will result in a beneficial human computer relationship. The same pattern has emerged regarding research into the use of chatbots (or pedagogical agents) in educational settings; namely, that pedagogical agents are required to fill specific roles successfully and adhere to notions of human 'realism' in order to be successful in educational settings. Veletsianos and Russell (2014) have provided an excellent summary of the historical and theoretical foundations of the pedagogical agents used today, returning to their roots in the 1970s *Intelligent Tutoring Systems*. Most notably, they have addressed the diversity of pedagogical agent 'roles' which must now be filled, reinforcing the belief that Turing's goal of a 'human computer' is not yet achievable.

Whilst Frude remains optimistic, others are less so. For example, Kiran and Verbeek (2010) suggest that with every technological development what is at stake is the unique meaning of humanity. They do however, suggest that what is needed is that:

> "in order to be able to trust oneself to technology, two specific conditions need to be met. First, the technology in question needs to leave room to develop an explicit relation to its mediating role, rather than being dominating or overpowering. And second, human beings need to have the ability to "read" the mediating roles of the technology, and the skills to "appropriate" it in specific ways"

> (Kiran & Verbeek, 2010, p. 424)

When taking Kiran and Verbeek's perspective, then, the ability to trust oneself to chatbot technology is influenced by the roles which chatbots (and students) are required to fulfil. As social theories of learning have come to underpin most pedagogical approaches in the 21st Century, the use of pedagogical agents has shifted from predominantly instructional and informational roles to those of tutors, coaches (Payr, 2003) and learning companions (Kim & Baylor, 2006). Consequently, pedagogical agents are typically theorised as tools to support (social) cognitive development, and research associated with this has addressed the social interactions of pedagogical agents and students, considering the humanistic qualities of agent-student interaction and focusing on aspects such as users' cognitive styles (Lee 2010), the realism of the pedagogical agent (Gong 2008), and agent appearance (for example, Garau et al., 2003; Dunsworth & Atkinson, 2007).

Whilst more research is necessary, trends seem to indicate the formation of a relationship between student and pedagogical agent, which is informed by factors such as students' and agents' gender and ethnicity (Moreno & Flowerday, 2006). Yet this research has tended to neglect the social context within which students interact with pedagogical agents, with little attention paid to the educational topics studied, the sensitivity of discussions, and distinctions between participants of different genders, ages, and ethnicities. With the notable exception of Culley and Madhavan (2013), the topic of ethics in student-agent interactions has also been neglected. Yet the notion of trustworthiness has been identified as essential in the formation of an emotional connection between a user and an agent (Savin-Baden, Tombs, Burden, & Wood, 2013).

The potential formation of an emotional connection between the user and the agent is affected by the concept of social presence, in which users might feel 'present' in an interaction with a chatbot. When disclosing sensitive information, the vulnerabilities and potential risk associated with trust – even of a pedagogical agent or chatbot – are that much higher. Consequently, the use of agents in both commercial and educational settings may disrupt the ways in which interaction in online settings occurs; it is suggested here that there is a greater need to understand the ways in which individuals relate and disclose information to agents. Whilst understanding of student-agent interactions has improved significantly in recent years, as this section has shown, research has failed to distinguish between practices across different disciplines. Further, there has been little awareness of how the context impacts upon student-agent interactions and student willingness to disclose information to a pedagogical agent.

Evidence has shown that many users are not only comfortable interacting with high-quality pedagogical agents, but that an emotional connection can be developed between users and pedagogical agents, resulting in a more positive engagement experience. These findings should be considered in relation to the work of Lessler and O-Reilly (1997), who, amongst others, have found that self-administered surveys can yield more truthful responses than interview methods. This is particularly so when respondents are reporting on sensitive, personal or intricate information. Furthermore, as early as 1977 Lucas et al found that a computer developed for eliciting evidence about alcohol consumption was found to be highly acceptable by patients and they appeared to report more honestly to it than to their psychiatrists (Lucas, Mullin, Luna & McInroy, 1977). Hasler, Touchman and Friedman (2013) found, in a comparison of human interviewees with virtual world chatbots (pedagogical agents in non-learning situations), that chatbots and human interviewees were equally successful in collecting information about their participants' real live backgrounds. Pedagogical agents, being neither human interviewees nor text-based surveys, therefore pose an interesting opportunity for the educator seeking to facilitate student discussion of sensitive topics.

The disclosure of information, especially of sensitive information, requires the formation of a trust relationship (Wheeless & Grotz, 1977). Corritore, Kracher and Wiedenbeck (2003) propose that websites can be the objects of trust, in which trust is "an attitude of confident expectation that one's vulnerabilities will not be exploited" (2003:70). For them, the concepts of risk, vulnerability, expectation, confidence and exploitation play a key role in information disclosure in an online environment. It would appear that such findings can also be applied to pedagogical agent situations. This emotional connection has been found to be one of the strongest determinants of a user's experience, triggering unconscious responses to a system, environment or interface (Éthier, Hadaya, Talbot & Cadieux, 2008). These feelings strongly influence our perceptions, enjoyment, and pleasure, and influence how we regard our experiences at a later date. Emotional design at the basic level involves minimising common emotions related to poor usability such as boredom, frustration, annoyance, anger and confusion. Dennerlein, Becker, Johnson, Reynolds and Picard (2003) stated that during a computer task, systems usability may play a role in creating stressful situations that manifest themselves into various exposures to biomechanical stressors. Thus emotional design should also focus on invoking positive emotions associated with acceptance of the system and continued usage (such as inspiration, fascination, perception of credibility, trust, satisfaction, appeal, attachment).

Research studies

This paper reports on the findings of two studies designed to explore student experiences of engaging with pedagogical agents on sensitive topics. The studies sought to explore the extent to which pedagogical agents influence or affect a person's reactions and responses with regards to truthfulness, disclosure and personal engagement. In these mixed-method studies, we adopted Stake's (1983) responsive evaluation methodology, a pragmatic approach in which attention is given to the information and issues that those involved in the evaluation want to know about and the questions to which they want answers. This evaluation has been used to study an organization or curriculum in such a way as to contribute to a review of policy and/or decision-making within the organization. Therefore evaluation here is undertaken in relation to specific situations, contexts and questions.

