

# **Pedagogical Agents: Pedagogical Interventions via Integration of Task-oriented and Socially Oriented Conversation**

Annika Silvervarg

*Linköping University, Sweden*

Magnus Haake

*Lund University, Sweden*

Lena Pareto

*University West, Sweden*

Betty Tärning

*Lund University, Sweden*

Agneta Gulz

*Linköping University and Lund University, Sweden*

Please address all correspondence to:

Agneta Gulz

Department of Computer Science, Linköping University, Sweden &  
Lund University Cognitive Science, Sweden

phone: +46-46-2223269

fax: +46-46-2224424

agneta.gulz@liu.se; agneta.gulz@lucs.lu.se

Paper presented at the  
Annual Meeting of the American Educational Research Association  
New Orleans, April 8-12, 2011

# **Pedagogical Agents: Pedagogical Interventions via Integration of Task-oriented and Socially Oriented Conversation**

Annika Silvervarg

*Linköping University, Sweden*

Magnus Haake

*Lund University, Sweden*

Lena Pareto

*University West, Sweden*

Betty Tärning

*Lund University, Sweden*

Agneta Gulz

*Linköping University and Lund University, Sweden*

*The paper discusses the motivation for and outcome of the addition of socially oriented so called “off-task” conversational abilities to an existing “teachable agent” (TA) in an educational mathematics game. The purpose of the extension is to affect constructs known to promote learning, such as self-efficacy and engagement, as well as enhancing students’ experiences of the game. A comparison of students that played the educational game with the off-task interaction included to those who played without it, indicate that the former had a more positive experience of the game, and that they also learnt more in the sense of teaching their TA better. The potential for pedagogical interventions in this and similar systems is discussed as well as differences found between high- and low-achievers.*

By conversational pedagogical agents, CPAs, we refer to computer-generated characters in a pedagogical context that engage in spoken or written conversation with students. Some CPAs may use non-verbal conversational channels, such as gestures and facial expressions, but this paper limits itself to conversation in written language. Various research groups have developed CPAs for different domains like physics, mathematics, foreign languages, programming, and many others, and several evaluative studies have shown that CPAs can be effective as tutors (e.g. Graesser, Chipman, Haynes & Olney, 2005; Kumar, Gweon, Joshi, Cui & Rose, 2007).

There is a large variety in how CPAs are designed and what specific pedagogical strategies they exploit, but all engage, in some way or other, in *task-oriented conversation* such as: elaborating on students' answers, asking questions regarding the domain or task, correcting misconceptions, asking students to elaborate on their examples, providing hints and directions. In other words, they engage in conversation that clearly pertains to the learning material and tasks in question.

During the past decade some researchers have developed CPAs that in addition to carrying out task-oriented conversation engage in *relation oriented or socially oriented conversation* with students, i.e. conversation with no (apparent) relation to the learning material and tasks. Examples are: reassuring or cheering up a student, carrying on small-talk, engaging in mutual self-disclosure. Relational or social behaviors can also be realized via non-verbal communication, but this is outside the scope of this paper.

Reasons for adding a capability for socially oriented conversation in a CPA include: i) increased overall engagement and receptivity (Cooper & Baynham, 2005) ii) improved recall of the learning material through emotional engagement (Hamann, 2001) in particular because social experiences activate the reward circuitry of the brain, which helps cement newly learned associations (Chen, Shohamy, Ross, Reeves & Wagner, 2009), iii) promotion of trust and rapport-building (Bickmore, 2003) and finally that students may feel more at ease with a learning task or topic (Kim, Wei, Xu, Ko & Llieva, 2007). For a more extensive presentation of these and other reasons, see Gulz, Haake, Silvervarg, Sjöden & Veletsianos (2011) and Silvervarg, Gulz & Sjöden (2010).

This paper describes a teachable agent that has been extended with a social conversation module. We discuss the justification for the extension and present an empirical study that evaluates the effects of the extension in terms of student experiences and learning. However, first we present selected examples of other CPAs capable of socially oriented conversation.

