

Animated Pedagogical Agents: Do they advance student motivation and learning in an inquiry learning environment?

Hans van der Meij^a, Jan van der Meij^b, and Ruth Harmsen^a

^aUniversity of Twente, Faculty of Behavioral Science, Department of Instructional Technology, P.O. Box 217, 7500 AE Enschede, The Netherlands

^bUniversity of Twente, Faculty of Behavioral Science, ELAN Institute for Teacher Education, Science Communication & School Practices, P.O. Box 217, 7500 AE Enschede, The Netherlands

h.vandermeij@utwente.nl; j.vandermeij@utwente.nl; ruth_harmsen@hotmail.com

Abstract

Student behavior in inquiry learning environments has often been found to be in need of (meta)cognitive support. Two pilots revealed that students might also benefit from motivational support in such an environment. An experiment with 61 junior high school students (ages 14-16) compared three conditions related to motivational support: a motivating agent (female image and voice), the agent's voice only, or no support. The support provided addressed two vital components of motivation: task-relevance and self-efficacy belief. The learning environment covered a topic in physics, a domain for which a gender difference in self-efficacy has frequently been reported. The effects of both gender and condition were investigated. Overall, students showed gains in self-efficacy belief, perceptions of task-relevance, and learning. Effects related to gender and condition included the finding that: (1) when the task was more difficult, the self-efficacy belief of the girls tended to increase for the Agent and Voice condition while staying equal in the Control condition, whereas that of the boys increased in the Control condition but decreased for the Agent and Voice condition, and (2) girls tended to learn more in the Agent and Voice condition while boys did better in the Control condition. The discussion addresses the question of how to create an agent that fulfills basic requirements of credibility (external properties) and task-specific support (internal properties).

Keywords: Animated Pedagogical Agents, Motivation, Inquiry learning, Simulation-based learning environments

Introduction

Inquiry learning approaches invite students to engage actively in a variety of activities that have been found to relate to knowledge development, such as orienting, formulating hypotheses, experimenting, and drawing conclusions ([de Jong, 2006b](#)). Reviews invariably reveal that students need support when engaging in such inquiry activities; pure inquiry learning is less conducive to learning than direct instruction or guided-inquiry learning (e.g., [Alfieri, Brooks, Aldrich, & Tenenbaum, 2011](#); [Eysink, de Jong, Berthold, Kolloffel, Opfermann, & Wouters, 2009](#); [Kirschner, Sweller, & Clark, 2006](#); [Scalise, Timms, Moorjani, Clark, Holtermann, & Irvin, 2011](#)). Therefore, considerable effort is spent on creating scaffolds that support the inquiry processes of students in inquiry learning environments. A large majority of these support tools are directed toward the development of (meta)cognitive processes ([de Jong, 2006a](#)).

The extent to which students actively engage in inquiry processes and gain knowledge from these activities depends not only on their (meta)cognition, but also on their motivation. Students must be willing to explore the learning environment, and they must persist and exert effort when faced with difficult tasks ([Wang, Johnson, Mayer, Rizzo, Shaw, & Collins, 2008](#)). These activities are at risk when students fail to perceive the relevance of engaging in pertinent tasks and activities and when they have little confidence in their ability to handle problems they encounter in the learning environment. In other words, motivation, “the process whereby goal-directed activity is instigated and sustained” ([Pintrich & Schunk, 2002, p. 5](#)), is also a precondition of successful inquiry learning.

Research on student behavior in inquiry learning environments occasionally reports findings that might signal motivational deficiencies. For instance, [Hagemans, van der Meij, and de Jong \(submitted\)](#) found that students rarely went back to restudy assignments they had initially completed incorrectly. Because the learning environment clearly flagged answers as being (in)correct, the relative dearth of restudying behavior is likely to be due to lack of motivation. Many studies have reported lack of exploration by students of a sufficient set of factors or variables and random or haphazard rather than systematic and complete testing of variables and conditions, which may be overcome by cognitive support (e.g., [de Jong & van Joolingen, 1998](#); [Mayer, 2004](#)), but which may also be signs of lack of motivation.

These findings do not constitute proof of motivational problems, of course. At best, they provide circumstantial evidence that motivation may be at stake. A study on motivational support in inquiry learning environments should therefore first ascertain whether such support might be needed and helpful. This was done in two pilots with students from the target audience. The first pilot showed that student motivation for task engagement in an inquiry learning environment on a science topic was indeed relatively low. In the second pilot, significant motivational gains were observed for students working in the same environment who received motivational comments from an experimenter. These pilots provided justification for conducting an experiment on the role of an Animated Pedagogical Agent (APA) in improving student motivation and learning in this environment. Because research repeatedly reports the presence of gender differences in high school students' motivation for science (e.g., [Catsambis, 1995](#); [Lau & Roeser, 2002](#); [Mattern & Schau, 2002](#); [Osborne, Simon, & Collins, 2003](#); [Yeung, Kuppan, Kadir, & Foong, 2011](#)), we paid particular attention to this issue.

Animated Pedagogical Agents (APAs)

Animated Pedagogical Agents (APAs) are broadly described as lifelike characters that guide users through multimedia learning environments. In educational settings this definition is usually refined by adding that the APA is intended to enhance learning (e.g., [Craig, Gholson, & Driscoll, 2002](#); [Moreno, Reislein, & Ozogul, 2010](#); [Shaw, Johnson, & Ganeshan,](#)

1999). There are APAs in many different guises, including humans (e.g., [AutoTutor - Graesser & McNamara, 2010](#)), animals (e.g., [Herman the Bug - Moreno, Mayer, Spires, & Lester, 2001b](#)), and inanimate objects (e.g., [Microsoft's Clippy - Haake, 2009](#)).

An important stimulus for the inclusion of APAs in multimedia learning environments is that they can humanize the user experience. In fact, some of the first APAs were introduced into electronic environments specifically for the purpose of making the users' interactions with the system more life-like (e.g., [André, Mailer, & Rist, 1996](#); [Bates, 1994](#); [Cassell, 2000](#); [Lester, Converse, Kahler, Barlow, Stone, & Bhogal, 1997](#); [Paiva & Machado, 1998](#); [Picard & Klein, 2002](#)). Considerable design efforts went into creating a likeable, intelligent, credible, and trustworthy agent that improved user interactions with the system. Users' experiences with the system would then be tested and their views on the agent's qualities elicited. When agents began to be used in education, attention shifted toward the design of agents that could affect student motivation and learning.