In both studies, students were asked to respond to a *Student Life* survey about student lifestyle topics (finances, alcohol, plagiarism, drugs and sexual health). The survey was designed to increase levels of sensitivity over time and across subjects, meaning that finances, alcohol, and plagiarism were perceived as less sensitive while drugs and sexual health were perceived as more sensitive topics.

First study

Twelve students (m = 4; f = 8, both postgraduate and undergraduate) were recruited to participate in the first research study. This study sought to explore the potential influence of pedagogical agents on sensitive topics. Data were collected through the following methods:

- A thirty minute online questionnaire with an interactive pedagogical agent. Students were able to choose a pedagogical agent from a selection of eight agents, who varied in age, gender, and ethnicity. Students were asked questions on the topics listed above and coded data were analysed using non-parametric methods.
- A face-to-face interview on the topic of the student's experience of using the interactive pedagogical agent (average 30 mins). The interview was audio recorded and transcribed for data analysis purposes. Analysis was undertaken through an interpretive interaction approach (Denzin, 1989), which illustrated findings from the quantitative data and provided further insight into the student experience of engaging with the pedagogical agent on sensitive topics.

The student responses that were obtained when engaging with the pedagogical agent were coded to reveal how many idea units were included in their responses. Students' answers to the four substantive questions under each of the five topic areas (finances, plagiarism, sexual health, drugs and alcohol) were included in this analysis. For comparison purposes, the students' answers to three questions from the post-pedagogical agent interview were also coded for idea units.

Data were examined to assess if there were differences in disclosure levels between students who stated that they had been more truthful in their responses to the pedagogical agent than they would have been to an interviewer, and respondents who stated that their truthfulness would be unaffected by the format of questioning. When the two groups of participants were compared (more truthful vs same level of honesty) on the quantity of information produced in their responses to the two interview formats, an interesting pattern was revealed (see Figure 1). As can be seen, there was no difference across interview formats in the number of idea units produced by the group who believed they

were more truthful when they interacted with the pedagogical agent ($Z = -1.572$, $p = 0.116$, $d = 1.019$). However, the group who believed that their truthfulness was not influenced by the interviewer type produced significantly more detailed responses when they interacted with the interviewer than when they interacted with the pedagogical agent ($Z = -2.023$, $p = 0.043$, $d = 1.439$). It is important not to conflate detailed answers with truthful answers, but it would seem that for these respondents, interacting with a human interviewer resulted in more detailed interaction.
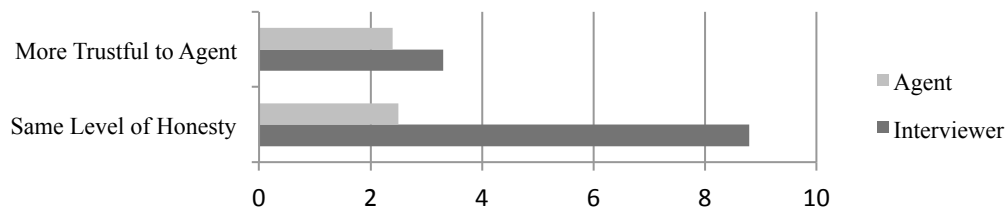

Figure 1: Students' levels of disclosure to questions, by 'truthfulness' groups

The post-survey interviews provided additional insight into these findings. Three key themes were identified, which were truthfulness, emotional engagement, and personalisation. With regard to truthfulness, students oriented their discussions around the *realism* of the pedagogical agent, comparing it to either a person, or to a questionnaire. For example, some students were emboldened by their awareness that the pedagogical agent was not real, and felt able to share truthful answers to sensitive questions without impressions of judgement, which they had come to expect in face-to-face discussions on sensitive topics. Others suggested that they disclosed more truthful information to a pedagogical agent because it invoked impressions of social presence, such as feeling present in an environment with another individual. For some, this emerged from feeling more confident discussing sensitive topics with a 'person'; other students suggested that they would have felt guilty if they had not disclosed truthful information because of the agent's realism. Truthfulness, then, was closely related to students' emotional engagement with the pedagogical agent, in which students felt a personal connection to the agent. Findings here suggested that the greater the emotional engagement, the more positive the experience. The final theme revealed the importance of personalisation in student-agent interactions, which focused upon using pedagogical agent technology to accommodate the differences between individuals and provoke and encourage choice.

Findings in relation to perceptions of the personalisation of the chatbot suggested that students' emotional engagement with pedagogical agents was influenced by the degree to which they were able to personalise or relate to the pedagogical agent. Consequently, students' abilities to trust the pedagogical agent and thus disclose truthful information were informed by the degree to the pedagogical agent was personalised. Within the second study, particular attention was paid to the ways in which the personalisation of the chatbot and the emotional engagement with the chatbot might influence notions of trust and truthfulness in relation to pedagogical agents.

Second study - methodology
Based on the findings of the first study, the second study explored the effect of using agents under different survey conditions, on levels of information disclosure. This study sought to

provide additional insight into the trends identified in the first study by answering the following research questions:

1. Does changing the level and condition of contact with the pedagogical agent influence the level of disclosure?
2. To what extent do perceptions of the pedagogical agent influence the relationships formed with the agent?

The findings presented here are from a larger study exploring the influence of pedagogical agents. Four different survey conditions were applied, in which students responded to an adapted version of the *Student Life* survey from the first study.

1) A non-agent survey delivered in one session.
2) An identical agent-based survey delivered in one session (short-term agent).
3) An identical agent-based survey delivered across five sessions (long-term agent, exploring length of engagement).
4) An agent-based survey delivered across five sessions and including additional 'icebreaker' questions at the start of each session (long-term agent with icebreakers, exploring depth of engagement).

