Pedagogical agents as learning companions: the impact of agent emotion and gender

Y. Kim,* A.L. Baylor† & E. Shen†

*Department of Instructional Technology, Utah State University, 2830 Old Main Hill, Logan, UT, USA
†Department of Educational Psychology and Learning Systems, Florida State University, 307 Stone Building, Tallahassee, FL, USA

Abstract The potential of emotional interaction between human and computer has recently interested researchers in human–computer interaction. The instructional impact of this interaction in learning environments has not been established, however. This study examined the impact of emotion and gender of a pedagogical agent as a learning companion (PAL) on social judgments, interest, self-efficacy, and learning. Two experiments investigated separately the effects of a PAL’s emotional expression and empathetic response. Experiment 1 focused on emotional expression (positive vs. negative vs. neutral) and gender (male vs. female) with a sample of 142 male and female college students in a computer literacy course. Experiment 2 investigated the impact of empathetic response (responsive vs. non-responsive) and gender with 56 pre-service teachers. Overall, the results yielded main and interaction effects of PAL emotion and gender on the dependent variables. In particular, the PAL’s empathetic response had a positive impact on learner interest and self-efficacy; PAL gender had a positive impact on recall. The findings imply that the emotion and the gender of the digital learning companion could be utilized to optimize college students’ motivation and learning.

Keywords affective computing, human–computer interaction, learning companions, pedagogical agents, social interaction, virtual peers.

Introduction With the advance of technology, we can design computer-based learning environments to support simulated social interactions between learner and computer. The social context and social interaction among participants in the learning process are considered crucial for intellectual and social development (Piaget 1995; Rogoff 1995; John-Steiner 1996; Matusov & Hayes 2000; John-Steiner & Mahn 2003). Granted, the ideal form of instruction might well be one-on-one human tutoring (Bloom 1984), where a learner can benefit from individualized cognitive guidance through dynamic social interaction. But given the challenges of providing such an ideal environment of one-on-one human tutoring, computer-based instruction has been developing systems for providing individualized cognitive guidance more flexibly than is possible with one-on-one human tutoring. Some of those systems have reported success (Koedinger & Anderson 1997), but their impact was – as is to be expected – far weaker than that of human tutoring. So, it seems prudent to ask what has been missed in conventional computer-based tutoring systems as compared with human tutoring. Recent research in social cognition seems to support the need to include social context in computer-based learning.

Pedagogical agents are animated life-like characters (Johnson et al. 2000) designed to enhance learning and/or motivation by simulating social interaction with...
a learner. The provision of simulated social interaction may distinguish pedagogical agents from traditional computer-based tutoring, seemingly offering a unique instructional impact (Kim & Baylor, 2006). In particular, the benefits of peer–peer interaction for a learner’s cognitive and affective attainments have been supported both theoretically (Tudge et al. 1996; Bandura 1997; Matusov & Hayes 2000) and empirically (Jacobs et al. 1996; Kumar & Harizuka 1998; Mathes et al. 1998; Davenport & Howe 1999; Fuchs et al. 1999; Topping & Ehly 2001). Computer-based learning environments might emulate such interaction by integrating a pedagogical agent as a learning companion (PAL) to simulate human–peer interaction.

**PAL emotion**

Given that emotion influences an individual’s rational thinking, decision making, social memory, judgements, and learning (Adolphs & Damasio 2001;Forgas 2000), efforts have been made to implement emotional interaction between human and computer (Picard 2000; Burleson et al. 2004). The positive impact of this interaction in gaming environments is often reported (Klein et al. 2002; Scheirer et al. 2002). Thus, it sounds plausible that a learning companion equipped with emotional capabilities might play positive influential roles in learning environments.

As human–computer interaction seems generally to imitate human-to-human interaction in the real world (Reeves & Nass 1996), so might a learner be able to build virtual social relations, e.g. virtual friendship, with a virtual companion (Bickmore et al. 2005; Gulz 2005). For instance, children who played with the virtual peer Sam listened to Sam’s stories carefully and mimicked Sam’s linguistic styles, which might validate the virtual peer as actually playing a social role (Ryokai et al. 2003). The children might feel an affiliation with Sam that in turn might induce their behavioural changes. The implication is that a virtual companion should be perceived by a learner as believable and natural, perhaps comrade-like, so as to have meaningful instructional impact. At the centre of believability of an agent might be its capability to express emotions (Bates 1994).

Furthermore, an individual’s emotion is closely related to and influenced by social contexts (Martin 2000; Saarni 2001). In classrooms, the affective states of teachers and peers function as a social context, influencing a learner’s affective characteristics, e.g. the learner’s emotions, self-conception, and motivation (Sutton & Wheatley 2003). Teachers’ expressions of their emotions influence students’ attributions of their successes or failures (Weiner 2000). Also, students having a negatively affective instructor typically experience negative affect and handicap themselves significantly more than do students having a positively affective instructor. Given those implications, it is highly plausible to expect that a learning companion’s emotion might play a role in influencing the affective and cognitive characteristics of a learner in computer-based environments. Also, an important mechanism of emotional development from childhood through adolescence is socialization by peers (Asher et al. 1996). In simulating a human peer, a virtual companion may model various emotional expressions and reactions to guide a learner’s positive emotional experiences in the process of learning.