Variouly labeled as social cue, social agency, or persona effect (henceforth simply the persona effect), the claim has been examined that the sheer presence of an APA primes a social interaction schema that positively influences student motivation (e.g., [Atkinson, Mayer, & Merrill, 2005](#); [Choi & Clark, 2006](#); [Lester, et al., 1997](#); [Mayer, 2005](#); [Moreno, Mayer, Spires, & Lester, 2001a](#); [Moundridou & Virvou, 2002](#); [Reeves & Nass, 1996](#)). The findings from these empirical studies have been equivocal. In part, this may have stemmed from the variety of measures of motivation that were used. Motivational effects of the APA were assessed with regard to utility, entertainment, enjoyment of working with the learning environment or the learning task, perceptions of task difficulty, self-efficacy belief, domain or task interest, willingness to continue, and satisfaction.

The mixed outcomes of these studies could also be due to the underspecified nature of the APA, with regard to both its attributes and the content that it provides to the learner. That is, it has become increasingly clear that there is often more to the persona effect than simply an embodied interface (the agent's image and voice) that enlivens the interactions with the computer. In addition to the discovery that variations in the visual and auditory presence of an APA affect the outcomes, it has also been both argued and observed that APAs often provide important content that affect motivation and learning (e.g., [Clark & Choi, 2005](#); [Dehn & Van Mulken, 2000](#)).

The present study addresses several of these issues. It moves away from measuring a lot of different aspects of motivation without having them unified by a motivational theory, by concentrating on the two key tenets from expectancy-value theory, namely perceptions of task-relevance and self-efficacy belief ([Eccles & Wigfield, 2002](#)). It also moves beyond an examination of the influence of the sheer presence of an APA by investigating the effects of an APA that is intended specifically to be motivating (mAPA). Only a few empirical studies examining effects of mAPAs on motivation and learning have compared the agent conditions to a no agent control condition (see [Heidig & Clarebout, 2011](#)). The following section summarizes these studies.

Motivating Animated Pedagogical Agents (mAPAs)

[Plant, Baylor, Doerr, and Rosenberg-Kima \(2009\)](#) examined gender influences of a male and a female mAPA serving as a role model. The mAPA promoted engineering by stressing the people-oriented nature and social benefits of this profession, and countering stereotypes about women and engineering. Students in the control condition did not interact with an agent. Changes in junior high school students' attitudes and interest regarding engineering-related fields were measured, along with their scores on a math performance test. A significant main effect for self-efficacy belief was reported, among other findings. The self-efficacy of both girls and boys increased significantly after interacting with the agent,

regardless of agent gender. However, the female agent produced significantly higher outcomes on the math posttest than the male agent or no agent.

[Domagk \(2010\)](#) conducted two studies in which an agent instructed students about perceptions of depth and Gestalt laws, manipulating the agent's social appeal. In the first, the mAPA was either likable, neutral, or dislikable in appearance (image). A control condition did not include an agent. Motivational measures administered during training yielded one significant outcome. The likable agent led to higher state motivation (a mixture of task enjoyment and self-efficacy belief) than the dislikable agent. No effect of the mAPA on retention was found, but there was an interaction between mAPA type and performance on transfer items from the posttest, where the students who had experienced the likable mAPA outperformed all other groups. The second study systematically varied both image (likable versus dislikable) and voice (likable versus dislikable). In contrast to Study 1 no effects of mAPA type on motivation during training were found. One significant effect for learning was discovered, namely that the agent with a dislikable image and voice negatively affected transfer when compared to all other groups. In other words, this type of agent was found to hurt performance.

Arroyo and colleagues examined the effects of mAPAs on student motivation in several studies, paying particular attention to the influence of gender ([Arroyo, Woolf, Cooper, Bursleson, & Muldner, 2011](#); [Arroyo, Woolf, Royer, & Tai, 2009](#)). In these studies the mAPA played the role of learning companion in an intelligent tutoring system (ITS) on math for high school students. The agent's motivational messages were derived from Dweck's ([1999, 2007](#)) research on human motivation, which draws attention to the role of attributions and effort. Specifically, the agent challenged the students' conception of math ability as stable, emphasizing more the influence of effort and perseverance in learning math. Students were randomly assigned to a female mAPA, a male mAPA, or a no-agent control condition. The first study ([Arroyo, et al., 2009](#)) found no significant effects. The second study ([Arroyo, et al., 2011](#)) reported a significant effect of the female mAPA on math attitudes. Boys as well as girls indicated a greater liking for math and also reported having developed more confidence in their math skills. Specific measures of motivation taken during and after training further revealed a gender-specific effect of the mAPA on self-efficacy belief. For girls only, there was a significant increase in self-confidence in solving math problems, with both the male and the female mAPA. No effect of the mAPA on learning was found.

The study reported in this paper presented a female mAPA as a learning companion in an inquiry learning environment for science on the topic of Motion. The choice of the agent's role was in line with the aforementioned findings from [Plant, et al. \(2009\)](#) and [Arroyo, et al. \(2011\)](#). For females, who maybe need more encouragement in science, the gender of the agent accorded with Bandura's ([1997](#)) claim that model-target similarity increases the effectiveness of the model. The mAPA focused on perceptions of task-relevance and self-efficacy belief. Design strategies for stimulating these motivational aspects have been put forward in Keller's model addressing attention, relevance, confidence, and satisfaction (ARCS-model; [1987, 2010](#)), and when implemented have repeatedly been found to lead to significant motivational gains for students in regular educational settings.

Research Questions

The aim of this study was to enhance motivation and learning in an inquiry learning environment. The fact that the environment dealt with science was important for considering gender effects. Three experimental conditions were compared. In the *Agent* condition, students received the support of a female mAPA who was presented as a peer student. In the *Voice* condition, the learners received the same motivational comments audibly but without the image of the agent. The *Voice* condition is a check on the criticality of the agent's image,

and an exploration of an alternative, low-cost method of support ([see Clark & Choi, 2005](#)). In the *Control* condition no motivating support was given.

The study addresses three research questions:

(1) Do conditions and gender affect changes in *motivation during training*?

Perceptions of task-relevance and self-efficacy belief were measured at three points during participants' working with the learning environment. The first measurement, at the start, could establish whether there was an initial gender difference in the target audience. Our prediction was that there would be a greater increase in motivation while interacting with the learning environment for the Agent and Voice conditions compared to the Control condition, and further that the Agent condition would increase more than the Voice condition. It was further expected that girls would improve more than boys in the Agent and Voice conditions, because the female agent/voice should be a stronger model for girls.