In condition four, students were asked three initial questions at the beginning of their first interaction with the pedagogical agent. Their responses were then used to drive questions at the beginning of each of the four subsequent interactions. The agent was able to ask both social questions (how are you, how's your week, what are you doing today, what course are you studying, are you going on holiday soon, how is your family) and interest questions (on sports, film, television, and music). For example, if the student expressed an interest in certain types of films at the beginning of the first session, the pedagogical agent would then ask if they had seen a recent film release at the beginning of a subsequent session. The purpose of these questions was to help facilitate a rapport between the pedagogical agent and the student and, in doing so, increase the depth of engagement in the interaction.

All students were asked to complete the *Student Life* survey twice, both with and without a pedagogical agent. All students completed survey condition one, and then additionally either condition two, three or four; thus three groups were formed. The outcome variable was the level of disclosure measured by words typed in response to each survey. 117 participants were recruited from both undergraduate and postgraduate courses across a variety of disciplines and provided complete data through completion of both the agent and non-agent versions of the questionnaire (see Table 1).

Group	Male	Female	Mean age (standard deviation)
1. Non-agent and short term agent (standard deviation)	5	28	20.15 (3.50)
2. Non-agent and long term agent (standard deviation)	6	32	22.79 (7.5)
3. Non-agent and long term agent with	6	40	20.72

icebreakers (standard deviation)			(5.36)
Overall (standard deviation)	17	100	21.23 (5.8)

Table 1: Age and gender of participants involved in the study

A website was designed by *Daden Ltd.* to allow students to access an online questionnaire which was comprised of 3 demographics questions and 24 items relating to the student lifestyle topics of finances, alcohol, plagiarism, drugs and sexual health. The website presented the questions first and then provided some information regarding that topic, before moving onto the next topic. When participants were using the agent version of the questionnaire, they were given an option to pick from ten agents who ranged in age, gender, ethnicity and appearance. Agent voices were chosen to match appearances, but were computer-synthesized. The figure below provides an example of the website when completing an agent-based survey, which is no longer operational.
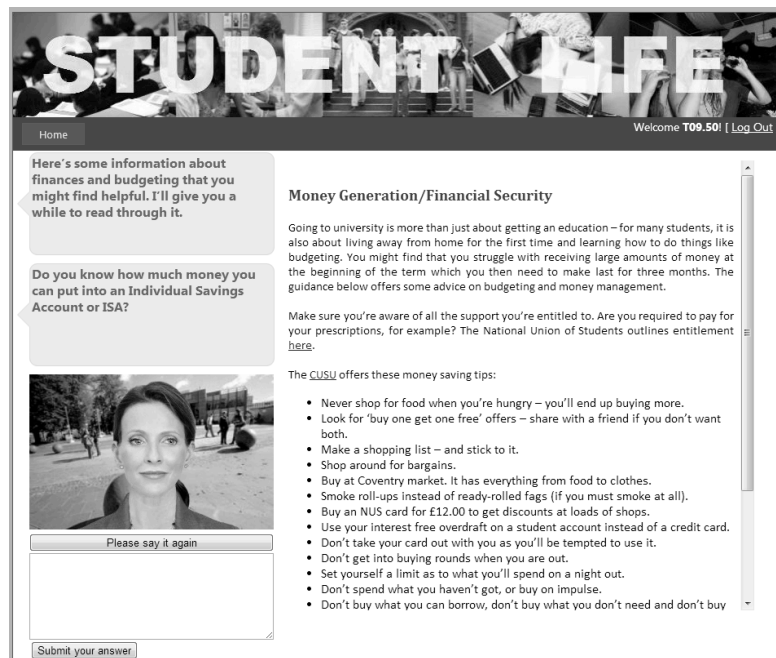

Figure 2: Example of a pedagogical agent survey on the Student Life website

There was a gap of between two and three weeks between completing the agent and non-agent surveys. The order in which the three groups completed the agent and non-agent surveys was randomised so as to reduce order effects. After completing the agent-based survey, students then completed a short follow-up survey based at Bristol Online Surveys (http://www.survey.bris.ac.uk/). This asked students about their likes and dislikes with regard to the agent, and asked them to explain their pedagogical agent choice.

Quantitative findings from the second study suggested that for sensitive topics, individuals were more likely to divulge more information to the pedagogical agent than to a standard questionnaire. On the topics of sexual health and drug use (which were deemed to be more sensitive than the other topics of financial management, alcohol, and plagiarism), individuals divulged more information when engaging with the pedagogical agent over a

longer period of time. However, this effect was negated by the inclusion of the additional engagement questions. The following section outlines these findings and the methods by which they were reached.

Quantitative and survey findings
Non-parametric tests were used, due to the data not satisfying the assumptions of parametric statistical tests. Data from individual groups were found to deviate from a normal distribution in terms of both skewness and kurtosis. Transformation of the data was attempted; however, doing so did not rectify the problems of normality.

	Finances	**Alcohol**	**Plagiarism**	**Drugs**	**Sex**	**Overall**
Group 1 *Compares short-term agent condition to non-agent survey condition*						
Participant *n* =	33	33	33	33	33	33
No agent	31	27	37	62	19	184
agent	35	27	39	67	18	195
Group 2 *Compares long-term agent condition to non-agent survey condition*						
Participant *n* =	38	37	37	37	31	31
No agent	42	36	43	**72**	**16**	220
agent	34	32	48	**80**	**22**	242
Group 3 *Compares long-term agent with additional engagement questions condition to non-agent survey condition*						
Participant *n* =	46	46	46	44	38	38
No agent	**38**	**34**	41.5	60	20	216
agent	**28**	**26**	42	69	17.5	197

Table 2: Median word length of the answers given by participants in response to online Student Life surveys

Figures 3 and 4 (below) illustrate the differences between agent and non-agent responses in the two long-term agent groups.
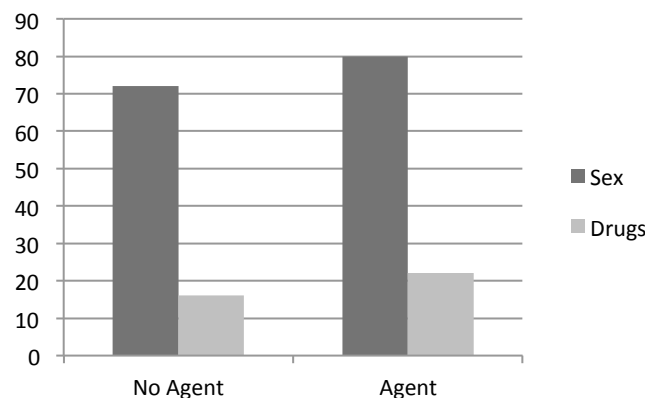