**PAL gender**

Gender-related social stereotypes in the real world seem consistently projected to computing environments. (Carli 1999, 2001). In a study with college students, Lee (2003) showed that when a computer was categorized as male or female, users perceived a male computer as more credible and conformed to the advice of the male computer on masculine topics such as sports, whereas they perceived a female computer as more credible and conformed to the advice of the female computer on feminine topics such as cosmetics or fashion. This projection is demonstrated more clearly with anthropomorphized agents. Baylor and Kim (2004) found that both male and female college students’ expectations and perceptions of pedagogical agents were significantly differentiated by agent gender. Likewise, Moreno et al. (2002) reported that learners applied gender stereotypes to animated agents, and these stereotypic expectations affected their learning. Given that in everyday life males and females differ in their emotional expressions, their empathic accuracy, and their emotional behaviours (Brody 1999), when investigating a learning companion’s emotional impact, it is natural to inquire about the impact of PAL gender.

Although human–computer interaction research suggests the instructional potential of the emotion and
gender of a learning companion, few studies have provided the empirical evidence for their impact. This study thus implemented two classroom-based experiments to examine the impact of the emotion and gender of a learning companion on college students’ social judgements of its instructional functionality, on their interest, and self-efficacy beliefs in the task, and on their learning. In general, agent affect is defined as an agent’s ability to recognize the learner’s affective states, to express its own emotion, and to respond to the learner’s (Picard 1997; Hudlicka 2003). Each capability requires unique technologies and resources for implementation. For instance, affect recognition is typically engineered by hardware technologies (Burleson et al. 2004). The affective expression and response of an agent are engineered by software technology using scripting languages and tools. Hence, it is possible to examine the impact of each capability separately to determine their unique instructional efficacy. This study examined the impact of a PAL’s emotional expression (Experiment 1) and empathetic response (Experiment 2) separately in two controlled experiments. Detailed descriptions of the experiments follow.

**Experiment 1**

This experiment investigated the impact of a learning companion’s emotional expressions (positive vs. negative vs. neutral) and gender (male vs. female) on learners’ affective and cognitive characteristics, as determined by four measures: the learners’ social judgements of the companion’s instructional functionality, their interest in the task and in working with the companion, their self-efficacy beliefs in the task, and their learning.

**Participants**

For the study population, college students were considered appropriate, given that users’ perceptual reactions to pedagogical agents are varied (Gulz 2004), especially among the users in their 20s (Bickmore & Picard 2003). Convenience sampling was used to obtain the participants, who were new to the task domain of instructional design. A total of 142 college students in a computer-literacy course in a large public university located in the Southeast US voluntarily participated in the study, implemented as optional to other class activities. Approximately 40% of the participants were male and 60% were female. Sixty-seven per cent of the participants identified themselves as Caucasian; 13% as African American; 10% as Hispanic; 2% as Asian; and 8% as other. The average age of participants was 20.25 years (sd = 2.27), with a range of 18 through 32. A pretest question asking about their prior experience in the domain on a Likert scale did not result in any statistically significant variations across the experimental groups.

**Instructional module, E-Learn**

The instructional module was E-Learn, an agent-based research environment introducing novice learners to the basic concepts and procedures of instructional design, so that they might design an e-learning class. E-Learn included three phases of instructional planning: Introduction, Goals, and Planning. Intro presented a case scenario in which the participants, playing the role of instructor, were asked to convert a classroom-based course on time management for freshmen into an e-learning course. The participants’ tasks were to acquire knowledge about the concepts and procedures of instructional planning (presented by a learning companion) and to write their plans for the e-learning course in Goals and Planning, the second and third phases. In Goals, they wrote instructional goals and objectives. In Planning, they wrote instructional strategies, activities, and sequences. As students entered the E-Learn program, the learning companion, named Chris, appeared and introduced himself/herself as a peer. As students proceeded to the next steps, Chris provided context-specific information and help messages at their request. The Appendix presents the screen excerpts of E-Learn.

**PAL design**

Male and female agents, both named Chris, were developed using Poser 5 (www.e-frontier.com), a 3D-image/animation-design tool and Mimic Pro 2 (www.pluginz.com), a voice/affect-editing tool. The animation files created in Poser 5 were converted to Macromedia Flash (www.adobe.com) movies for compression and were later integrated into E-Learn. Chris was designed to look about 20 years old, was casually dressed, and spoke informally, sometimes using slang. The
participants estimated Chris’s age at an average of 20.39 years ($sd = 7.94$) with a range of 15 through 32.