(2) Do conditions and gender affect *motivation after training*? Our prediction was that motivation for future tasks after training would be higher for the Agent and Voice conditions compared to the Control condition, and further that the Agent condition would score higher than the Voice condition. In the event that a gender difference existed at the start, a further question was whether this would level off in the Agent and Voice conditions.

(3) Do conditions and gender affect *learning gains*? Our prediction was that the students in the Agent and Voice conditions would have higher learning gains than those in the Control condition, and further that the Agent condition would show greater gains than the Voice condition.

Method

Participants

Participants were 61 third-year secondary school students (31 girls and 30 boys, between 14 and 16 years old). Participants were randomly assigned to conditions. Stratification was applied in order to have similar distribution for gender within the three conditions.

Materials

Learning environment

The participants worked with a simulation-based inquiry learning environment on the physics topic of Motion, built in the SimQuest authoring environment ([de Jong, van Joolingen, Veermans, & van der Meij, 2005](#); [van Joolingen & de Jong, 2003](#)). The topic was a compulsory part of the curriculum for the participants. Participants already had some prior knowledge of kinematics, but not on the topic of uniformly accelerated motion that was discussed in the learning environment. In the learning environment the topic is divided into three subtopics (displacement and time, speed and velocity, and acceleration), following the principle of model progression ([White & Frederiksen, 1990](#)). Thereby, the difficulty increases with each subtopic. The subtopics are each being covered in nine to ten assignments for a total of twenty-nine assignments with corresponding feedback. When an assignment is opened, a simulation interface also opens, showing one or more graphs, and an animation of one or two cars (see Figure 1). Each assignment poses a multiple-choice question that encourages students to explore or experiment with the simulation. The learning environment automatically gives students immediate textual feedback on the correctness of their answer.

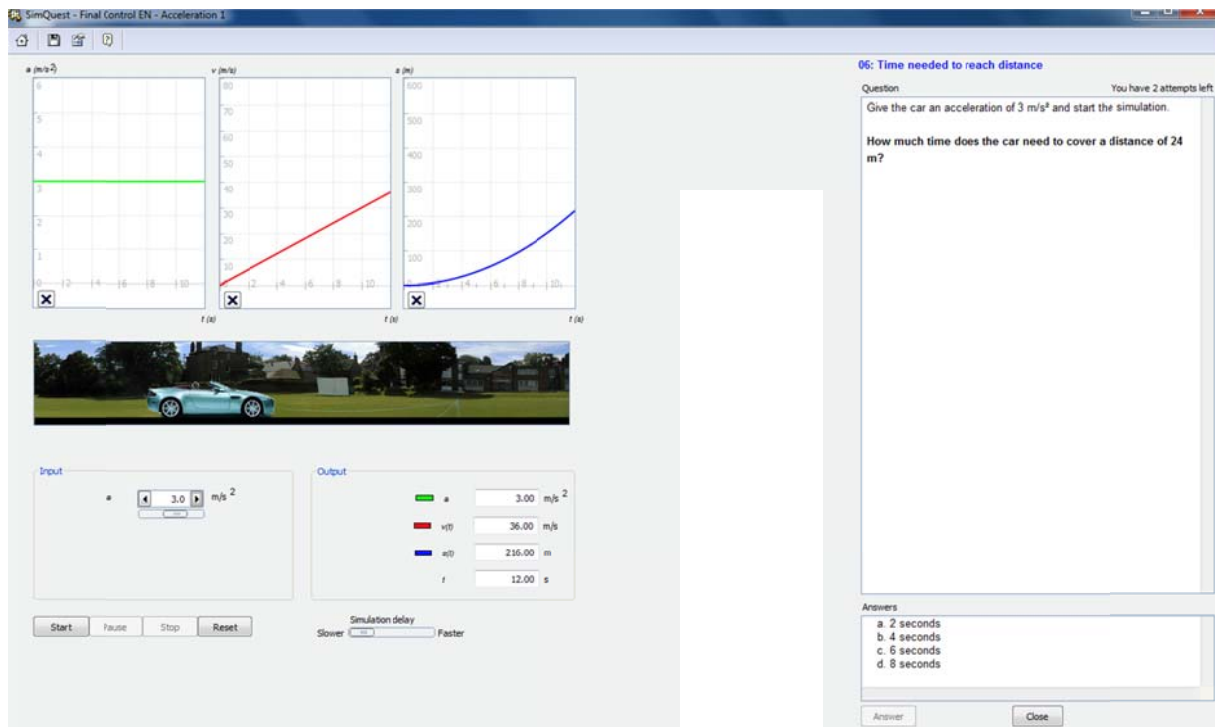


Figure 1. Interface for the Motion environment; Control and Voice conditions.

The Agent condition was enhanced with a female mAPA called Emma (see Figure 2), who was created by combining SimQuest, Elckerlyc ([van Welbergen, Reidsma, Ruttkay, & Zwiers, 2010](#)), and Loquendo (<http://www.loquendo.com>). Elckerlyc is a BML (behavioral markup language) compliant behavior realizer for generating multimodal verbal and nonverbal behavior for Virtual Humans (VHs). Animation in Elckerlyc is generated using precise temporal and spatial control offered by procedural motion and the naturalness of physical simulation. From Elckerlyc we only displayed the portrait of the VH. Emma's facial expressions (e.g., neutral, happy, or sad) and head movements were attuned to her comments. Emma's expression as well as here eye blinks and movements were neutral between comments. Emma's voice was created with the standard Dutch voice of the Loquendo program. Unfortunately, Loquendo has only a limited set of pre-programmed words and sentences expressed with the correct tonality and emotion. Non-programmed (self-constructed) words and sentences are pronounced more monotonously. Emma was always present in the assignment interface. She automatically gave motivational comments two seconds after the student had opened an assignment and one second after an assignment had been answered. In the Voice condition, students received motivational comments in the same way, uttered in Emma's voice but without the corresponding image of Emma. The Control condition received no motivational support.