Figure 3: Median number of words for the responses to sexual health and drug related topics in Group 2
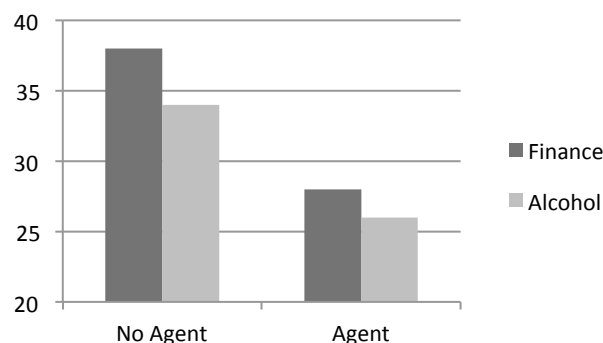
Figure 4: Median number of words for the responses to financial and alcohol related topics in Group 3

Wilcoxon Signed Ranks tests were conducted to examine the differences further. Participants in the long term agent group responded with longer answers to questions from an agent instead of a non-agent, when discussing drugs ($Z = 2.023$, $p < 0.05$, $r = 0.238$) and sexual health ($Z = 2.952$, $p < 0.05$, $r = 0.375$). However, participants in this same group responded with shorter answers to an agent than a non-agent when discussing finances ($Z = 1.956$, $p < 0.05$, $r = 0.202$) and alcohol ($Z = 2.475$, $p < 0.05$, $r = 0.255$).

The Revised Self-Disclosure Scale (Wheeless & Grotz 1976) was also used to assess the inherent levels of trust and self-disclosure (depth and amount subscales) in the participants when talking about themselves. When looking at the data as a whole and also subdivided into groups based on the type of questionnaire that was completed, data suggested that there was no correlation between the self-disclosure subscales and the amount of information disclosed to the pedagogical agent. There was no correlation between the findings of this survey and the amount students' disclosed to the pedagogical agent. Nor did there appear to be a link between the responses to this survey and students' qualitative responses when sharing their perceptions of the trustworthiness of the pedagogical agent. There are two possible reasons for this finding. Firstly, it may be that the measure chosen was not suitable for this particular study. Alternatively, these findings might suggest that truthfulness is not as closely related to self-disclosure as has been posited in other studies such as Steel (1991), who indicated that for there to be self-disclosure there must first exist some level of trust. Instead, the key factor in disclosure to the pedagogical agent would seem to be the student-agent relationship formed, rather than inherent willingness to trust on the part of the user. Consequently, subsequent surveys designed for the purpose of this study paid particular attention to student perceptions of the pedagogical agent.

Students' choices of agent were attributed to four key factors: random choice, comfort, appearance, and relatability. Nearly a quarter of students (22%) stated that their choice of agent had been 'random choice'. However, 30% of students seemed to seek out familiar situations in which they might feel comfortable disclosing information. Most of these responses stated that the agent was chosen because of a 'friendly' appearance, whereas others commented that they chose an agent of the opposite gender as they felt most comfortable in those situations. Other responses noted that the agent's seeming professionalism – and associated qualities of confidentiality and lack of judgement – were

particularly important, whilst a group of students chose the agent because of a sense of familiarity, for example because they resembled an uncle, or looked like a friend, or shared the same name as a respondent's boyfriend.

A further 30% of students chose the pedagogical agent because of its appearance: either because it conformed to their perceptions of attractiveness, or because it looked unrealistic – one cartoon-type design of agent was provided whilst the rest used photo-realistic designs. Here, most students commented that speaking to a photo-realistic agent seemed incompatible with their understanding that they were speaking to a computer. The final choice of agent was attributed to students' ability to relate directly to the pedagogical agent. Twenty-two per cent of students stated that they could relate to the agent, either because they shared the same name, or were of a similar age. The largest category of students in this group stated that they chose the pedagogical agent because it was similar in appearance to them, for example because the agent most closely represented their race or ethnicity. Many of these findings identified the pedagogical agent as a trustworthy entity; either as an individual associated with professionalism and confidentiality, or as an individual in the student's life deemed to be approachable and relatable.

These findings were replicated in a further survey which asked participants to describe their experience of using the agent by selecting several words that best described how they felt about the avatar that they had chosen (see Figure 5 below). Over 50% described their chosen pedagogical agent based on the perception of it being 'friendly', 'professional' or 'approachable'.
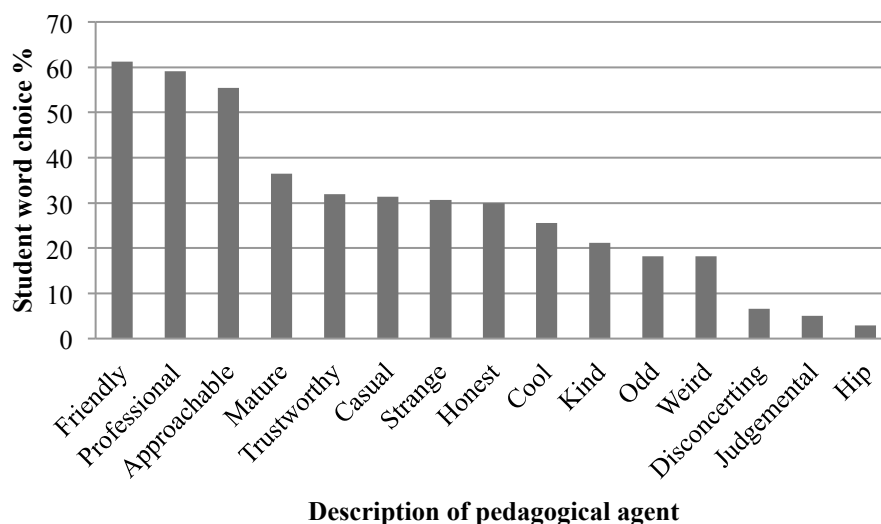
Figure 5: Terms used to describe the pedagogical agent