**Independent variables**

**PAL emotional expression**

The PALs’ emotional expressions were achieved through verbal and facial expressions, tone of voice, and head movements, as supported by human emotion research indicating that people express and perceive emotions mostly through facial expressions, sounds, and body movements, together with verbal manifestations. According to Keltner and Ekman (2000), the face is the primary source for expressing distinct emotions nonverbally. The distinctive features of individuals’ voices also have a powerful influence on how people decipher emotional messages (Bradley & Lang 2000; Bachorowski & Owren 2002). Bodily movements are clearly differentiated according to positive or negative feelings (Cacioppo et al. 1993; Chen & Bargh 1999).

PAL emotion had three expressions: positive, negative, and neutral. Psychologists typically classify affect as positive if it involves pleasure (e.g. happiness or satisfaction) and as negative if it includes distress (e.g. frustration or anger) (Ottati et al. 1997; Sutton & Wheatley 2003). Thus, in the positive emotion condition, Chris had a happy, smiling face and an engaging posture, with eye contact and head nodding. The background tone was red. The PAL also verbally expressed its mood, such as ‘. . . this task looks fun’ and ‘. . . completing this will be rewarding.’ The participants perceived positive Chris as significantly more ‘happy looking’ than negative Chris ($P < 0.001$). In the negative emotion condition, Chris had a somber and rather frowning face and an aloof posture, with evasive eye gaze and less head nodding. The background tone was blue. Chris expressed verbal mood with such statements as ‘I don’t feel like doing this, but we have to anyway.’ These emotional comments were very brief and did not affect, in general, total instructional time across the conditions. The participants perceived negative Chris as significantly more ‘sad looking’ than positive Chris ($P < 0.001$). The verbal expressions of positive versus negative Chris were sharply contrasted, given that the purpose of the study was to investigate whether the agents’ positive and negative emotions affect the learners differently. The agents’ emotional states were clearly communicated to the learners to ensure the validity of the findings. In the neutral condition, Chris did not express emotions at all. The background tone was grey. Overall, the adjustment of the emotion parameters in the voice/affect editing tool, Mimic Pro 2, determined the degree of positive, negative, and neutral expressions. The factual information provided by Chris was identical across the three conditions. The Appendix presents the positive and negative variations of Chris.

**PAL gender**

Either a male or a female version of Chris was included depending on the experimental conditions. The two versions were identical in comments, gestures, and emotional expressions, differing only in image and voice. Given that voice was a significant indicator for social presence (Nass & Brave 2005), voices of male and female college students were recorded.

**Dependent variables**

**Social judgements**

Social judgements referred to learners’ judgements of PAL instructional functionality and persona as measured by a questionnaire modified from Agent Persona Instrument (Baylor & Ryu 2003). The questionnaire consisted of three sub-measures: ‘facilitating learning’ (four items), ‘engaging’ (four items), and ‘human-like’ (four items). Learners rated their PAL on a scale ranging from 1 (Strongly disagree) to 5 (Strongly agree). Item reliability of each category was evaluated as coefficient $\alpha = 0.91, 0.80,$ and $0.81$, respectively.

**Interest**

Getzels (1966) defined interest as a ‘disposition organized through experience which impels an individual to seek out particular objects, activities, understandings, skills, or goals for attention or acquisition’. In this study, interest referred to a learner’s disposition towards working with the PAL and towards the task of instructional planning. Anderson and Bourke (2000) suggested that the range of interest is best expressed on the scale of ‘interested–disinterested.’ According to the suggestion, a questionnaire consisting of seven items was developed, with a scale ranging from 1 (Strongly disagree) to 5 (Strongly agree). Item reliability was assessed as coefficient $\alpha = 0.87$. Interest was measured before and
after the intervention. To control for test–retest effect, a split-half technique was used, where the first four items out of seven were implemented in the pretest, and the last five items were implemented in the post-test.

Self-efficacy
In general, self-efficacy is defined as an individual’s beliefs in his/her competency to perform a particular task required to reach a goal (Bandura 1986, 1997; Weiner 1992). In this study, self-efficacy referred to the learners’ beliefs about their competency in the task of e-learning design. The direction of self-efficacy is best captured by ‘I can’ versus ‘I can’t’ (Weiner 1992) or ‘How sure are you?’ (Bandura & Schunk 1981; Pajares 1996). A questionnaire with five items was scaled from 1 (Strongly disagree) to 5 (Strongly agree). Item reliability was evaluated as coefficient $\alpha = 0.95$. Learners’ self-efficacy beliefs were measured before and after the intervention. To control for test–retest effect, a split-half technique was used, where the first two items out of five were implemented in the pretest, and the final three items were implemented in the post-test.