### **Previous work on socially oriented conversation in CPAs**

CPAs with a capability for socially oriented conversation broadly belong to two different categories. CPAs in the first category exhibit *on-task sociability*, that is they will and cannot digress into other topics than those that pertain to the learning task and domain. However, in connection with task-oriented conversation, they exhibit social behavior such as displaying encouragement, assurance, agreement, and praise. One example is the *cooperative co-learner* (Maldonado et al., 2005) that in addition to on-task conversation about the domain of English language idioms, compliments and shows concern and encouragement when the difficulty level of the questions increases or when the student fails on a question (e.g. "You'll get the next one"). As another example consider the *Low social* and *High Social* agents (Ai, Kumar, Nguyen, Nagasunder & Rosé, 2010) in a system for supporting collaborative design learning regarding thermodynamics. Student pairs can chat with each other as well as with the tutor CPA, where the percentage of social turns by the CPA (showing solidarity with a student who has difficulties, agreeing or showing tension release) is varied from 0% for the *No Social* agent, to 15% for the *Low social* and to 30% for the *High social*. As a third example Wang, Johnson, Mayuer, Rizzo, Shaw & Collins (2008) developed a model of socially intelligent tutorial dialogue on the basis of

politeness theory. The *polite tutor agent* provided tutorial feedback to promote learner face and mitigate face threat, whereas the standard tutor agent provided direct feedback that disregarded learner face.

The second category of socially oriented CPAs contains those that exhibit *off-domain sociability*<sup>1</sup>. They are able to go outside the task(s) and domain(s) and engage in conversation that involves small-talk-like topics, self-disclosure, personal narratives, etc. Although the work by T. Bickmore does not deal with pedagogical applications per se, it is central in this context. In (2003) he coined the term *relational agent*, an agent designed to develop and maintain long-term, socio-emotional relations with users, and he has conducted a large number of studies that compare relation-oriented with strictly task-oriented agents and explore various off-domain sociability features. The value of autobiographical stories in agents is investigated by Bickmore, Schulman & Yin (2009), with reference to Jakobson's (1960) *phatic* function of dialogue: to keep the communication channel open so that primary functional messages can be conveyed. The authors (Bickmore et al., *ibid*) propose that autobiographical storytelling by an agent is a central means for maintaining user engagement in an intervention over time – which can be crucial for educational applications. However, they also point at the importance that the stories that an agent tells are *truly* engaging.

Kumar et al. (2007) compared two software versions for letting student pairs engage in collaborative learning of mathematics via a chat. Both versions contained *cognitive support agents*, but one also contained *social dialogue agents*, designed to show personal interest in the students by asking them to reveal their personal preferences about extra-curricular domains. The preferences were used as input when the math problems were constructed, with the intention that the social dialogue should give students the impression that the agent takes personal interest in them. The addition of the social dialogue agents turned out to have a strong positive effect on the attitude that students displayed towards agents and a slight positive effect on learning outcomes.

Mehlman, Häring, Bühling, Wissner & André (2010) present research on a learning game for expressing conceptual knowledge through qualitative reasoning models. A set of CPAs are included, and among them a *quizmaster agent*, that besides asking questions and giving feedback makes small talk utterances and humorous distractions unrelated to the quiz domain. This is modelled on how quizmasters in famous television shows countervail participants' stress and provide a more enjoyable form of competition.

### **Extending a teachable agent based game extended with social off-task conversation**

The underlying game is a mathematics game that trains basic arithmetic skills with a focus on grounding base-ten concepts in spatial representations (Pareto, Haake, Lindström, Sjöden & Gulz, under revision; Pareto, Schwartz & Svensson, 2009). It employs a board-game design with a variety of game modes and levels of difficulty. When a student has learnt to play one particular board game, she may teach it to her *Teachable Agent* (TA) (Biswas, Katzlberger, Bransford & Schwartz, 2001) In the *observation mode* the TA “watches” the student play and picks up on game rules and on the student's responses to multiple-choice questions, such as

---

<sup>1</sup> Which can be combined with on-task sociability.

“Why did you choose this card?” The student then chooses one answer from the listed potential explanations (but only one correct answer), including a “Don’t know” option. Proper (or improper) choices of cards and answers promote corresponding skills in the TA throughout the game. In the *try-and-be-guided mode*, the agent is allowed to propose cards. The student either accepts the agent’s suggestion or rejects it and exchanges the agent’s card for another one. In the latter case the agent asks, via the multiple-choice-format, why the student thinks her card was a better choice. For its choices the agent uses its current knowledge, that is, if it has reached a threshold of “knowing a rule well enough”, then it may apply this rule. For more information on the AI in the system we refer to Pareto (2010) and to Pareto et al. (ibid.), which also describes the underlying pedagogical master-apprentice model, which differs from the more common teacher-student model in TA-systems.