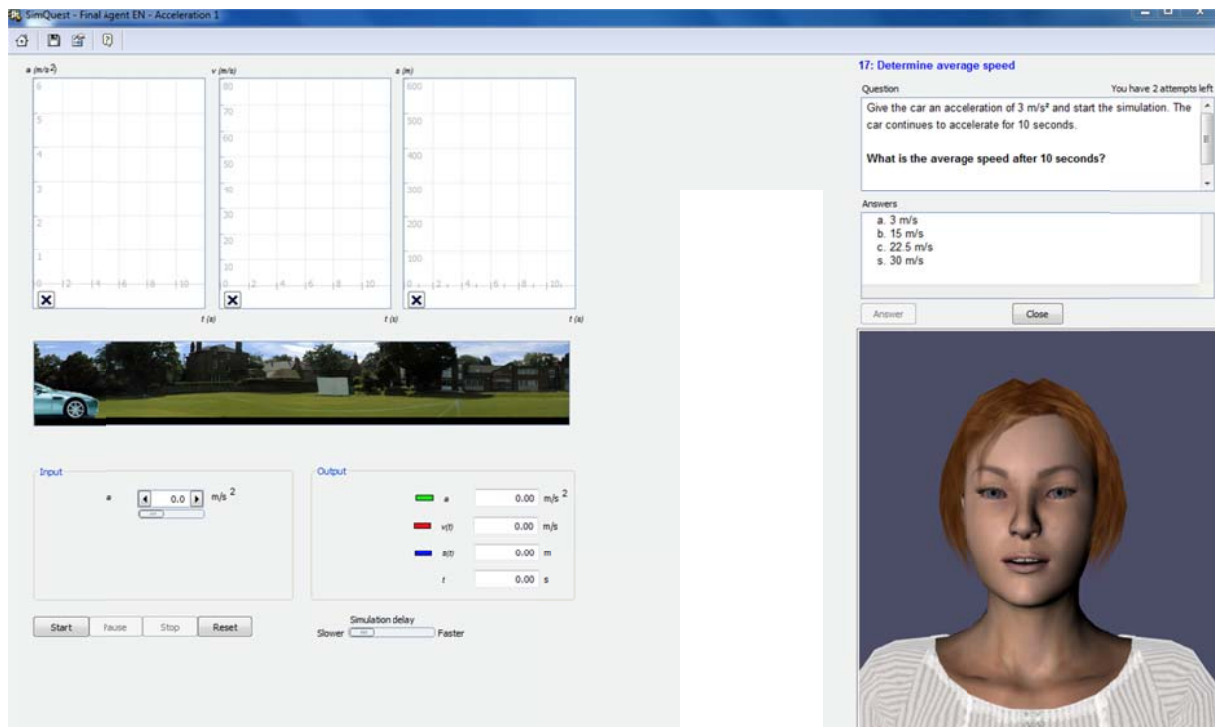


Figure 2. Interface for the Motion environment; Agent condition.

The motivational comments for the mAPA were based on Keller's ARCS-model (1987, 2010) from which the confidence component maps onto self-efficacy belief. Figures 3a and 3b summarize the main instructional strategies for stimulating perceptions of task-relevance and self-efficacy belief.

Relevance strategy	Description	Examples (Agent voice)
R1: Goal orientation	<i>Illustrate task-relevance</i> Let students see that what they are learning will help them realize a certain goal in the future by providing statements or examples of the utility of the instruction. Either present goals or have learners define them.	"Great. Now I know how to compute velocity. This will help me in the other assignments on motion."
R2: Motive matching	<i>Address student needs and desires</i> Make instruction responsive to learner motives and values by providing personal achievement opportunities, positive role models, leadership responsibilities, and cooperative activities.	"Why do I need to learn this? Let's just go ahead and try, and hope that it will become clear later on."
R3: Familiarity	<i>Link to student experiences</i> Make materials and concepts familiar by providing concrete examples and analogies related to the learner's work or background.	"This is just like biking. To speed up you just have to push the pedals harder."

Figure 3. Relevance strategies used in constructing the agent's comments.

Self-efficacy belief strategy	Description	Examples (Agent voice)
C1: Learning requirements	<i>Express belief in students' capacities</i> Establish trust and positive expectations by explaining the requirements for success and the evaluative criteria.	"Let's see. This is another assignment about time. The last one went well. So I should be able to solve this one, too."
C2: Success opportunities	<i>Ensure success and gradually increase complexity</i> Increase belief in competence by providing many, varied, and challenging experiences that increase learning success.	"This is like the previous one. When I just read the assignment carefully, I will do fine."
C3: Personal control	<i>Indicate that student effort matters</i> Provide feedback that ascribes success to personal effort, and use techniques that offer personal control (if possible).	"Yeah. Maybe I can do these tasks after all."

Figure 4. Self-efficacy belief strategies used in constructing the agent's comments.

Questionnaires and tests

All questionnaire responses were given on a Likert scale. Depending on the question, the response anchors were given as *not relevant – extremely relevant*, *extremely poorly – extremely well*, and so on, as appropriate. Participants answered each question on paper by marking a cross at the right place on an unmarked line of 10 centimeters long. The participant's score was measured in centimeters with one decimal place (maximum 10).

To assess *motivation during training* students were asked one question (R) on perceived task-relevance (i.e., "How relevant are these assignments in your opinion?") and one (SE) on self-efficacy belief (i.e., "How well do you think you will do on these assignments") after they had acquainted themselves with the first assignment of a new subtopic. Students answered the two motivation questions after assignments 1 (R1 and SE1), 10 (R2 and SE2), and 21 (R3 and SE3). The interface instructed students on these moments that they were expected to request these motivation questions from the experimenter.

To assess *motivation after training* students were asked six questions, identical to those asked during training, on perceived task-relevance (Ra1, Ra2, and Ra3) and self-efficacy belief (SEa1, SEa2, and SEa3) for related hypothetical assignments on the three subtopics. As they were not actually going to work on these assignments, the participants had to imagine doing so.

A post-training *Agent* questionnaire, administered only in the Agent condition, asked eleven questions on the credibility of the mAPA (e.g., "Emma said what I also thought" and "I felt just like Emma did."). After removing two questions, reliability was good with a Cronbach α of 0.89.

The learning environment logged all user actions. From these logs we analyzed the number of completed assignments and errors.

Learning was assessed with a *pre-test and post-test*. Each test consisted of 27 multiple-choice items with four answer alternatives. The tests measured three subtypes of knowledge (see Jonassen et al., 1993): conceptual knowledge (e.g., "What does the slope of the graph in the velocity-time diagram represent?"), structural knowledge (e.g., "Which set of concepts is representative of only uniformly accelerated motion?"), and procedural knowledge (e.g., "How do you determine the displacement of a car based on the information that is displayed

in the velocity-time graph?”). The items in the pre-test and post-test measured the same underlying constructs, but were phrased differently.

Procedure

About one week prior to the experiment students took the pre-test, for which a maximum time of 20 minutes was given. Training took place in a separate room where students could work individually with the learning environment. Students were first informed about the procedure for the training session. Next, the use of the learning environment was briefly explained. Students in the Control condition were allotted 35 minutes for their training. Students in the Voice and Agent condition received an additional 5 minutes to compensate for time that might be lost in listening to the agent. Directly after finishing training and completing the post-training questionnaires, students took the post-test. Students were given a maximum of twenty minutes to complete the post-test.