Learning
Learning was measured by open-ended recall and application post-test questions. In the recall question, students were asked to write all the ideas conveyed by the PALs about designing an e-learning class. To score recall, two instructional designers with master’s degrees in Instructional Design counted independently the legitimate idea units from the information provided by the PAL in the module. Before scoring students’ answers, the raters agreed upon a total of 15 idea units. The student’s recall was then scored in terms of those idea units, a procedure used in previous studies (Mayer & Gallini 1990; Plass et al. 1998; Kim et al. 2006). Interrater reliability was evaluated with Cohen’s $\kappa = 0.94$. In the application question, the participants were asked to write a brief e-learning plan according to a new scenario. Students’ instructional plans were scored by the two instructional designers, using a scoring rubric scaled 1 (Very poor) through 5 (Excellent). The scoring rubric (Baylor & Kim 2004; Kim et al. 2006) focused on how specific the students’ plans were in terms of the topic and instructional strategies. Interrater reliability was evaluated as Cohen’s $\kappa = 0.97$.

Procedures
The experiment was conducted during a regular session of a computer-literacy course. Participants were randomly assigned to one of the six conditions. The researcher administered the experiment with assistance from the course instructors. The participants first logged onto the web-based E-Learn module by entering demographic information. They answered questions on prior experience, pretest interest, and self-efficacy. Then they performed the task. The participants were given as much time as they needed to finish the entire process, approximately 30 min, with individual variations. Lastly, the participants answered the post-test questions.

Design and analysis
The study used a $3 \times 2$ factorial design, in which the variables included PAL emotional expressions (positive vs. negative vs. neural) and PAL gender (male vs. female). Given the multiple dependent measures, multivariate analysis techniques were used to control for the inflation of family-wise error rates. For learning and social judgements, a multivariate analysis of variance (MANOVA) was conducted. For interest and self-efficacy, a multivariate analysis of covariance (MANCOVA) was conducted, with prior interest and self-efficacy as covariates. The significance level for all the analyses was set at $\alpha < 0.05$.

Results
A preliminary analysis of the data did not show any evidence of violations of statistical assumptions. Examination of scatter plots supported the assumption of normality and revealed linear relationships for all tests. Box’s test of equality of covariance supported the equal covariance assumptions for multivariate analyses.

Social judgements
The MANOVA conducted as protected testing revealed the significant main effect for PAL emotional expressions, Wilks’ Lambda $= 0.70$, $F_{6,252} = 8.33$, $P < 0.001$, partial $\eta^2 = 0.16$. The univariate results revealed a significant effect on ‘facilitating learning’, $F_{2,128} = 3.75$, $P < 0.05$. Students who worked with the positive PAL
(M = 2.59, sd = 1.07) rated the PAL as ‘facilitating to their learning’ significantly more than did students with the negative PAL (M = 2.12, sd = 0.93). The standardized effect size for this difference was Cohen’s d = 0.47, indicating a medium effect. Also, students who worked with the neutral PAL (M = 2.63, sd = 0.94) rated the PAL as ‘facilitating to their learning’ significantly more than did students with the negative PAL (M = 2.12, sd = 0.93). The standardized effect size for this difference was Cohen’s d = 0.55, indicating a medium effect. The results also revealed a significant effect on ‘engaging’. F_{2,128} = 14.77, P < 0.001. Students who worked with the positive PAL (M = 3.06, sd = 0.97) rated the PAL as significantly more ‘engaging’ than did students with the negative PAL (M = 2.25, sd = 0.83). The standardized effect size for this difference was Cohen’s d = 1.09, indicating a large effect. Also, students who worked with the neutral PAL (M = 3.12, sd = 0.77) rated the PAL as significantly more ‘engaging’ than did students with the negative PAL (M = 2.25, sd = 0.83). The standardized effect size for this difference was Cohen’s d = 1.09, again a large effect.

MANOVA indicated the marginal main effect for PAL gender (P = 0.07, partial \( \eta^2 = 0.05 \)). The univariate results showed that students working with the male PAL rated the PAL more positively than did students working with the female PAL on all the sub-measures: ‘facilitating learning’ (P < 0.05), ‘engaging’ (P < 0.05), and ‘human-like’ (P < 0.05). Thus, the authors were confident in concluding that the male PAL was perceived more positively than the female PAL.

The overall MANOVA revealed a marginal interaction effect of PAL emotions and gender (P < 0.07, partial \( \eta^2 = 0.05 \)). A visual inspection of this relationship suggested a partial interaction between PAL emotional expression and gender. To test the interaction in detail, partial-interaction tests were conducted. The tests examined the interaction effects between one treatment (PAL gender) by every orthogonal contrast on the other treatment (PAL emotional expression), controlling for the inflation of family-wise error rates. The results of the tests revealed three significant interactions between gender and emotional expression. First, the positive and negative emotions interacted with gender, F_{1,133} = 5.61, P < 0.05. When Chris expressed positive emotions, students judged male Chris more favourably than female Chris. This interaction, however, was not significant when Chris expressed negative emotions. Second, the positive and neutral emotions interacted with gender, F_{1,133} = 6.00, P < 0.05. When Chris expressed positive emotions, students judged male Chris more favourably. This interaction was not significant when Chris did not express emotions (neutral). Third, the positive versus negative and neutral emotions interacted with gender, F_{1,133} = 7.73, P < 0.01. When Chris expressed positive emotions, students judged male Chris more favourably than female Chris. This interaction was minimal when Chris expressed negative and neutral emotions. To summarize, only in the positive emotion conditions did the emotional expression and the gender of the learning companion interact significantly to influence the learners’ social judgements, mirroring in part the interaction between gender and emotions in the real world.