Qualitative interview findings

After interacting with the pedagogical agent, seventeen students across groups 2 and 3 participated in semi-structured interviews lasting 15 minutes on average. These interviews focused on students' experiences of interacting with the pedagogical agent and the difference between the agent and non-agent surveys. The findings are presented here with

pseudonyms representing students' genders and cultural backgrounds. Results supported the findings from the first study, with several students suggesting that they answered more honestly due to feelings of guilt in disclosing dishonest information to a 'real person', or reacting to a perceived lack of judgement from the pedagogical agent. This study's focus on the increased length and depth of the engagement also provided additional insight into the themes of emotional engagement and personalisation. For example, conversational flow between the student and pedagogical agent played an important role in emotional engagement, as one student commented:

> *"[The pedagogical agent] kind of built up that fake rapport beforehand, which was quite interesting. Like questions, when going onto the website, you wouldn't expect to be asked those things. So it got you thinking and it got you engaged."* (Lauren)

Here, Lauren seemed to have made certain assumptions about the pedagogical agent based upon her experience of previous surveys, suggesting that the survey would be on-topic and not seek to facilitate an emotional connection. In asking the additional icebreaker questions, the agent failed to comply with those expectations and instead provoked a sense of 'difference', capturing Lauren's attention. However, the inclusion of the additional icebreaker questions did not facilitate increased *emotional* engagement and thus trust in this situation. Lauren was cognizant of the fact that the pedagogical agent was a computer program and thus, to her, these questions conveyed a compliance with conversational norms yet failed to facilitate the emotional engagement that typically follows in interpersonal interactions. For other students such as Monica, however, these questions enabled them to feel as though they were conversing with a friend:

> *"This lady was talking to me, asking me questions, generally asking me how I am. And do I like a holiday. You know, as you would do, talking friendly with someone before you're gonna speak to them about a particular situation. And that's exactly what they did …. You felt like you were talking to someone. And even though you knew they weren't, and I'm an intelligent woman, obviously I knew, but I felt very relaxed, because I was talking to someone."* (Monica)

Whilst Monica, like Lauren, was aware of the fact that the pedagogical agent was not a person, its observance of conversational norms did increase her sense of emotional engagement. Monica appeared to feel that she was present in a panoptical space, in which someone else was listening and cared about her responses. Similarly, Emily commented that *"even though you weren't talking to a person, you were"*. The pedagogical agent thus encompassed a kind of in-between space in which something or someone else was engaging in the interaction, but also was not. Here, students were responded to the seemingly disparate cues – the realistic appearance and actions of the pedagogical agent, versus its presence on a computer screen and their prior knowledge that it was not a person. Yet for some students, such as Kate, the pedagogical agent's attempt to engage in off-topic conversation was confusing or even insulting, due to their knowledge that the pedagogical agent *was* an agent:

"The questions at the beginning are just patronising and pointless. I know I'm talking to a computer program so small talk is just an insult of one's intelligence and not at all warming or positive." (Kate)

For Kate, the trust which she may have placed in the pedagogical agent was damaged by the perceived deception incurred by the inclusion of the additional icebreaker questions. Whilst Kate recognised the purpose behind the use of the icebreaker questions - to create a warming and positive environment – she resented what she perceived as an attempt to trick her into experiencing those emotions. Thus in this case, the attempt to improve the emotional engagement between the pedagogical agent and the student had the opposite effect.

The appearance and personalisation of the pedagogical agents was also found to inform the emotional engagement between agent and student:

"I chose an Indian woman because I am also Indian so I could relate to her." (Priti)

"I preferred the English accent on the agents and stuck with one with a friendly face and female because I am female and I think I could relate to her better than a man." (Amelia)

"[My agent] seemed like quite a sensible guy. I thought, you can trust guys, so I'll pick him. I prefer speaking to guys, so I was like, 'I'll pick a guy'." (Kate)

"I prefer talking to a female in these circumstances." (Mark)

Students tended to choose pedagogical agents with whom they could relate and trust, as has already been shown. In many cases, students made distinctions based upon gender and ethnicity.

Discussion

The findings from the main study suggest that both the topic under discussion and the length of engagement with the agent are important in determining information disclosure levels. In the main study, students who answered the *Student Life* survey over a period of two weeks (but did not answer the off-topic questions designed to increase depth of engagement) disclosed significantly more information to the agent when discussing sexual health and drug use, but not on the financial, plagiarism, and alcohol questions. These findings would seem to suggest that a stronger emotional connection – leading to increased disclosure – might be forged over longer periods of time. No such significance was found in students who answered the survey designed to increase both length and depth of engagement, or in students who answered the agent survey in the one-off session, although similar trends can be observed. Interviews undertaken by Kim (2007) have yielded data indicating that agent characteristics of 'friendliness' are particularly important, as illustrated by our findings in Figure 6. Veletsianos and Miller (2008) have speculated that longer-term interactions with an agent may have a positive effect on emotional engagement. Thus far, empirical studies of agents have neglected this area; these findings, although preliminary,

reinforce the argument that examinations of longitudinal agent engagement are essential (Veletsianos & Russell, 2014).

Findings from the main study also suggest that the increased disclosure in the long-term agent setting might be negated by the inclusion of the 'icebreaker' questions designed to increase the depth of interaction. This demonstrates that the inclusion of these questions certainly has an effect, but perhaps not the desired one. We have identified three possible reasons for this finding. Firstly, Oguchi (1991) has found that self-disclosure depends upon the goal one sets out to achieve in an interaction. As students in this study were informed that they were participating in a survey on student lifestyle topics, the inclusion of unexpected and additional questions may have influenced their levels of self-disclosure. Further, early and more recent work in human-computer interaction suggests that it is important not to over play any implications of artificial intelligence, as the human user would inevitably be disappointed and thus ultimately respond negatively to the software (for example Gaines & Shaw, 1984; Culley & Madhavan, 2013). Lessons from the Turing tests would seem to be particularly important here, in that chatbots have been found to be most effective when designed to fill specific roles. Qualitative findings from this study suggest that the chatbot's consistency with or divergence from expected conversational roles can be particularly important in the emotional engagement between pedagogical agents and students.