Results

Research Question 1: Do conditions and gender affect changes in motivation during training?

Because the slope of the within-subject factor (time) was found to differ for the first two moments of measurement and the last two moments, two separate analyses were conducted. A repeated measures ANOVA with condition and gender as fixed factors, and with the scores for *task-relevance* (R1 and R2) at the first and second measurement points as dependent variables showed a trend for time, $F(1,55) = 3.12, p = 0.08$. There was a slight increase in the students' perceptions of task-relevance. The same analysis for the last two measurement points (R2 and R3) revealed no significant effect.

A similar repeated measures ANOVA with condition and gender as fixed factors, and with the scores for *self-efficacy belief* (SE1 and SE2) at the first and second measurement points as dependent variables showed a significant effect of time, $F(1,55) = 29.71, p = 0.00$. The scores for self-efficacy rose considerably, $d = 1.42$. Between-subjects analyses indicated that the average score of the girls was significantly lower than that of the boys, $F(1,55) = 30.34, p = 0.00, d = 1.43$. The same analysis for the last two measurement points (SE2 and SE3) revealed a trend for *time * gender * condition*, $F(1,52) = 2.41, p = 0.10$. This outcome is in line with the prediction. Figure 3 shows that the self-efficacy belief of the girls continued to rise in the Agent and Voice conditions, and stayed equal in the Control condition. For boys the self-efficacy belief decreased in the Agent and Voice conditions, but rose in the Control condition (see Figure 4). Between-subjects analyses indicated that the average score of the girls was significantly lower than that of the boys, $F(1,55) = 18.11, p = 0.00, d = 1.11$.

Together with an effect for self-efficacy belief, also a change in *error rate* was found. Compared to the error rates on the assignments for the subtopics one (11.8%) and two (13.4%) the error rate for the third subtopic almost doubled (25.5%). A repeated measures ANOVA with condition and gender as fixed factors, and with the scores for the error rates for the first and second subtopic showed a *gender* effect, $F(1,55) = 12.01, p = 0.00, d = 0.93$ (subtopic 1: girls 15.5% and boys 7.4%; subtopic 2: girls 17.2% and boys 10.0%). For the second and third subtopic there was a significant effect of time, $F(1,52) = 26.94, p = 0.00, d = 1.39$. For the third subtopic the error rates of the girls (26.5%) was comparable to that of the boys (24.5%).

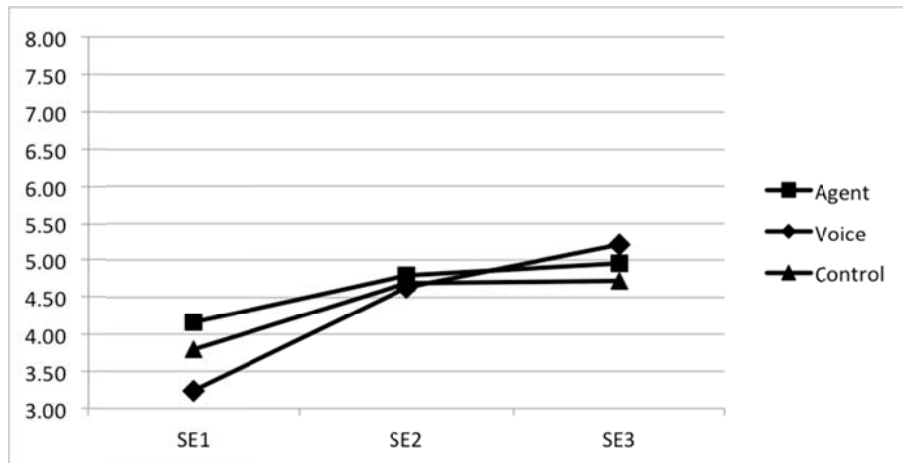


Figure 5. Self-efficacy belief of girls.

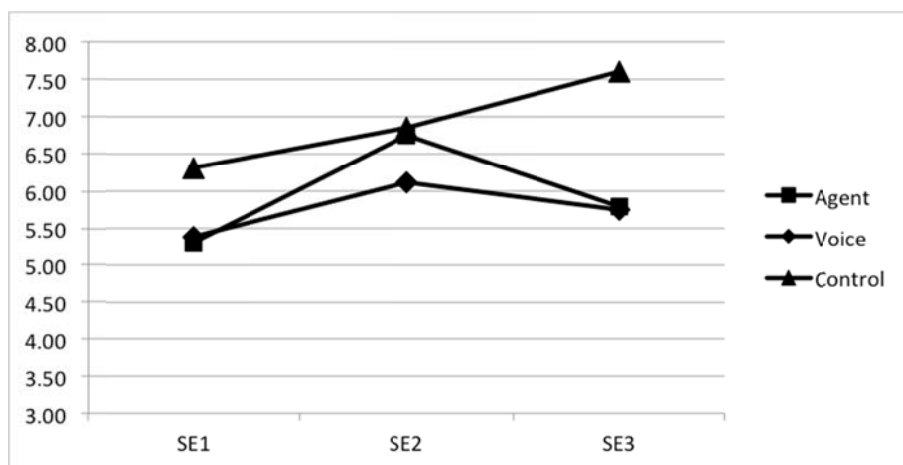


Figure 6. Self-efficacy belief of boys.

Research Question 2: Do conditions and gender affect motivation after training?

A repeated measures ANOVA with condition and gender as fixed factors and mean scores for perceptions of *task-relevance* before and after (R1 and Ra) training as dependent variables showed a significant effect of *time*, $F(1,55) = 8.47$, $p = 0.00$, $d = 0.76$. Students' perceptions of task-relevance increased (see Table 1). There was no effect of condition or gender.

A similar repeated measures ANOVA with condition and gender as fixed factors and mean scores on *self-efficacy belief* before and after (SE1 and SEa) training as dependent variables, showed a significant effect of *time*, $F(1,55) = 64.66$, $p = 0.00$, $d = 2.00$, and a trend for *time * gender*, $F(1,55) = 3.50$, $p = 0.07$. The interaction effect indicated that the average self-efficacy belief of the girls increased more than that of the boys. Between-subjects analyses indicated that the average score of the girls was significantly lower than that of the boys, $F(1,55) = 20.58$, $p = 0.00$, $d = 1.18$ (see Table 2).