Interest

First, MANCOVA revealed a significant main effect for PAL gender, Wilks’ Lambda = 0.893, F_{5,127} = 3.05, P < 0.05, partial \( \eta^2 = 0.11 \). Students who worked with male Chris showed significantly higher interest in the task and in working with Chris than did students working with female Chris. There was no significant main effect for PAL emotional expression on interest.

Second, MANCOVA indicated a significant interaction effect between emotional expressions and gender, Roy’s Largest Root = 1.0, F_{5,128} = 2.51, P < 0.05, partial \( \eta^2 = 0.11 \). However, the univariate analysis did not show significance in this interaction effect. A visual inspection of the data suggested the interactive relationship of the two: the learners’ interest in male Chris’ conditions was differently patterned by the emotion types. A simple trend analysis was conducted to test the statistical significance of this finding. The results indicated a significant linear relationship of PAL gender only in the positive emotion condition, F_{1,126} = 5.21, P < 0.05. When Chris expressed positive emotions, students who worked with male Chris showed higher interest than did those working with female Chris. This result was consistent with the result on social judgements.

Self-efficacy

There were no significant effects for PAL emotional expression and gender on self-efficacy.
Learning

ANOVA revealed a significant main effect for PAL gender on learning, Wilks’ Lambda = 0.83, $F_{2,65} = 6.66$, $P < 0.01$, partial $\eta^2 = 0.17$. The univariate results revealed a significant main effect for gender on recall, $F_{1,66} = 7.08$, $P < 0.01$. Students who worked with male Chris ($M = 1.52$, $sd = 1.84$) achieved significantly higher recall scores than students working with female Chris ($M = 0.65$, $sd = 0.89$). The standardized effect size for this difference was Cohen’s $d = 0.60$, a medium effect according to Cohen’s guidelines.

In summary, the results from Experiment 1 revealed, first, that the emotional states of the digital learning companions clearly differentiated the learners’ social judgements of them. Second, the PAL’s gender and emotional expression interacted in positive emotions to influence learners’ social judgements and interest. Third, the male companion had a superior impact to the female companion on interest and recall. Given the significant effect of the emotional expression and the gender of the PALS from Experiment 1, the authors were further motivated to investigate another aspect of agent affective capabilities, emotional response. Experiment 2 was conducted to determine whether PAL empathetic responses and gender related to the learners’ affect and cognition.

Experiment 2

Experiment 2 investigated the effects of PAL empathetic responses (responsive vs. nonresponsive) and gender (male vs. female) on social judgements, interest, self-efficacy, and recall.

Participants

Participants were 56 pre-service teachers (11 male and 45 female) enrolled in a course in introductory educational technology. Convenient sampling was used with the consideration that, unlike in Experiment 1, the participants had to express their affective states and thus had to be more seriously engaged in their learning to be affectively aroused. The learning task was an integral part of their coursework; the intervention was implemented as a required class activity. The participants’ performances were reflected in the course grades. Sixty-eight per cent of the participants identified themselves as Caucasian; 12% as African American; and 20% as other. The average age of participants was 20.71 years ($sd = 2.92$) with a range of 18 through 34. As in Experiment 1, a pretest question asking about the participants’ prior experience in the domain did not result in any statistical differences across the experimental groups.

Instructional module

In the web-based module, the learners’ task was to develop an instructional plan for sixth graders who were learning the economic concept of supply and demand. The module included five steps: Introduction, Case Study, Blueprints, Plan, and Assessment, in each of which a PAL provided learners with context-specific information and suggestions. The steps were indicated by large buttons located at the top of the screens. The Introduction briefly explained learners’ task. The Case Study described a scenario to teach Anna, a sixth grader, the economic concept of supply and demand. In Blueprints, the participants wrote instructional goals or objectives in a text-box field. In Plan, the participants wrote instructional strategies and activities. In Assessment, the participants described the assessment plans to test Anna’s learning.

In addition, to enable a learner to express his/her affective states, a panel of emoticons (i.e. icons expressing emotions) popped up when the learner clicked a navigation button to move to the next phase. When the learner expressed affect by clicking an emoticon, the PAL verbally responded or not according to experimental conditions. The emoticons reflected six affective states commonly occurring in learning situations, as suggested by Kort et al. (2001): Interest, Boredom, Confidence, Anxiety, Satisfaction, and Frustration. The Appendix presents the emoticons and sample screen excerpts.