In other words, the basic TA-system contains a (strictly regulated) form of on-task conversation, via a multiple-choice format. A simple form of on-task sociability is involved as well, for instance the agent may praise the student when she earns points in the game.<sup>2</sup> For the study presented in this paper, the game architecture was extended with a module where the student can engage in conversation with the TA, writing freely by means of the keyboard (in contrast to the multiple-choice format in the on-task conversation) and bring up basically any topic in a chat-like manner.

We refer to this chat-like conversation as *off-task* conversation and distinguish within it between *on-domain* conversation and *off-domain* conversation – the former referring to chat conversation related to school, math, the math game, etc., and the latter to any other topic. The off-task conversation is implemented as a mixed-initiative dialogue strategy, which allows both the agent and the user to direct the dialogue by introducing new topics and posing questions. The agent keeps a history of the topics in the dialogue, both from the current and previous sessions. *On-task* and *off-task* conversations have very different formats, but are still designed as two interrelated and complementary activities. The interconnecting factor is the persona of the agent, which integrates task and domain knowledge with off-domain knowledge (e.g. the agent is a 11-year old that goes to school and is learning math in the game, but also has interests such as music and film). For switching between playing the game and chatting we use the metaphor of regular breaks between lessons in school.

#### *Aims and relations to other systems*

The off-task conversation is in the first place a means to enrich the game and its motivational qualities for the novel age group of 12-14 year old users. Informal pre-studies revealed that these users required more variation than the younger students who became very engaged by the game in its basic form (Lindström, Haake, Sjödén & Gulz, 2011). Bickmore’s work and arguments on how social conversation with agents may be a means to maintain engagement in an intervention over time, was a main source of inspiration. Our aim is accordingly to enhance students’ experiences and to increase their inclination to continue using the game over time. A further aim is to exploit the off-task conversation for pedagogical

---

<sup>2</sup> But a TA cannot coach with respect to whether answers to questions are wrong or right, since it is a TA and not a teacher /instructor agent.

interventions such as influencing students math self-efficacy and attitudes toward math. It is worth to point out that our work, like the work by others related above, approach off-task conversation in terms of its *pedagogical power* – not in terms of being pedagogically detrimental in taking attention from the learning task (e.g. Rowe, McQuiggan, Robison & Lester, 2009). We return to this in the discussion.

Enhancing students' experience of the game can be achieved in various ways. For some individuals it is a question of adding variability in order to countervail boredom. For others is a question of making the learning domain of mathematics more appealing and making students less tense or nervous (compare Mehlman et al. (ibid.) above). This in turn relates to the potential for more dedicated pedagogical interventions, such as the work by Kim et al. (2007) on affecting students math self-efficacy and detracting “math anxious” students from perceived inability to confront mathematical learning content. For such interventions to work, trust in the agent is crucial. Bickmore (2003) has shown that small talk and conversational storytelling can contribute to building such trust.

For our system we have taken inspiration from all of the above mentioned research. Nevertheless, our system is unique in involving a *Teachable agent* capable of off-task, off-domain, sociability. A teachable agent provides specific challenges as well as specific advantages when it comes to developing its off-domain sociability. Among the advantages we find: this kind of pedagogical agent is integral to the learning software from the very start – not an ‘additional’ feature that may or may not be included in the software; a TA is in essence a social actor, someone that you shall teach and that shall learn from you; also with on-task interaction and conversation only, a social relation between students and TAs develops (Lindström et al., 2011; Chase, Chin, Oppezzo & Schwartz, 2009). In other words, in a TA system there is already a social relation to build on. But to develop an off-task conversation and a social actor related to this, in a consistent and believable manner, is a challenge. In the on-task-conversation the TA appears inferior to the student, oblivious from the start, learning everything from the student. In the off-task-conversation one may for various reasons strive for a more equivalent peer-relation. For details on how we approach this challenge of designing a peer, yet retaining some of the protégée-effect (Chase et al., 2009), see Gulz et al. (2011).

The present study focuses on the chat's potential to increase motivation and engagement by comparing groups of students who use the original game with those who use the extended version, i.e. with and without chat. Apart from students' experiences of the game, we study their perception of the role of the TA, their self-efficacy (i.e. beliefs about their competency in playing the math game), and their learning accomplishments. We also study possible differences between low- and high-achievers.