Boys and girls gave different credibility ratings for the agent. A statistically significant effect for gender was found, $F(1,19) = 6.77$, $p = 0.02$, $d = 1.57$. Girls rated the mAPA substantially higher and were more uniform in their appraisals than boys (girls' $M = 5.43$, $SD = 1.11$; boys' $M = 3.50$, $SD = 2.18$).

Table 1: Mean task-relevance scores for boys and girls.

	Boys		Girls	
	During	After	During	After
Agent	5.61 (1.96)	6.13 (2.56)	4.80 (1.83)	5.07 (1.80)
Voice	4.73 (1.72)	5.05 (1.92)	5.45 (1.66)	5.83 (1.74)
Control	6.04 (2.32)	6.02 (2.35)	5.61 (1.11)	5.65 (0.99)
<i>Total</i>	<i>5.46 (2.02)</i>	<i>5.73 (2.27)</i>	<i>5.27 (1.56)</i>	<i>5.50 (1.55)</i>

Scales have a maximum of 10; higher values indicate more positive appraisals.

Table 2: Mean self-efficacy scores for boys and girls.

	Boys		Girls	
	During	After	During	After
Agent	5.95 (1.37)	6.99 (1.18)	4.64 (1.24)	5.82 (1.88)
Voice	5.75 (1.88)	5.60 (1.69)	4.15 (1.59)	5.08 (1.82)
Control	6.92 (0.92)	7.48 (1.47)	4.38 (1.28)	5.23 (1.43)
<i>Total</i>	<i>6.20 (1.49)</i>	<i>6.69 (1.63)</i>	<i>4.40 (1.34)</i>	<i>5.39 (1.70)</i>

Scales have a maximum of 10; higher values indicate more positive appraisals.

Research Question 3: Do conditions and gender affect learning?

A repeated measures ANOVA, with condition and gender as fixed factors and scores on pre-test and post-test as dependent variables showed a significant effect of *time*, $F(1,55) = 60.23$, $p = 0.00$. Scores on the post-test ($M = 12.43$, $SD = 4.20$) were substantially higher than on the pre-test ($M = 8.56$, $SD = 2.68$) with an effect size of $d = 2.01$. Between-subjects analyses signaled a trend for *gender* $F(1,55) = 3.39$, $p = 0.07$, and *condition * gender*, $F(2,55) = 2.71$, $p = 0.08$. Detailed analyses indicated that the existing gender difference in pre-test scores tended to level off in the Agent and Voice conditions, with boys making less progress than girls. In contrast, boys had larger knowledge gains in the Control condition. Table 3 shows the pre-test and post-test scores of boys and girls in the three experimental conditions.

Table 3: Pre-test and post-test scores of boys and girls in the three experimental conditions

	Boys		Girls	
	Pre-test	Post-test	Pre-test	Post-test
Agent	9.00 (2.87)	13.30 (4.79)	8.55 (2.88)	12.09 (5.11)
Voice	9.20 (2.94)	11.00 (3.77)	8.20 (2.30)	12.90 (3.00)
Control	9.70 (2.79)	14.80 (4.73)	6.70 (1.64)	10.50 (2.51)
<i>Total</i>	<i>9.30 (2.78)</i>	<i>13.03 (4.58)</i>	<i>7.84 (2.41)</i>	<i>11.84 (3.78)</i>

The maximum score was 27.

Discussion and Conclusions

This study examined whether student motivation and learning in an inquiry learning environment could be improved with a motivational agent (mAPA). The agent primarily addressed task-relevance and self-efficacy belief, the two key constructs from expectancy-value theory (Eccles & Wigfield, 2002). The design strategies from Keller's ARCS-model (1987, 2010) were used for constructing the agent's comments on these aspects.

Significant effects of the mAPA on student motivation during training were found only for self-efficacy belief. The data indicated that there was an overall increase in this belief from the first to the second measurement point. Because there was no effect of condition or gender, the re-appraisal of the students may have stemmed from having developed a better understanding of what the learning environment expected from them and how it supported their behavior. After this attunement, the self-efficacy belief of the girls kept rising in the two experimental conditions and stayed equal in the control condition. For boys the self-efficacy belief decreased in the Agent and Voice conditions, but rose in the Control condition. Analyses revealed that the shift coincided with a strong increase in error rates. For the third

subtopic students made almost twice as many mistakes in their assignments as for the first two subtopics. Perhaps the rise in error rates, in combination with the gender difference in credibility rating of the agent, worked out especially favorable for girls, making them more susceptible to the comments of the mAPA. With regard to the agent appraisals it is important to note that the agent deliberately used the word “I” to give students the impression that this was really a fellow student, rather than a guide because the latter would be more likely to use the word “you” in her comments. The idea was that the model-target similarity would become more salient when students could explicitly identify with the agent. For the girls this seemed to have worked out well.

An alternative explanation for the interaction may be that the effect of the motivating agent hinged on the strength of the student’s initial self-efficacy belief. That is, the girls’ low self-efficacy score at the beginning of the training may signal a fragile starting point, one that is likely to increase with support under special circumstances (i.e., an unexpected rise of the task difficulty level).

Analyses of the motivation scores after training indicated that students’ perceptions of task-relevance and self-efficacy had significantly increased under the influence of having worked with the inquiry learning environment. There was not a main effect of condition, nor of gender. A significant interaction for time and gender showed that, compared to boys, the self-efficacy belief of girls had increased more, but still remained significantly below that of the boys. The responses to the motivation questions after training refer to new assignments for the subtopics that students had just been practicing on. In other words, these data signal not just an increase over time, but of what the students take away with them from having had this experience.

Substantial learning gains in all conditions speak favorably of the basic set-up of the inquiry learning environment. That is, all groups benefitted from the guidance that was offered through the model progression and assignments in the environment ([de Jong & van Joolingen, 1998](#)). Learning may also have been positively influenced by attractiveness of the interface. Students had already indicated in the pilots that they liked to work with the inquiry learning environment; they found the environment an attractive method of learning about the topic of motion. However, [Dehn and Van Mulken \(2000\)](#) warned that when the interface is visually attractive and engaging from the start, the presence of an APA might not make much of a difference.

An interesting trend for an interaction between condition and gender on learning gains was found. Detailed analyses indicated that this trend stemmed from boys doing better in the control condition and girls doing better in the experimental conditions. Considering the fact that motivational support can only indirectly contribute to an effect on learning, the presence of a trend in the predicted direction is a promising finding.

When research on APAs moved into the arena of education, this led to a gradual shift in attention regarding the agent’s features. From an emphasis on the agent’s visual and auditory presence (external properties), the design efforts began to concentrate more on the agent’s actions or instructional methods (internal properties). For educational purposes the latter are considered to be critical ([Moreno, 2005](#)), but the distinction is not as black and white as it may appear. There are two caveats.