Independent variables

**PAL empathetic responses**

Empathetic responses, i.e. whether or not Chris the PAL responded with empathy to a learner’s affect, were categorized as responsive or nonresponsive. In the responsive condition, the PAL verbally responded or not according to experimental conditions. The emoticons reflected six affective states commonly occurring in learning situations, as suggested by Kort et al. (2001): Interest, Boredom, Confidence, Anxiety, Satisfaction, and Frustration. The Appendix presents the emoticons and sample screen excerpts.
responded with, ‘I am so glad you are interested.’ When a learner clicked a ‘Frustrated’ button, the PAL said, ‘Everybody can be frustrated once in a while, just hang in there.’ The affective responses were brief and did not affect the overall instruction time. In the nonresponsive condition, the PAL did not respond when a learner expressed affect; the module simply led the learner to the next phase. In both conditions, the factual information provided by the PAL was identical. Overall, the PAL demonstrated a positive mood by smiles and a pleasant tone of voice in both conditions.

**Gender**
As in Experiment 1, the same male or female PAL, both named Chris, was included according to experimental conditions.

**Dependent variables**
As in Experiment 1, the same four dependent variables – learners’ social judgements, interest, self-efficacy, and learning – were included. The same measures were used, except for learning, which included the recall test only.

**Procedures**
This experiment was implemented as a mandatory class activity during a regular session of an introductory educational-technology course. Other than that, overall procedures were consistent with Experiment 1. Participants were randomly assigned to one of the four conditions. The participants first logged onto the web-based module by entering demographic information. After answering pretest questions, they performed the task, taking as much time as they needed. This session took approximately an hour, with individual variations. Lastly, the participants answered the post-test questions.

**Design and analysis**
The study employed a $2 \times 2$ factorial design, in which variables included PAL empathetic responses (responsive vs. nonresponsive) and PAL gender (male vs. female). For social judgements, MANOVA was conducted. For interest and self-efficacy, MANCOVA was conducted, with pretest interest and self-efficacy as covariates. For learning, two-way ANOVA was conducted. The significant level was set at $\alpha < 0.05$.

**Results**
A preliminary analysis of the data did not indicate any violations of the assumptions for the parametric statistics used in the study. Examination of scatter plots supported the assumption of normality and revealed linear relationships for all tests. Box’s test of equality of covariance supported the equal covariance assumptions for multivariate analyses. Levine’s test for homogeneity of error variances supported the equal variance assumption for univariate analyses.

**Social judgements**
MANOVA revealed a significant main effect for PAL gender, Wilks’ Lambda = 0.85, $F_{3,46} = 3.08$, $P < 0.05$, partial $\eta^2 = 0.15$. The univariate results revealed a significant effect on ‘facilitating learning’, $F_{1,48} = 3.8$, $P < 0.05$. Students who worked with the male PAL ($M = 3.56, sd = 0.64$) rated the PAL as significantly more ‘facilitating [to] their learning’ than did students working with the female PAL ($M = 3.14, sd = 0.82$). The standardized effect size for this difference was Cohen’s $d = 0.57$, indicating a medium effect. Second, the results revealed a significant effect on ‘human-like’, $F_{1,48} = 6.95$, $P < 0.05$. Students who worked with the male PAL ($M = 3.59, sd = 0.52$) rated the PAL as significantly more ‘human-like’ than did students working with the female PAL ($M = 3.14, sd = 0.69$). The standardized effect size for this difference was Cohen’s $d = 0.74$, indicating a medium-large effect. Lastly, the results revealed a significant effect on ‘engaging’, $F_{1,48} = 4.11$, $P < 0.05$. Students who worked with the male PAL ($M = 3.79, sd = 0.52$) judged the PAL as significantly more ‘engaging’ than did students working with the female PAL ($M = 3.51, sd = 0.43$). The standardized effect size for this difference was Cohen’s $d = 0.59$, a medium effect.

**Interest**
MANCOVA revealed a significant main effect for PAL empathetic responses, Wilks’ Lambda = 0.53, $F_{3,20} = 3.54$, $P < 0.05$, partial $\eta^2 = 0.47$. Students who worked with the responsive PAL showed significantly higher
interest in the task and working with the PAL than did students with the nonresponsive PAL.

**Self-efficacy**

**MANCOVA** revealed a significant main effect for PAL empathetic responses, Wilks’ Lambda = 0.71, $F_{3,31} = 4.29$, $P < 0.01$, partial $\eta^2 = 0.29$. Students who worked with the responsive PAL showed significantly higher self-efficacy than did students who worked with the nonresponsive PAL.

**Learning**

There was no significant main or interaction effect for PAL empathetic responses and gender on students’ recall.

