Pedagogical Agents as Learning Companions: The Impact of Agent Emotion and Gender

Yanghee Kim, *Utah State University*
A. L. Baylor
E. Shen

Available at: https://works.bepress.com/yanghee_kim/6/
Pedagogical Agents as Learning Companions: The Impact of Agent Affect and Gender

Abstract This study examined the impact of affect 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 PAL emotional expressions and empathetic responses. Experiment I focused on PAL emotional expressions (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 II investigated the impact of PAL empathetic responses (responsive vs. non-responsive) and gender with 56 pre-service teachers. Results yielded a main effect for emotional expressions on social judgments ($p < .05$) and main effects for empathetic responses on interest ($p < .05$) and on self-efficacy ($p < .05$). For PAL gender, Experiment I yielded main effects on interest ($p < .05$) and on recall ($p < .01$); Experiment II yielded a main effect on social judgments ($p < .05$). Also, Experiment I yielded an interaction trend of PAL emotional expressions and gender on social judgments. Findings imply that the presence of “happy smiles” might not be sufficient to influence learners; rather, that a desirable PAL should be responsive to a learner’s affective states and that gender-related social stereotypes in the real world seem to be applied to PAL/learner relationships.

Keywords: Pedagogical agents; Learning companions; Virtual peers, Human/computer interaction; Social interaction; Affective computing

Introduction

Pedagogical agents might provide an opportunity to simulate social interaction between human 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). 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 attempted to afford individualized cognitive guidance. Some of those systems reported success (Koedinger & Anderson, 1997), but their impact was—as is to be expected--far weaker.

than that of human tutoring. Hence, it seems natural to inquire about what has been missed in conventional computer-based tutoring systems, compared to human tutoring. The recent emphasis on social cognition seems to suggest 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 make pedagogical agents distinct from traditional computer-based tutoring, seemingly offering a unique instructional impact (Kim & Baylor, in press). 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). By integrating pedagogical agents, computer-based learning environments might emulate such interaction. Thus, the goal is to design a pedagogical agent as a learning companion (PAL) to serve as a computerized virtual peer (Ryokai et al., 2003) to simulate human-peer interaction.

The affective capabilities of a PAL might increase the believability and naturalness of the PAL (Bates, 1994) and so be an essential precedent for building social relations and facilitating social interaction with a learner. As human/computer interaction is generally similar to human-to-human interaction in the real world (Reeves & Nass, 1996), so a learner might be able to build virtual social relations – e.g., virtual friendship -- with a PAL (Bickmore et al., 2005; Gulz, 2005) when the PAL seems believable and natural. 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 affiliated with Sam, which might induce their behavioral changes. The implication is that a PAL should be perceived by a learner as a comrade so as to have meaningful instructional impact. To be perceived as friend-like, a PAL must be believable and natural (Dautenhahn et al., 2002; Bickmore & Cassell, 2004;). At the center of believability and naturalness might be a PAL’s ability to demonstrate affect (Bates, 1992; Dautenhahn, 1998; Ortony, 2002), because affect is an integral natural part of human social and intellectual functioning and greatly influences building relations among people.

**PAL Affect**

Affect influences an individual’s rational thinking, decision-making, social memory, judgments, and learning (Adolphs & Damasio, 2000; Forgas, 2000). More specifically, our momentary moods influence daily social interactions (Martin, 2000) and such information processing as attention, memory, and social judgments (Clore, 2001). Also, an important mechanism of emotional development from childhood through adolescence is socialization by peers (Asher et al., 1996). Simulating a human peer, a PAL may model various emotional expressions and reactions to guide a learner’s positive emotional experiences in the process of learning.

Also, an individual’s affect 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 PAL affect might play a role in influencing the affective and cognitive characteristics of a learner in computer-based environments.

The implementation of affect in computing environments has interested researchers in human/computer interaction. Many efforts have been made to develop the algorithms and systems architecture for the affect implementation (Ball & Breese, 2000; Hudlicka, 2003; Partala & Surakka, 2003), sometimes from the engineering perspective (Burleson et al., 2004; Picard, 2000). Few studies have rigorously investigated the instructional impact of agent affect. Those studies that indicated the positive impact of agent affect were implemented in gaming environments (Klein et al., 2002; Scheirer et al., 2002). The implications of those studies may differ in learning environments.