One, the agent should probably be sufficiently believable or credible to affect the learner. This requires attention to each and every facet of the agent’s presence. In studies that specifically attend to voice, it is frequently mentioned that it should have certain qualities that are hard to achieve in computer-generated voices. For this reason, the agent’s comments are typically recordings of a human voice (see [Domagk, 2010](#); [Hershey, Mishra, & Altermatt, 2005](#)). In our study, the agent’s voice was completely computer-generated, and at times

somewhat monotonous. Still, it appeared to have worked well. Whether or not qualities such as likability and use of the proper clothing for representing a certain role are also important for the agent's credibility is an open question. The same can be said for the agent's nonverbal, affective expressions. In our study, the mAPA regularly ventilated her moods nonverbally to complement her comments, just as the motivational agent did in the studies of Arroyo et al. (2009, 2011). It is not clear whether these expressions really mattered. Based on the findings that the effects for the Agent and Voice condition were similar in our study, questions the importance of the agent's embodiment. It suggests that, for this learning environment, it may not be necessary to spend a considerable effort on realizing a technically demanding solution where a more simple and low-cost method may also suffice (see Clark & Choi, 2005).

Two, it is not entirely clear or self-evident which actions or methods are vital for a *pedagogical* agent to affect motivation and learning. In Moreno's (2005) discussion of the agent's internal properties, strategies are mentioned such as cognitive load reduction, external memory expansion, feedback, modeling, and guidance. The list could be expanded with worked examples, advance organizers, and adjunct questions, among others. All of these features are clearly relevant for education. They have proven their value in many empirical studies and form candidates for the design of a pedagogical agent. But what about the agent's general communication strategies that have been on the agenda of agent studies right from the start? Here, too, one should assume that the agent needs to satisfy minimal communicational demands. In addition, recent research suggests that educational benefits may also be found from APAs based on specific communication theories. That is, Wang, et al. (2008) developed and tested the effects of a mAPA whose communication strategies, dubbed 'motivational tactics', were based on Brown & Levinson's (1987) politeness theory (i.e., freedom of choice or student autonomy, and approval or performance feedback). Their experiment pitted two politeness conditions against each other. In one, students communicated with an impolite mAPA who disregarded learner face (e.g., "You did not save your factory parameters. Save them now."). In the other, a polite mAPA gave students feedback that promoted learner face (e.g., "How about if we save our factory now?") and mitigated face threats (e.g., "Do you want to save the factory settings?"). A difference was predicted but not found for the two groups in students' self-efficacy belief, but a significant effect was found for learning. Students who had worked with the polite agent scored better on a learning test taken after training. Somewhat surprisingly, an earlier study had yielded exactly the opposite results (Wang, Lewis Johnson, Rizzo, Shaw, & Mayer, 2005). That is, self-efficacy belief was raised significantly more by the polite agent than by the impolite agent, but agent condition had no effect on learning.

The studies by Wang et al. (2005, 2008) exemplify a design approach in which the agent's actions are based on a mixture of general guidelines for communication (e.g., politeness norms or rules) and specific guidelines for educational settings (e.g., student autonomy and performance feedback). This is probably a necessary condition for designing the agent's internal properties. Agent communications should not be based solely on what (psycho)linguistic theory dictates as proper, nor can agents attend solely to educationally relevant actions. The equivocal outcomes for Wang et al. indicate that considerable fine-tuning of the agents' verbal behaviors may be needed to achieve robust results.

To conclude, there is nothing amiss with an emphasis on designing the agent's internal properties, as long as sufficient attention is also paid to the agent's external properties. The question that then invariably remains is when an agent is credible enough. Research indicates that this question is more difficult than it seems. Also, one should probably aim for complex mixed designs of APAs to influence student motivation and learning. It may be an illusion to believe that it is possible to design a "purely" motivational agent that works. The multifaceted

nature of designing a ‘mixed’ agent may be an important reason why so many empirical studies have reported equivocal outcomes ([see Heidig & Clarebout, 2011](#)).

References

- Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Educational Psychology Review*, *103*, 1-18.
- André, E., Mailer, J., & Rist, T. (1996). *The PPP persona: a multipurpose animated presentation agent*. Paper presented at the workshop on advanced visual interfaces, Gubbio, Italy.
- Arroyo, I., Woolf, B. P., Cooper, D. G., Burleson, W., & Muldner, K. (2011). *The impact of animated pedagogical agents on girls' and boys' emotions, attitudes, behaviors and learning*.
- Arroyo, I., Woolf, B. P., Royer, J. M., & Tai, M. (2009). *Affective gendered learning companion, International Conference on Artificial Intelligence and Education*, Brighton, England.
- Atkinson, R., Mayer, R., & Merrill, M. (2005). Fostering social agency in multimedia learning: Examining the impact of an animated agent's voice. *Contemporary Educational Psychology*, *30*, 117-139. doi: 10.1016/j.cedpsych.2004.07.001
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman and Company.
- Bates, J. (1994). The role of emotion in believable agents. *Communications of the ACM*, *37*, 122-125.
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage* (Vol. 4): Cambridge University Press.
- Cassell, J. (2000). Embodied conversational interface agents. *Communications of the ACM*, *43*, 70-78.
- Catsambis, S. (1995). Gender, race, ethnicity, and science education in the middle grades. *Journal of Research in Science Teaching*, *32*, 243-257. doi: 10.1002/tea.3660320305
- Choi, S., & Clark, R. E. (2006). Cognitive and affective benefits of an animated pedagogical agent for learning english as a second language. *Journal of Educational Computing Research*, *34*, 441-466.
- Clark, R. E., & Choi, S. (2005). Five design principles for experiments on the effects of animated pedagogical agents. *Journal of Educational Computing Research*, *32*, 209-225.
- Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features and redundancy. *Journal of Educational Psychology*, *94*, 428-434.
- de Jong, T. (2006a). Scaffolds for computer simulation based scientific discovery learning. In J. Elen & R. E. Clark (Eds.), *Dealing with complexity in learning environments* (pp. 107-128). London: Elsevier Science Publishers.
- de Jong, T. (2006b). Technological advances in inquiry learning. *Science*, *321*, 532-533.
- de Jong, T., & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, *68*, 179-201.
- de Jong, T., van Joolingen, W. R., Veermans, K., & van der Meij, J. (2005). Authoring discovery learning environments: In search for reusable components. In J. M. Spector & D. A. Wiley (Eds.), *Innovations in instructional technology: Essays in honor of M. David Merrill* (pp. 11-28). Mahwah, NJ: Erlbaum.
- Dehn, D. M., & Van Mulken, S. (2000). Impact of animated interface agents: a review of empirical research. *International Journal of Human Computer Studies*, *52*, 1-22.
- Domagk, S. (2010). Do pedagogical agents facilitate learner motivation and learning outcomes?: The role of the appeal of agent's appearance and voice. *Journal of Media Psychology: Theories, Methods, and Applications*, *22*, 84.
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality and development*. Philadelphia: The Psychology Press
- Dweck, C. S. (2007). The perils and promises of praise. *Educational Leadership*, *65*, 32-39.
- Eccles, J. S., & Wigfield, A. (2002). Motivation beliefs, values, and goals. *Annual Review of Psychology*, *53*, 109-132. doi: 10.1146/annurev.psych.53.100901.135153