Secondly, Culley and Madhavan have cautioned that "as the agent becomes increasingly morphologically similar to a human, it is also likely that operators will engage in correspondence bias more frequently by ascribing human motivations, reasoning abilities and capabilities to this non-human system" (2013: 578). Consequently, the inclusion of questions designed to facilitate increased engagement might have resulted in students feeling a sense of talking to a person and thus perhaps being less willing to disclose information. We derive this explanation from the findings of first study in which qualitative responses indicated that student willingness to disclose sensitive information to an agent was attributed partially to the agent being *almost* like a person. As the long-term agent was designed to create an emotional connection by asking off-topic questions, it may have shifted in student perceptions to seeming more humanistic and thus likely to seem 'judgemental' of responses to sensitive questions.

Thirdly, the lessening of student disclosure in this particular setting might also be attributed to split-attention effect and a lack of conversational realism. The agent in the main study asked how students were, but was unable to respond when asked "How are you?". Students' qualitative responses to this would seem to reinforce arguments by Nass and Yen (2012) that humans respond to computers in the same way they do to humans, and thus form the same expectations of conversational interaction. Agent ability to respond to conversational norms and engage in bridging topics has been identified as particularly important in users' perceptions of realistic conversations and this should also be considered in future studies. This study suggests that particular attention should be paid to design when seeking to facilitate increased depth of engagement, and that emotional engagement and the way that the chatbot is personalised is particularly important in sensitive settings. Whilst this study focused on student lifestyle issues, we contend that the same might apply in counselling or healthcare educational situations. Earlier studies on disclosure to agents have largely failed to consider the influence of the sensitivity of questions; our findings suggest that different topics even across the same study can yield different levels of disclosure.

In the wider context of this study, what seems to be increasingly evident is that many people who use social networking sites choose to be what Westin et al (1991) have termed privacy pragmatists, in that they are prepared to share information for personal gain, such as by using Facebook, to maintain networks with friends and family. However, more recent research by Brandtzæg et al (2010), compared the experiences and usage of younger and older Facebook users; they found that all age groups had diverse friends and acquaintances, but that younger users were more skilled in their Facebook usage, and that both groups displayed different open public profiles.

We suggest that research into pedagogical agents is too limited and currently based entirely on social theories of learning, so that there is a need to consider the socio-cultural implications of their use in education, and draw on the work of theorists such as Frude (1983). In this study, students disclosed significantly more information on sensitive topics when engaging with the pedagogical agent over a longer period of time. These findings would seem to suggest that Frude's goal of the formation of beneficial relationships between humans and computers is not insurmountable, and that these beneficial relationships might be best formed in circumstances in which human to human interaction is discomforting, such as on sensitive topics. This is particularly important since it would seem that such technologies are likely to become a part of students' daily lives outside the educational arena, and that as we learn more about the areas in which technology can be effective, the relationship between humans and computers may shift and adapt in unknown ways.

In the 21st century we lead augmented lives, lives that can be changed and enhanced through technology and surgery. Julian Huxley developed the concept of augmented reality technology and transhumanism, suggesting that augmented reality technology (1927; 1957) would enable the human race to develop itself through science and technology. For Aldous Huxley (1927) education was the road to emancipation, but this was blocked by unalterable bio-genetical profiles, resulting in humans being caught between different psycho-biological capabilities and the requirements of society. The debates still, to a large extent, focus on posthumanism and transhumanism. Transhumanists believe that human enhancement technologies should be made widely available, that individuals should have broad discretion over which of these technologies to apply to themselves, and that parents should normally have the right to choose enhancements for their children-to-be. Posthumanism posits the idea that there are new forms of human existence which blur the boundaries between humans, nature and machines, suggesting an ideal situation in which the limitations of human biology are transcended, replaced by machines. Indeed Haraway argued for the cyborg, "we are all chimeras, theorized and fabricated hybrids of machine and organism; in short, we are cyborgs. The cyborg is our ontology; it gives us our politics" (Haraway, 1985: 118). However, as Herbrechter argues:

> "The temptation has therefore been to see posthumanism as the 'natural' successor – in analogy with the popular idea that AI, cyborgs or digital machines function as successors to the human species – to the still too humanist postmodernist/poststructuralist paradigm. Which means of course that the poststructuralist theory responsible for the birth of this posthumanism supposedly merely has a 'midwife' function and thus needs to be 'overcome'."
>
> (Herbrechter, 2013, p. 328)

Whether we are cyborgs or not, our existence is augmented, and our responses to 'machines' increasingly illustrate that we are prepared to trust them and reveal sensitive information to them (Savin-Baden et al, 2014). For example, Hasler et al. (2013) found, in a comparison of human interviewees with virtual world chatbots (pedagogical agents in non-learning situations), that chatbots and human interviewees were equally successful in collecting information about their participants' real live backgrounds. Although students in this study were increasingly comfortable in disclosing information to pedagogical agents, they still imposed human expectations and requirements on the pedagogical agent. However, whilst technology may be progressing towards the kind of transhumanism considered by Huxley (for example, people having chips embedded in themselves), participants in this study subscribed to the human-technology binary and were seemingly troubled by attempts to step away from this, for example, by the inclusion of the additional engagement questions. Thus we argue that there is a need to consider the social and ethical implications of pedagogical agent use, and suggest arguments that technology enhanced learning is more about technology than learning is somewhat misplaced. The number of projects funded by the European Union over the last 10 years might seem to suggest this is the case, as Bayne argues:

> "'TEL', far from being an unexceptionable and neutral term simply in need of clearer definition, is in fact a deeply conservative discourse which reduces our capacity to be critical about digital education, and fails to do justice equally to the disruptive, disturbing and generative dimensions of the academy's enmeshment with (digital) technology"

> (Bayne, 2014, p. 348)

However this stance is based on Biesta's (2006) critique of the 'learnification' of education. Biesta suggests that a focus on learning 'makes it far more difficult, if not impossible, to ask the crucial educational questions about content, purpose and relationships' (Biesta, 2012: 36), but does not really explain why this might be the case. We suggest that Biesta's ideals do not fit well with the digital age; rather they are suggestive of teachers who feel the need to teach people something and to manage the content that they think needs to be taught.