**Discussion**

The study examined the impact of the emotion and the gender of a learning companion on learners’ affective and cognitive gains as measured by their social judgements, interest, self-efficacy, and learning. Table 1 presents the summary of the results. Overall, the results showed that the learners’ affective and cognitive characteristics were influenced by the digital peer’s emotion and gender, as is the case in human peer–peer relations. This indicated that the learning companions played a social role, in that varying the emotion and gender produced varying degrees of learners’ social perceptions and their interest and self-efficacy beliefs in the task as well as their learning. More important, the study revealed that the instructional impacts of differing emotional capabilities (i.e. emotional expressions and empathetic response) of the learning companions were clearly differentiated, implying that each capability might serve varying instructional goals.

The effect of emotional expression

The learners’ social judgements about the instructional functionality of the learning companion varied according to its emotional states. Positive emotional expressions had a constructive impact on the learners’ judgements of the learning companion. Students who worked with the PAL that expressed positive emotions judged the PAL as significantly more facilitating to their learning and as more engaging than did students with the PAL expressing negative emotions. The results are consistent with classroom emotion research favouring teachers’ positive affect (Juvonen & Wentzel 1996; Lewis 2001). As in human-to-human interaction, the PAL’s negative emotions were not welcomed by the learners.

However, the PAL’s emotional states did not change the learners’ interest and self-efficacy nor their learning. A possible reason might be the lack of range in the PAL’s emotions. That is, the PAL expressed one constant type of emotion in each condition – all happy, all sad, or no emotion – throughout the module. Learners who were randomly assigned to one condition might have been less aware of the PAL’s emotions, perhaps not sufficiently aware to change their motivation and learning. Indeed, human emotion research indicates that individuals’ attention to their feelings mediates the effect of their feelings in general (Clore et al. 2001). This speculation also seems plausible when we consider that positive and neutral PALs were not judged differently but rather that both PALs were perceived favourably. Another explanation might be that the students were performing an optional task and that, in general, their motivation to learn the topic might be lower. In any case, the weak impact of agent emotional expressions on learning has been indicated in previous studies on pedagogical agents. Although emotionally expressive agents have been actively developed (Ball &

---

**Table 1. Summary of the results.**

<table>
<thead>
<tr>
<th></th>
<th>Emotional expressions (Experiment 1)</th>
<th>Empathetic responses (Experiment 2)</th>
</tr>
</thead>
</table>
| Social judgements         | • Positive or Neutral PAL > Negative PAL  
Positive male PAL > Positive female PAL | • Male PAL > Female PAL |
| Interest                  | • Positive male PAL > Positive female PAL |                           |
| Self-efficacy             | –                                    | • Responsive PAL > Nonresponsive PAL |
| Learning (recall)         | • Male PAL > Female PAL               | • Responsive PAL > Nonresponsive PAL |

PAL, pedagogical agent as a learning companion.
Breese 2000; Lester et al. 2000; Mori et al. 2003), the positive impact of expressive agents has been limited only to students’ perceptions. The current study confirmed this phenomenon: ‘Smiley faces’ may make a learner smile (by changing the learner’s social perceptions) but may not be sufficient to increase learning and/or motivation.

The effect of empathetic response

Students showed both higher interest and higher self-efficacy when the learning companion responded with empathy to their affective states. These results again reflect classroom settings, where students’ motivation and self-concept were increased when students understood that their teachers cared about them (Juvonen & Wentzel 1996). Likewise, when the learning companion showed that it cared about a learner’s affect by verbally responding with empathy, the learner’s interest and self-efficacy in the task were enhanced. This positive impact of empathetic responses implies that, for the agent to be effective, its emotion should be responsive to the learner’s when possible. Remember that in Experiment 1 the learning companions’ emotional expressions per se influenced neither the learners’ interest nor their self-efficacy. Rather than being simply a happy talking head, a PAL should respond to a learner’s affective states and flexibly adapt its feedback to a learner’s affective needs at the moment.

However, the presence of the learning companion’s empathetic responses did not influence learning. This finding confirms the current knowledge of affective pedagogical agents research, which lacks empirical support for increased learning (Towns et al. 1998; Ball & Breese 2000; Lester et al. 2000; Rizzo 2000). Hence, even the empathetic responses proven to be valuable should be used judiciously, and only when warranted to address instructional needs. Typically, in instructional settings, there are different goals and emphases, focusing on cognitive skill acquisitions or on affective gains such as behavioural or attitudinal changes. Given the results, PAL empathetic responses might be more effectively used to enhance a learner’s affect than knowledge and skill acquisition.

The effect of gender

The results of Experiment 1 indicated the positive impact of the male PAL on social judgements and interest. Students who worked with the male PAL perceived him more favourably and showed significantly higher interest in the task and in working with him than did students with the female PAL. Moreover, students who worked with the male PAL recalled more ideas than did students with the female PAL. Higher interest in working with the male PAL might better engage the learners in the task, later bringing higher recall. However, the students’ recall scores across the conditions were overall very low, ranging from 1.52 (highest) through 0.65 (lowest) out of the total of 15 points. So even with the statistical significance, the authors reserve judgement on the generalizability of the results.