- Eysink, T. H. S., de Jong, T., Berthold, K., Kolloffel, B., Opfermann, M., & Wouters, P. (2009). Learner performance in multimedia learning arrangements: An analysis across instructional approaches. *American Educational Research Journal*, *46*, 1107-1149. doi: 10.3102/0002831209340235
- Graesser, A., & McNamara, D. (2010). Self-regulated learning in learning environments with pedagogical agents that interact in natural language. *Educational Psychologist*, *45*, 234-244. doi: 10.1080/00461520.2010.515933
- Haake, M. (2009). *Embodied pedagogical agents. From visual impact to pedagogical implications*. Department of Design Sciences, Lund University, Sweden.
- Hagemans, M., van der Meij, H., & de Jong, T. (submitted). Regulating the inquiry learning process. The effects of a concept map based support tool for simulation-based learning.
- Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning? *Educational Research Review*, *6*, 27-54. doi: 10.1016/j.edurev.2010.07.004
- Hershey, K., Mishra, P., & Altermatt, E. (2005). All or Nothing: Levels of Sociability of a Pedagogical Software Agent and its Impact on Student Perceptions and Learning. *Journal of Educational Multimedia and Hypermedia*, *14*, 113-127.
- Keller, J. M. (1987). Development and use of the ARCS model of instructional design. *Journal of instructional development*, *10*, 2-10.
- Keller, J. M. (2010). *Motivational design for learning and performance: The ARCS model approach*. New York, NY: Springer Verlag.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*, 75-86. doi: 10.1207/s15326985ep4102_1
- Lau, S., & Roeser, R. W. (2002). Cognitive abilities and motivational processes in high school students' situational engagement and achievement in science. *Educational Assessment*, *8*, 139-162. doi: 10.1207/s15326977ea0802_04
- Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., & Bhogal, R. S. (1997). *The persona effect: affective impact of animated pedagogical agents*. Paper presented at the Proceedings of the SIGCHI conference on human factors in computing systems.
- Mattern, N., & Schau, C. (2002). Gender differences in science attitude-achievement relationships over time among white middle-school students. *Journal of Research in Science Teaching*, *39*, 324-340.
- Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist*, *59*, 14-19. doi: 10.1037/0003-066x.59.1.14
- Mayer, R. E. (2005). *Cambridge handbook on multimedia learning*. New York, NY: Cambridge University Press.
- Moreno, R. (2005). Multimedia learning with animated pedagogical agents. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 507-523). New York, NY: Cambridge University Press.
- Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001a). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction*, *19*, 177-213.
- Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001b). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction*, *19*, 177-213.
- Moreno, R., Reislein, M., & Ozogul, G. (2010). Using virtual peers to guide visual attention during learning. *Journal of Media Psychology: Theories, Methods, and Applications*, *22*, 52-60.
- Moundridou, M., & Virvou, M. (2002). Evaluating the persona effect of an interface agent in a tutoring system. *Journal of Computer Assisted Learning*, *18*, 253-261.

- Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, *25*, 1049-1079. doi: 10.1080/0950069032000032199
- Paiva, A., & Machado, I. (1998). *Vincent, an autonomous pedagogical agent for on-the-job training*. Paper presented at the Intelligent tutoring systems.
- Picard, R. W., & Klein, J. (2002). Computers that recognise and respond to user emotion: Theoretical and practical implications. *Interacting with Computers*, *14*, 141-169.
- Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education. Theory, research, and applications. 2nd Edition*. Upper Saddle River, NJ: Merrill Prentice Hall.
- Plant, E. A., Baylor, A. L., Doerr, C. E., & Rosenberg-Kima, R. B. (2009). Changing middle-school students' attitudes and performance regarding engineering with computer-based social models. *Computers & Education*, *53*, 209-215.
- Reeves, B., & Nass, C. (1996). *How people treat computers, television, and new media like real people and places*: CSLI Publications and Cambridge University Press.
- Scalise, K., Timms, M., Moorjani, A., Clark, L. K., Holtermann, K., & Irvin, P. S. (2011). Student learning in science simulations: Design features that promote learning gains. *Journal of Research in Science Teaching*, *48*, 1050-1078. doi: 10.1002/tea.20437
- Shaw, E., Johnson, W. L., & Ganeshan, R. (1999). *Pedagogical agents on the web*. Paper presented at the third annual conference on autonomous agents, Seattle, Washington, United States.
- van Joolingen, W. R., & de Jong, T. (2003). SimQuest: Authoring educational simulations. In T. Murray, S. Blessing & S. Ainsworth (Eds.), *Authoring tools for advanced technology educational software: Toward cost-effective production of adaptive, interactive, and intelligent educational software* (pp. 1-31). Dordrecht: Kluwer Academic Publishers.
- van Welbergen, H., Reidsma, D., Ruttkay, Z. M., & Zwiers, J. (2010). Elckerlyc, a BML Realizer for continuous, multimodal interaction with a virtual human. *Journal on Multimodal User Interfaces*, *3*, 271-284.
- 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*, 98-112. doi: 10.1016/j.ijhcs.2007.09.003
- Wang, N., Lewis Johnson, W., Rizzo, P., Shaw, E., & Mayer, R. E. (2005). *Experimental evaluation of polite interaction tactics for pedagogical agents*. Paper presented at the International Conference on Intelligent user Interfaces.
- White, B. Y., & Frederiksen, J. R. (1990). Causal model progressions as a foundation for intelligent learning environments. *Artificial Intelligence*, *42*, 99 -57.
- Yeung, K. S., Kuppan, L., Kadir, M. S., & Foong, S. K. (2011). Boys' and girls' self-beliefs, engagement, and aspirations in physics. *International Journal of Learning*, *17*, 397-418.