## **Method**

### *Participants and procedure*

38 female and 42 male 12-14 year olds from three classes in a Swedish school participated in the study. The students were assigned a value (low, middle or high) for math achievement by their teacher, where 18 were classified as low, 39 as middle and 23 as high. Each

class was divided into two groups with an even distribution of gender and math achievement in each group. All students got to play the game during three lessons. The NoC group used the game without the chat module. The WithC group used the game with the off-task module where a chat-break was offered after every two game sessions. During the first three of these breaks one had to chat with the agent until the break ended after three minutes and the chat was closed. For the coming breaks the students were offered a choice between chatting with the agent or continuing to play, and when chatting there was always a choice to end the chat before the break was over. The students in the NoC group spent on average a total of 105 minutes with the game and the students in the WithC group 120 minutes, in order to make the time spent on the math game sessions equal for both groups.

### *Instruments*

To evaluate the effect of the social chat on learning and experience of the game, a combination of data from questionnaires and computer-generated logs were used.

For an estimation of students' learning, we exploited system log data on the knowledge level of the TAs, which according to the Learning by Teaching paradigm reflects the students' own knowledge level. The students' learning outcome was calculated based on the agent's final knowledge level in relation to how many times the student had played and trained the agent.

After the last game session all students filled out a questionnaire with 18 statements scaled from 1 (Strongly disagree) to 7 (Strongly agree). The questionnaire included the areas: i) game experience, e.g. if interesting, challenging, easy to concentrate, ii) experience of the role of the TA in facilitating learning and increasing enjoyment in the game, iii) self-efficacy beliefs regarding the game play and one's role as teacher. Statements for i) and ii) were developed based on Anderson & Bourke (2000), the self-efficacy measurement according to guidelines from Bandura & Schunk (1981).

In order to get information on chat use and conversational behavior in the chats, we furthermore analyzed the chat logs for the WithCgroup – and in addition the logs for the on-task conversations for the same students. The analysis was conducted with a focus on the following five questions:

- i) How productive were students' off-task conversations in terms of driving a conversation further instead of being passive/non-informative in one's utterances or being counter-productive, i.e. producing nonsense?
- ii) How productive were the students' on-task conversations in terms of approaching many questions and getting the right answers for the TA instead of choosing the 'I don't know' answer or telling the TA the wrong answer?
- iii) Will the degree of productiveness in on-task conversation and in off-task conversation concur or not for students or not?

iv) To what extent do students control of the chat in terms of quitting the chat before the system ends it and refraining from starting a chat? Which students engage in such control and when?

v) What do students do when the chat log indicates that they have become very non-engaged in the conversation? Do they quit? Do they refrain from starting the next chat? Do they continue in the same non-engaged way?

### *The off-task and on-task productivity measurements*

On average, the number of phrases exchanged by the TA and an individual student were 130, of which half were uttered by the agent and half by the student. Each phrase was categorized in terms of how productive it was as a response to the previous dialogue or to start a dialogue. There were three categories. *Productive responses* were either an initiative to start a new topic to discuss (e.g., “Let’s talk about music”), a question (e.g., “What music do you like?”), or a reasonable and informative answer (e.g., “rock music”). *Non-informative responses* include phrases like “mhm”, “aha” and “okay” (when okay is not an informative answer to a posed question such as “What do you think about math?”). *Contra-productive responses* include obscenities, insults and nonsense letters. Two researchers encoded the dialogues separately based on agreed on principles, and the codes coincided in more than 98%.

A student off-task Productivity Measure (PM) was calculated as a function of the number of productive (P), non-informative (NI) and contra-productive (CP) phrases. Since both the productive and the contra-productive responses are stronger indications of willingness to engage in a discussion with the agent (positively or negatively) than the more passive non-informative responses, they are weighted twice as much. Hence, the productivity measure used is:

$$PM_{\text{off}} = 2 * P_{\text{off}} + NI_{\text{off}} - 2 * CP_{\text{off}}$$

In order to be comparable with the on-task productivity, the productivity measure is normalized to range within the scale [0,100].