Limitations

Certain limitations of this study need to be taken into account. Firstly, we have only measured disclosure levels, i.e. number of words disclosed in the agent and non-agent settings. From that, we are unable to derive the quality and truthfulness of the agent-student interaction. Future work will focus upon the content of these interactions and consider qualitative comparisons of the agent and non-agent responses. Secondly, there was a high female bias in the main study sample, meaning that conclusions cannot be drawn based upon gender. Future studies should also examine the trend towards increased disclosure from men, identified in the first study.

Based upon the findings from both studies, seven key implications have been identified:

- The adaptivity of the system and emotional connection to the pedagogical agent are intrinsic to the student's belief that they can trust and therefore be more truthful. By

capitalising on an understanding of user emotions there is an opportunity to enhance the level of individual connection with the learning environment and the sense of immersion offered.

- The amount of information divulged was dependent on how well the participant engaged with the pedagogical agent. For example, one student wanted to divulge more information but felt rushed by the pedagogical agent body language and movements. Another did not divulge as much information as he did in a paper questionnaire, due to associating the pedagogical agent with having a real conversation and 'boring' it with talking too much.

- An emotional design philosophy will ensure the psychosocial features of the environments as well as physical and cognitive requirements. This emotional connection with the pedagogical agents would seem to heighten the sense of immersion and therefore it is argued, the disclosure potential.

- Learning and engagement using pedagogical agents provides opportunities for displaying, testing and responding to the emotions of self and others in a safe and non-threatening environment. This can be either subject specific emotional skills (for example empathy) or non-subject specific in the general sense of emotional intelligence.

- Despite the sense that when asked directly students did not feel pedagogical agents encouraged them to be more honest in reality, when questioned on particular aspects of interaction, they did in fact disclose more to pedagogical agents than they believed they had. This would seem to imply pedagogical agents encourage more disclosure than the students themselves believe they have actually disclosed.

- It is important not to conflate detailed answers with truthful answers. It might be expected that the 'more truthful to the pedagogical agent' group would have given more detailed answers to the agent than the interviewer.

- The perceived realism of an agent both in terms of its appearance and conversational style can negate the potential effectiveness of agents being used long term.

Conclusion

These studies indicated that students disclosed more information to the pedagogical agent on sensitive topics when engaging with the agent over a longer period of time (suggesting the increased length of engagement may have resulted in a trusting relationship). Further, there was no relationship between students' disclosure and the trust measure, which suggests that disclosure is less about inherent willingness to trust and more about the relationship between the agent and the student. The findings also suggested a need to perhaps include a 'socially desirable responding' measure to future studies in this area, to allow the unpicking of 'truthfulness' and 'detailed responding' more precisely in relation to these data. Yet it is also vital to balance this with the knowledge that students disclosed less information to the pedagogical agent when the agent attempted to facilitate increased depths of engagement. Whilst this finding may be explained by the agent's lack of success in shifting smoothly between different conversational topics, a difficulty noted in many unsuccessful Turing tests, or the expectations that students had formed prior to their interaction with the agent, it also brings to mind Frude's (1983) ideal of beneficial relationships between humans and computers. Findings from this study suggest that this goal might perhaps be achievable through sustained human-computer interaction, as

opposed to one-off Turing tests, but also that individual students will react differently to the suggestion of intimate human-computer interaction.

References

Bailenson, J. N., Yee, N., Blascovich, J. & Guadagno, R. E (2008) *Transformed Social Interaction in Mediated Interpersonal Communication*, in E. Konijn, M. Tanis, S. Utz. & A. Linden (Eds.) *Mediated Interpersonal Communication*. Lawrence: Erlbaum Associates.

Bayne, S. (2014) *What's wrong with 'Technology-Enhanced Learning'?*, in S. Bayne, C. Jones, M. de Laat, T. Ryberg & C. Sinclair (Eds) *Proceedings of the 9th International conference on Networked Learning.*

Berrar, D. P. & Schuster, A. (2014) Computing Machinery and Creativity: lessons learned from the Turing test, *Kybernetes,* 43(1), 82-91.

Biesta, G. (2006) *Beyond Learning: democratic education for a human future.* Boulder, Colorado: Paradigm Publishers.

Brandtzæg, P. B., Lüders, M. & Skjetne, J. H. (2010). Too Many Facebook "Friends"? Content sharing and sociability versus the need for privacy in social network sites, *Journal of Human-Computer Interaction*, 26(11-12), 1006-1030.

Corritore, C. L, Kracher, B. & Wiedenbeck, S. (2003) On-line Trust: concepts, evolving themes, a model, *International Journal of Human-Computer Studies,* 58(6), 737-758.

Culley, K. E. & Madhavan, P. (2013) A Note of Caution regarding Anthromorphism in HCI Agents, *Computers in Human Behavior,* 29(3), 577-579.

Dehn, D. & Van Mulken, S. (2000) The Impact of Animated Interface Agents: a review of empirical research, *International Journal of Human-Computer Studies,* 52(1), 1-22.

Dennerlein, J., Becker, J., Johnson, P., Reynolds, C. & Picard, R. W. (2003) Frustrating Computer Users Increases Exposure to Physical Factors. Paper presented at International Ergonomics Association, Seoul, Korea, 24-29 August. http://affect.media.mit.edu/pdfs/03.dennerlein-etal.pdf.

Denzin, N. (1989) *Interpretative Biography*. London: Sage.

Dunsworth, Q. & Atkinson, R. (2007) Fostering Multimedia Learning of Science: exploring the role of an animated agent's image, *Computers & Education,* 49(3), 677-690.

Éthier, J., Hadaya, P., Talbot, J. & Cadieux, J. (2008) Interface Design and Emotions Experienced on B2C Web Sites: empirical testing of a research model, *Computers in Human Behavior,* 24(6), 2771-2791.