**PAL Gender**

Gender difference persists in academic interest and cognitive and interaction style and becomes more salient in educational computing (Cooper & Weaver, 2003), which is often attributed to social stereotypes associating computing more with males than with females (Clegg, 2001). Brody (1999) argues that social stereotypes are aligned with emotional functioning for the same or opposite gender distinctively and that as a consequence males

and females differ in their emotional expressions, empathic accuracy, and emotional behaviors.

The gender differences in the real world seem consistently projected to computing environments as well. In her 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 users perceived a female computer as more credible and conformed to the advice of the female computer on feminine topics such as cosmetic or fashion. 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 and colleagues (2002) reported that learners applied gender stereotypes to animated agents, and these stereotypic expectations affected their learning.

We have long known that the gender of peers contributes powerfully to differences in emotional experiences. It seems natural to assume that a learner’s reactions to a PAL might also be differentiated by PAL gender.

The purpose of this study was to examine the impact of PAL affect and gender on college students’ social judgments of PAL instructional functionality, their interest in the task and in working with the PAL, their self-efficacy beliefs in the task, and their learning in a computer-based environment. In general, PAL affect can be defined as a PAL’s ability to recognize a learner’s affective states, to express its own affect, and to respond to a learner’s affect (Picard, 1997; Hudlicka, 2003). Each capability requires unique technologies and resources for implementation. For instance, affect recognition is typically engineered by hardware technologies such as a pressure-sensitive mouse, a
BlueTooth wireless skin conductivity sensor, a Blue Eyes camera, a TekScan pressure sensor on a chair, and so on (Burleson et al., 2004). In contrast, affective expression and response are engineered by software technology using scripting languages and tools. Hence, investigating each affect separately might provide more accurate information on their individual effect in achieving a desired instructional outcome before investigating the integrated affect capabilities of a PAL. Therefore, this study examined the impact of PAL emotional expressions and empathetic responses separately in two controlled experiments. Detailed descriptions of Experiment I and II follow.

**Experiment I**

This experiment examined the impact of PAL emotional expressions (positive vs. negative vs. neutral) and gender (male vs. female) on learners’ affective and cognitive characteristics, as measured by the learners’ social judgments of PAL instructional functionality, their interest in the task and in working with the PAL, their self-efficacy beliefs in the task, and their learning.

**Participants**

Users’ perceptual reactions to pedagogical agents are varied (Gulz, 2004), especially among the users in their twenties (Bickmore & Picard, 2003). Thus, the population of this study was college students. A total of 142 college students in a computer-literacy course in a large public university located in the southeast United States voluntarily participated.

---

in the study, implemented as optional to other class activities. Approximately 40% of the participants were males and 60% were females. Sixty-seven percent of the participants identified themselves as Caucasian; 13% as African-American; 10% as Hispanic; 2% as Asian; and 8% as other. Convenience sampling was used to obtain participants new to the task domain of instructional design. The result of a pre-test question asking about their experience in the domain indicated that they were homogeneously novices. The average age of participants was 20.25 ($SD = 2.27$), with a range of 18 through 32.

**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: Intro, 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 PAL) 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, a PAL 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. Appendix A presents the screen excerpts of E-Learn.

**The design of the PALs**

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

**Independent Variables**

**PAL emotional expressions**

PAL 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. PAL expression had three affects: 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 affect condition, the PAL had a happy, smiling face and an engaging posture, with eye contact and head nodding. The background tone was red. The PAL verbally expressed its mood, such as “…this task looks fun” and “…completing this will be rewarding.” The participants perceived the positive PAL as significantly more “happy looking” than the negative PAL ($p < .001$). In the negative-affect condition, the PAL had a somber and rather frowning face and an aloof posture, with evasive eye gaze and less head nodding. The background tone was blue. The PAL expressed its mood verbally with such statements as “I don’t feel like doing this, but we have to anyway.” These affective comments were very brief and did not affect, in general, total instructional time across the conditions. The participants perceived the negative PAL as significantly more “sad looking” than the positive PAL ($p < .001$). In the neutral condition, the PAL did not express affect. 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 information provided by the PALs in E-Learn was identical across the three conditions. Appendix A presents the positive and negative variations of the PALs.