All students’ responses to on-task questions during game play are logged by the system. On-task questions are initiated by the system and consist of deep-level explanatory questions based on the student’s just-performed action, framed as the inquisitive teachable agent asking the master (i.e., the student) to explain *why* he or she took such action. One correct and a few incorrect but reasonable explanations based on typical misconceptions are provided, as well as a “don’t know” option. To prevent that guessing among the choices should be an effective strategy, one wrong answer must be compensated by two correct ones, and the agent will comment on the master providing different answers. Question difficulty will progress only when questions are satisfactory answered, and we have seen in more long-term usage of the system that the percentage wrong answers decline over time, indicating that students learn not to guess. Hence, each multiple-choice question can either be answered correctly, incorrectly, or indecisive. Correct answers are considered *productive* in two aspects: the student has identified the proper explanation which indicates a proper understanding, and the agent can proceed with more challenging questions. Correspondingly, incorrect answers are considered *contra-productive*.

Indecisive answers can either result from the student not wanting to, or not being able to answer the question. Therefore, we consider these responses to be non-informative (of the student's actual ability). Therefore, the on-task dialogue productivity is measured in the same way as for the off-task dialogue, i.e.,

$$PM_{on} = 2 * P_{on} + NI_{on} - 2 * CP_{on}$$

Just as for the off-task dialogue, the productivity measure is normalized to range [0,100].

## Results and analysis

Since not all students could participate in all three lessons or fill in the questionnaire, the final analysis included 29 females and 32 males (and for the parts based on chat logs, 13 females and 17 males).

### *Game experience and learning outcomes*

A comparison of the results on the questionnaire and the knowledge level of the trained agent for the NoC- and WithC-groups is presented in Figure 1. Items were clustered and an average score calculated for the game experience, the perceived importance of the agent's role in the system, and self-efficacy beliefs. The figure shows that students in the WithC-group tended to have a more positive game experience (diff=0.54, p=0.07), but there was no difference in the perceived role of the TA in the game, and marginal differences in self-efficacy beliefs. Also students in the WithC-group tended to reach better result in terms of how well they taught their TA (diff=0,3, p=0.07).

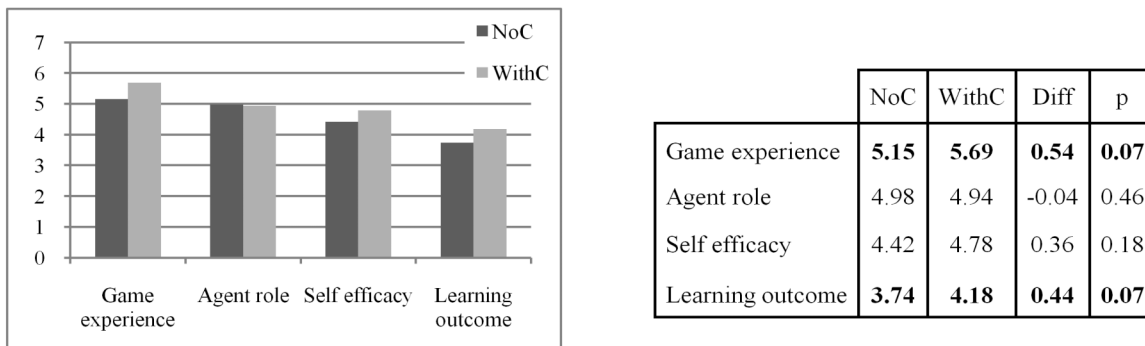


FIGURE 1. The diagram and table shows the difference between the NoC- and WithC groups regarding: game experience, the agent's role in the game, self efficacy, and the learning outcome in terms of how well they were able to train their agent.

In Table 1 we present subgroups in terms of students' achievements in mathematics, as given by teachers. We see no differences between the conditions for the low-achievers. However, for the medium and high achieving students the experience of the game is considerably more positive for the WithC condition (diff=0.71, p=0.04 and diff=0.91, p=0.09). High-achievers in

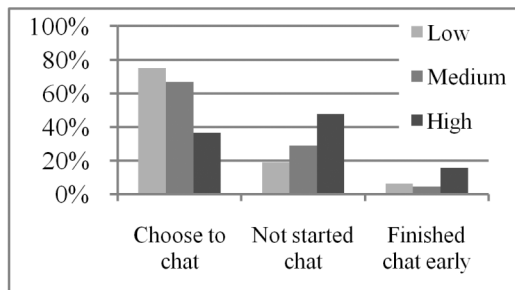
the WithC condition also rate their self-efficacy beliefs significantly higher (diff=0.93, p=0.04) and have a superior learning outcome (diff=7.65, p=0.06).

TABLE 1

The table shows the difference between the NoC- and WithC groups within subgroups divided based on achievements in mathematics, as indicated by teachers.