Frude, N. J. (1983) *The Intimate Machine.* London: Century.

Frude, N. J. (1984) *The Robot Heritage.* London: Century.

Frude, N. & Jandrić, P. The Intimate Machine – thirty years on, *E-Learning and Digital Media,* *XXX.*

Gaines, B. R. & M. L. G. Shaw (1984) *The Art of Computer Conversation: a new medium for communication*. Englewood Cliffs, New Jersey: Prentice/Hall International.

Garau, M., Slater, M., Vinayagamoorthy, V., Brogni, A., Steed, A. & Sasse, M. A. (2003) The Impact of Avatar Realism and Eye Gaze Control on Perceived Quality of Communication in a Shared Immersive Virtual Environment, in G. Cockton & P. Korhonen (Eds) *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: ACM Press.

Gong, L. (2008) How Social is Social Responses to Computers? The function of the degree of anthromorphism in computer representations, *Computers in Human Behavior,* 24(4), 1494-1509.

Haraway, D. (1985/1991) A Cyborg Manifesto: science, technology, and socialist-feminism in the late twentieth century, in D. Haraway (Ed) *Simians, Cyborgs and Women: The Reinvention of Nature*. New York: Routledge.

Hasler, B. S., Tuchman, P. & Friedman, D. (2013) Virtual Research Assistants: replacing human interviewers by automated avatars in virtual worlds', *Computers in Human Behavior,* 29(4), 1608-1616.

Heidig, S. & Clarebout, G. (2011) Do Pedagogical Agents Make a Difference to Student Motivation and Learning?, *Educational Research Review,* 6(1), 27-54.

Herbrechter, S. (2013) Posthumanism, Subjectivity, Autobiography, *Subjectivity*, 5, 327–347.

Huxley, A. (1927) *Proper Studies.* London: Chatto and Windus.

Huxley, J. (1927) *Religion without Revelation.* London: E. Benn.

Huxley, J. (1957) *Transhumanism.* London: Chatto and Windus.

Kiran, A. H. & Verbeek, P. (2010) Trusting Our Selves to Technology, *Knowledge, Technology and Policy,* 23(3-4) 409- 427.

Kim, Y. (2007) Desirable Characteristics of Learning Companions, *International Journal of Artificial Intelligence in Education, 17*(4), 371-388.

Kim, Y. & Baylor, A. L. (2006) A Social-cognitive Framework for Pedagogical Agents as Learning Companions, *Educational Technology Research and Development,* 54(6), 569-596.

Kim, Y. & Wei, Q. (2011) The Impact of Learner Attributes and Learner Choice in an Agent-based Environment, *Computers & Education,* 56(2), 505-514.

Lee, E. (2010) The More Humanlike, the Better? How speech type and users' cognitive style affect social responses to computers, *Computers in Human Behavior,* 26(4), 665-672.

Lessler, J., T. & O'Reilly, J., M. (1997) Mode of Interview and Reporting of Sensitive Issues: design and implementation of audio computer assisted self-interviewing, in L. Harrison & A. Hughes (Eds) *The validity of self-reported drug use: improving the accuracy of survey measurements*. Rockville: National Institute of Drug Abuse.

Lucas, R. W., Mullin, P.J., Luna, C. B. X & McInroy, D. C. (1977) Psychiatrists and Computers as Interrogators of Patients with Alcohol-Related Illness: a comparison, *British Journal of Psychiatry*, 131, 160-167.

Moreno, R. & Flowerday, T. (2006) Students' Choice of Animated Pedagogical Agents in Science Learning: a test of the similarity-attraction hypothesis on gender and ethnicity, *Contemporary Educational Psychology,* 31(2), 186-207.

Nass, C. & Yen, C. (2012) *The Man who Lied to his Laptop*. New York: Penguin Books.

Oguchi, T. (1991) Goal-based Analysis of Willingness of Self-disclosure, *Japanese Psychological Research*, 33(4), 180-187.

Payr, S. (2003). The Virtual University's Faculty: an overview of educational agents, *Applied Artificial Intelligence,* 17(1), 1–19.

Savin-Baden, M., Tombs, G., Burden, D. & Wood, C. (2013) 'It's Almost Like Talking to a Person': student disclosure to pedagogical agents in sensitive settings, *International Journal of Mobile and Blended Learning,* 5(2), 78-93.

Stake, R. (1983) *Responsive Evaluation*, in T. Husén & T. N. Postlethwaite (Eds) *International Encyclopaedia of Education: research and studies.* New York: Pergamon Press.

Steel, J. L. (1991) Interpersonal Correlates of Trust and Self-disclosure, *Psychological Reports*, 68(2) 1319-1320.

Turkle, S. (2011) *Alone Together*. New York: Basic Books.

Turing, A. M. (1950) Computing Machinery and Intelligence, *Mind*, *49*(236), 433–460.

Veletsianos, G. & Miller, C. (2008) Conversing with Pedagogical Agents: a phenomenological exploration of interacting with digital entities, *British Journal of Educational Technology*, 39(6), 969-986.

Veletsianos, G. & Russell, G. (2014) *Pedagogical Agents*, in M. Spector, D. Merrill, J. Elen & M. J. Bishop (Eds) *Handbook of Research on Educational Communications and Technology*, 4th ed, 759-769.

Westin, A. & Harris, L. and Associates (1991) *Harris-Equifax Consumer Privacy Survey*. Atlanta, GA: Equifax Inc.

Wheeless, L., & Grotz, J. (1976) Conceptualization and Measurement of Reported Self-Disclosure, *Human Communication Research*, 2(4) 338-346.

Wheeless. L. & Grotz, J. (1977) The Measurement of Trust and its Relationship to Self-disclosure, *Human Communication Research*, 3(3), 250-257.

Whitty, M. T. & Joinson, A. N. (2009) *Truth, Lies, and Trust on Internet*. London: Routledge.

Yee, N. (2006) The Demographics, Motivations and Derived Experiences of Users of Massively Multi-user Online Graphical Environments, *Presence,* 15(3), 309-329.