Experiment 2 supported the greater impact of the male PAL on learners’ social judgements but not on their interest nor on their recall. This differential result in Experiment 1 and II might be attributed to the differing ratios of learner gender in the two experiments. Note that Experiment 1 included 40% males and 60% females and Experiment 2 included 20% males and 80% females. To increase the validity of the findings, the experiments were conducted in natural environments, i.e. in regular college classrooms. The authors’ inability to control the enrolments resulted in a gender imbalance that may have affected the study.

Still, the greater impact of male agents over female agents has been indicated in previous studies. For instance, both male and female college students showed higher motivation and more positive perceptions of agents after they had worked with a male agent than after they had worked with a female agent (Baylor & Kim 2005). This phenomenon indicates that gender-related social stereotypes in the real world (Carli 2001) might be consistently applied to an agent–learner relationship. Future research is invited to investigate the proactive role of a simulated peer (PAL) to reduce such stereotypes in computing environments.

The interaction effect between PAL affect and gender

Experiment 1 yielded the significant partial interaction between PAL emotional expression and gender. When the PAL expressed positive emotions, students perceived the male PAL more favourably than the female PAL. This interaction was minimal when the PAL expressed negative or neutral affect. Experiment 2 did not indicate the interaction between empathetic
response and gender of the PAL. This result seemed legitimate, given the results of Experiment 1 indicating the interaction only in the contrasts between positive versus negative and/or neutral affect conditions. In Experiment 2, the PAL consistently expressed positive emotions across the conditions, regardless of the presence or absence of empathetic responses.

Incidentally, the results from Experiment 1 showed a three-way interaction of PAL gender, emotions, and learner gender on the learners’ social judgements. As reported, when the PALs expressed positive emotions, both male and female students rated the male PAL as more facilitating to their learning, more engaging, and more intelligent. However, when the PALs expressed negative emotions, male students tended to rate the female PAL as more facilitating, engaging, and intelligent, whereas female students tended to rate the male PAL as more facilitating, engaging, and intelligent. When the PALs did not express emotions (neutral condition), those differences were minimal. Nonetheless, the authors feel less confident in arguing for this three-way interaction, considering the limited number of the participants in each cell. So this variation invites further inquiry.

Conclusion

Although research on emotional interactions of humans with computers is ongoing, few studies have reported the instructional impact of these interactions in learning environments. This study was intended to determine the instructional utility of the emotion and gender of virtual learning companions. Some of the previous studies seem to view pedagogical agents simply as an extension of multimedia – a combination of texts, images, and animation (Moreno et al. 2001; Atkinson 2002; Craig et al. 2002). In contrast, this study confirmed that pedagogical agents were perceived as social entities towards which the learners projected social conventions and stereotypes in human-to-human relationships. The learners in the study seemed even to expect their learning companion to have a personality or character, by making clearly differentiated social judgements according to its emotion and gender. Furthermore, this study highlights the distinct instructional impacts of PAL emotional expressions and empathetic responses. Given the cost and technological difficulties of creating even one facet of agent affect, integrating emotional expression, response, and recognition seems exceptionally challenging. Before taking on that challenge, we need to be convinced that expending the effort is justified by the educational impact.

Several limitations in the study should be addressed in further research. First, the study was a one-time implementation of limited duration. Second, when PAL gender was an important variable, the unbalanced ratios of student gender – especially in Experiment 2 – seem to suggest cautious interpretation. Most notably, the technologies were limited in fully featuring natural affective interactions. The technologies for PAL emotional capabilities, more generally agent technologies, are still in their infancy. In particular, very little is known empirically about their effectiveness in learning environments. Hence, this study should be considered an initial exploration to provide the research community with a preliminary sketch of the instructional impact of PAL emotion and gender. Subsequent research is invited to deploy the variables of the study more fully and to confirm these findings.

Acknowledgements

This study was sponsored by National Science Foundation Grant No. IIS-0218692. The authors thank all the students who participated in the study.

References


Appendix 1: screen excerpts

### Experiment 1

<table>
<thead>
<tr>
<th>Positive Male</th>
<th>Positive Female</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Positive Male Screen Excerpt" /></td>
<td><img src="image2" alt="Positive Female Screen Excerpt" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Negative Male</th>
<th>Negative Female</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image3" alt="Negative Male Screen Excerpt" /></td>
<td><img src="image4" alt="Negative Female Screen Excerpt" /></td>
</tr>
</tbody>
</table>

### Experiment 2

<table>
<thead>
<tr>
<th>Case Scenario</th>
<th>Stage 1: Blueprints</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image5" alt="Case Scenario Screen Excerpt" /></td>
<td><img src="image6" alt="Stage 1: Blueprints Screen Excerpt" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Emoticons</th>
<th>Affective Response</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image7" alt="Emoticons Screen Excerpt" /></td>
<td><img src="image8" alt="Affective Response Screen Excerpt" /></td>
</tr>
</tbody>
</table>