	Low achieving				Medium achieving				High achieving			
	NoC	WithC	Diff	p	NoC	WithC	Diff	p	NoC	WithC	Diff	p
Game experience	4.92	3.58	-1.35	0.12	<b>5.24</b>	<b>5.95</b>	<b>0.71</b>	<b>0.04</b>	<b>5.12</b>	<b>6.03</b>	<b>0.91</b>	<b>0.09</b>
Agent role	4.58	3.31	-1.27	0.21	4.89	4.77	-0.12	0.41	5.34	5.71	0.37	0.3
Self efficacy	4.05	2.5	-1.55	0.11	4.48	4.85	0.37	0.25	<b>4.52</b>	<b>5.44</b>	<b>0.93</b>	<b>0.04</b>
Learning outcome	3.06	3.06	0.003	0.5	3.87	4.09	0.22	0.29	<b>3.91</b>	<b>4.68</b>	<b>0.77</b>	<b>0.06</b>

For the WithC group we analyzed the chat behavior for the different subgroups. As shown in Figure 2 there is a clear pattern where high achievers choose not to chat –by refraining from starting a chat or by quitting a chat before chat time is out – to a much greater extent than low and medium achievers. Medium achievers tended in turn to choose not to chat to a higher extent than low achievers.



	Total	Low	Medium	High
Choose to chat	56%	75%	67%	36%
Not started chat	35%	19%	29%	48%
Finished chat early	9%	6%	4%	16%

FIGURE 2. The diagram and table shows the difference in how low-, medium- and high achievers choose to chat when given the choice.

*Tendency to control the chat – quit early, refrain from next chat – in relation to unengaged conversation/chatting*

The already presented results, i) that high-achievers (and mid-achievers) tended to like the game better when the chat was included but low-achievers not, but ii) that high-achievers (and mid-achievers) were more inclined to quit and to refrain from starting the chat, may at first seem intriguing. A closer analysis of the chat behavior in those students groups provides us, however, with a possible line of explanation of the apparent contradiction.

This analysis started by an identification of those chat passages where a student clearly seemed to have lost her engagement in the conversation with the agent. One criterion is when a student repeats a blank, a dot, one single word, or meaningless strings, and continues to do so without getting back to a productive conversation. Another criterion is when a student goes on

with something that seems more like a monologue, sometimes including harassment, which does not relate to any of the utterances by the TA but runs as a monologue **on** its own. A third criterion is when the conversation gets very awkward or stuttering, sometimes (partly) because the TA is extremely repetitive and/or unable to utter anything of relevance and the conversational engagement in the student can be seen to decline in her utterances.

Among the students included in this analysis were 11 high-achievers, 6 low-achievers and 13 mid-achievers. Starting with the number of chats where instances of clearly lost engagement was found, 22 of these were high-achiever chats, 11 were low-achiever chats and 10 were mid-achiever chats. But of more interest are the different behavioral patterns for those chats with respect to student group (even though we cannot claim statistical significance given the limited number of students involved in the analysis).

For the 22 “extreme low-engagement” instances with high-achievers, there were 5 instances of quitting the chat and 13 of refraining from starting the next chat, but only 4 instances of getting on and also starting next chat. Reversely for the 11 instances with low-achievers, there were 8 instances of getting on and also starting next chat and only 3 instances of refraining from starting the next chat and no case of quitting a chat. For the 10 instances with mid-achievers, finally, there was an even distribution with 5 instances of getting on and also starting next chat and 5 instances of controlling the situation: 3 instances of quitting and 2 instances of refraining from starting next chat.

	Low-achievers	Medium	High-achievers
Quits the chat	0	3	5
Does not start next chat	3	2	13
Gets on with unengaged chat	8	5	4

TABLE 2

Actions taken in situations where the chat log indicates extreme low engagement in the conversation on the part of the student.

In other words, low-achievers **were** more inclined to continue a chat even when there is a strong indication that they **were** unengaged in what they **were** doing, whereas mid-achievers and most of all high-achievers, **were** more inclined in these situations to take control or action: they quit the chat and/or refrained **ed** from starting next chat.

Two aspects of this may contribute to why high-achievers (and mid-achievers) tended to like the game better when the chat was included but low-achievers did not. First, it is well known that the experience of having control over one’s situation plays an important role for a positive learning experience. Second, according to the behavioral differences described, low-achievers will tend to spend more time than mid- and high-achievers with a chat that they experience as non-engaging (boring, meaningless).

(Also, if an insufficient initiative in taking control – or an insufficient sense of control – when chatting exists and is a drawback for low-achievers, designers may need to think about possible measures to take.)



## Discussion

The primary result of the study is the indication that an added off-task conversation module can i) improve students' game experience and ii) is not necessarily a disadvantage in terms of learning accomplishment, but can to the contrary improve learning. This adds further support to our and others' approaches to the introduction of socially oriented off-task conversation as an integral learning element – in contrast to approaches where off-task behavior is regarded to divert attention from learning thereby reducing the pedagogical efficiency (e.g. Rowe et al., 2009). We hold both kinds of approaches valid, but advocate more nuances in the term “off-task behavior/conversation” in pedagogical contexts, and specifically for digital learning environments. The unit of learning is, we hold, a crucial parameter. For software meant to be used during a set, limited time and in relation to clear learning objectives, it may indeed be relevant to find means to control and even minimize off-task – and also relatively easy to determine whether a behavior indeed is unrelated to the curriculum in question. But in relation to a longer term learning context, another kind of balancing must be considered. Off-task behavior might be essential for the development of a relation between agent and student, which can be central for reaching certain learning goals in the long run. The TA game discussed in this paper offers a long term learning environment, and it is in view of this that we regard the off-task conversation or chat module promising.

A strong reason to believe in the pedagogical potential of a further developed chat-module that integrates off-task and on-task conversation – in this system and in similar systems – is the following: More than a third of the students spontaneously initiated chat conversations about the math game, wanting to discuss whether the TA found it difficult, whether it **thought** it **had** learnt much, whether it thought it went well, etc. With an agent that can carry on such conversation in a more productive way than in our present system, one may influence students who are engaged in the off-task conversation to become more engaged and make more effort in the game play.

## Acknowledgments

This material is based upon work supported by grants from the Swedish Knowledge Foundation, Wallenberg Global Learning Foundation, and Linköping University.

## REFERENCE LIST

Ai, H., Kumar, R., Nguyen, D., Nagasunder, A., & Rosé, C. (2010). Exploring the effectiveness of social capabilities and goal alignment in computer supported collaborative learning. In V. Aleven, J. Kay, & J. Mostow (Eds.), *LNCS 6095. Proc. of the 10th Int. Conf. on Intelligent Tutoring Systems, Part II* (pp. 134-143). Berlin/Heidelberg, Germany: Springer-Verlag.

Anderson, L. W., & Bourke, S. F. (2000). *Assessing Affective Characteristics in the Schools* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology*, 41, 586-598.

Bickmore, T. (2003). *Relational Agents: Effecting Change through Human-Computer Relationships*. Doctoral dissertation, Massachusetts Institute of Technology, Media Arts & Sciences.

Bickmore, T., Schulman, D., & Yin, L. (2009). Engagement vs. deceit: Virtual humans with human autobiographies. In Z. Ruttkay, M. Kipp, A. Nijholt, & H. H. Willhjálmsson (Eds.), *LNCS 5773: Proc. of the 9th Int. Conf. on Intelligent Virtual Agents* (pp. 6-19). Berlin/Heidelberg, Germany: Springer-Verlag.

Biswas, G., Katzlberger, T., Bransford, J., & Schwartz, D. (2001). Extending intelligent learning environments with teachable agents to enhance learning. In J. D. Moore, C. L. Redfield, & W. L. Johnson (Eds.), *Proc. of the 10th Int. Conf. on Artificial Intelligence in Education: AI-ED in the Wired and Wireless Future* (pp. 389-397). Amsterdam, The Netherlands: IOS Press.

Chase, C. C., Chin, D. B., Opezzo, M. A., & Schwartz, D. L. (2009). Teachable agents and the protégé effect: Increasing the effort towards learning. *Journal of Science Education and Technology*, 18, 334-352.

Chen, J., Shohamy, J. D., Ross, V., Reeves, B., & Wagner, A. D. (2009). *The Impact of Social Belief on the Neurophysiology of Learning and Memory*. Abstract for the Annual Meeting of the Society for Neuroscience. San Francisco, CA.

Cooper, B., & Baynham, M. (2005). *Rites of Passage: Embedding Meaningful Language, Literacy and Numeracy Skills in Skilled Trades Courses Through Significant and Transforming Relationships*. London: National Research and Development Centre for Adult Literacy and Numeracy.

Graesser, A., Chipman, P., Haynes, B., & Olney, A. (2005). AutoTutor: An intelligent tutoring system with mixed-initiative dialog. *IEEE Transactions in Education*, 48, 612-618.

Gulz, A., Haake, M., Silvervarg, A., Sjöden, B., & Veletsianos, G. (2011). Building a social conversational pedagogical agent – design challenges and methodological approaches. In D. Perez-Marin, & I. Pascual-Nieto (Eds.), *Conversational Agents and Natural Language Interaction: Techniques and Effective Practices*. Hershey, PA: IGI Global.

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences*, 5, 394-400.

Jakobson, R. (1960). Linguistics and poetics. In: T. A. Sebeok (Ed.), *Style in Language*. (pp. 130-144). Cambridge, MA: MIT Press.

Kim, Y., Wei, Q., Xu, B., Ko, Y., & Ilieva, V. (2007). MathGirls: Increasing girls' positive attitudes and self-efficacy through pedagogical agents. In R. Luckin, K. R. Koedinger, & J. Greer (Eds.), *Proc. of the 13th Int. Conf. on Artificial Intelligence in Education: Building Technology Rich Learning Contexts That Work* (pp. 119-126). Amsterdam, The Netherlands: IOS Press.

Kumar, R., Gweon, G., Joshi, M., Cui, Y., & Rose, C. P. (2007). Supporting students working together on math with social dialogue. *Proc. of the SLATE Workshop on Speech and Language Technology in Education* (pp. 96-99). Farmington, PA.

Lindström, P., Haake, M., Sjöden, B., & Gulz, A. (2011). Matching and mismatching between the pedagogical design principles of a math game and the actual practices of play. *Journal of Computer Assisted Learning*, 27, 90-102.

Maldonado, H., Lee, J., Brave, S., Nass, C., Nakajima, H., Yamada, R., Iwamura, K., & Morishima, Y. (2005). We learn better together: Enhancing eLearning with emotional characters. In T. Koschmann, D. Suther, & T. W. Chan (Eds.), *Proc. of the 7th Computer Supportive Collaborative Learning Conference: The Next Ten Years* (pp. 408-417). Mahwah, NJ: Lawrence Erlbaum.

Mehlmann, G., Häring, M., Bühling, R., Wißner, M., & André, E. (2010). Multiple agent roles in an adaptive virtual classroom environment. In J. Allbeck, N. Badler, T. Bickmore, C. Pelachaud, & A. Safonova (Eds.), *LNCS 6356: Proc. of the 10th Int. Conf. on Intelligent Virtual Agents* (pp. 250-256). Berlin/Heidelberg, Germany: Springer-Verlag.

Pareto, L. (2010). A teachable agent game for elementary school mathematics promoting causal reasoning and choice. In D. Pérez-Marín, I. Pascual-Nieto, & S. Bull (Eds.), *Proc. of the Int. Workshop on Adaptation and Personalization in E-B/Learning using Pedagogic Conversational Agents* (pp. 13-19). Hawaii, USA.

Pareto, L., Haake, M., Lindström, P., Sjöden, B., & Gulz, A. *A Teachable Agent Based Game Affording Collaboration and Competition – Evaluating Math Comprehension and Motivation*. (under revision).

Pareto, L., Schwartz, D., & Svensson, L. (2009). Learning by guiding a teachable agent to play an educational game. In V. Dimitrova, R. Mizoguchi, B. du Boulay, & A. Graesser (Eds.), *Proc. of the 14th Int. Conf. on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling* (pp. 662-664). Amsterdam, The Netherlands: IOS Press.

Rowe, J., McQuiggan, S., Robison, J., & Lester, J. (2009). Off-task behavior in narrative-centered learning environments. In V. Dimitrova, R. Mizoguchi, B. du Boulay, & A. Graesser (Eds.), *Proc. of the 14th Int. Conf. on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling* (pp. 99-106). Amsterdam, The Netherlands: IOS Press.

Silvervarg, A., Gulz, A., & Sjöden, B. (2010). Design for off-task interaction – rethinking pedagogy in technology enhanced learning. *Proc. of the 10th IEEE Int. Conf. on Advanced Learning Technologies* (pp.204-206). Sousse, Tunisia.

Wang, N., Johnson, W. L., Mayer, R. E., Rizzo, P., Shaw, E., & Collins, H. (2008). The politeness effect: Pedagogical agents and learning outcomes. *International Journal of Human Computer Studies*, 66, 96-112.