**PAL gender**

Either a male or female version of the PAL 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. The images of the male and female PALs are included in Appendix A. 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 judgments*

Social judgments referred to learners’ judgments of PAL instructional functionality and persona as measured by a questionnaire consisting of three sub-measures: “facilitating learning” (4 items), “engaging” (4 items), and “human-like” (4 items). Learners rated the PAL on a scale ranging from 1 (Strongly disagree) to 5 (Strongly agree). Item reliability of each category was evaluated as coefficient $\alpha = .91, .80, \text{and } .81$ 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 toward working with the PAL and toward 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 $\alpha = .87$. Interest was measured before and after the intervention. To control for test/retest effect, a split-half

technique was used, where the first 4 items out of 7 were implemented in the pre-test, and the last 5 items were implemented in the post-test.

*Self-efficacy*

In general, self-efficacy is defined as an individual’s beliefs in his or 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 vs. 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 $\alpha = .95$. 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 2 items out of 5 were implemented in the pre-test, and the final 3 items were implemented in the post-test.

*Learning*

Learning was measured by open-ended recall and application posttest 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 having Masters’ degree in Instructional Design counted independently the legitimate idea units from the information provided by the PAL in the module. Prior to scoring students’ answers, the raters agreed upon a total of 15 idea units presented by the PALs. Students’ recall was then scored in terms of the idea units. This procedure has been used in

previous studies (Mayer & Gallini, 1990; Plass et al., 1998; Kim et al., 2006). Inter-rater reliability was evaluated with Cohen’s Kappa = .94. In the application question, the participants were asked to write a brief e-learning plan according a new scenario. Students’ instructional plans were scored by the two instructional designers, given 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. Inter-rater reliability was evaluated as Cohen’s Kappa = .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 on to 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 minutes, with individual variations. Lastly, the participants answered the posttest 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 judgments, MANOVA (multivariate analysis of variance) was conducted. For interest and self-efficacy, MANCOVA (multivariate analysis of covariance) was conducted, with prior interest and self-efficacy as covariates. The significance level for all the analyses was set at $\alpha < .05$.

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 Judgments

MANOVA conducted as protected testing revealed the significant main effect for PAL emotional expressions, Wilks’ Lambda = .70, $F(6, 252) = 8.33, p < .001$, partial $\eta^2 = .16$. The univariate results revealed a significant effect on “facilitating learning,” $F(2, 128) = 3.75, p < .05$. 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 ($M = 2.12, SD = .93$). 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 ($M = 2.63, SD = .94$) rated the PAL as “facilitating to their learning” significantly more than did students with the negative PAL ($M = 2.12, SD = .93$). 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,” $F(2, 128) = 14.77, p < .001$. Students who worked with the positive PAL ($M = 3.06, SD = .97$) rated the PAL as significantly more “engaging” than did students with the negative PAL ($M = 2.25, SD = .83$). 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 ($M = 3.12, SD = .77$) rated the PAL as significantly more “engaging” than did students with the negative PAL ($M = 2.25, SD = .83$). 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 ($p = .07$, partial $\eta^2 = .05$). 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 submeasures: “facilitating learning” ($p < .05$), “engaging” ($p < .05$), and “human-like” ($p < .05$). 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 affect and gender ($p < .07$, partial $\eta^2 = .05$). A visual inspection of this relationship suggested a partial interaction between PAL affective expression and gender (Figure 1). 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 expressions), controlling for the inflation of family-wise error rates. The results of the tests revealed three significant interactions between PAL gender and affective expression. First, the positive and negative PAL emotions interacted with PAL gender, $F(1, 133) = 5.61, p < .05$. When the PAL expressed positive emotions, students judged the male PAL more favorably than the female PAL. This interaction, however, was not significant when the PAL expressed negative emotions. Second, PAL positive and neutral emotions interacted with PAL gender, $F(1, 133) = 6.00, p < .05$. When the PAL expressed positive emotions, students judged the male PAL more favorably. This interaction was not significant when the PAL did not express emotions (neutral). Third, PAL positive versus negative and neutral emotions interacted with PAL gender, $F(1, 133) = 7.73, p < .01$. When the PAL expressed positive emotions, students judged the male PAL more favorably than the female PAL. This interaction was minimal when the PAL expressed negative and neutral emotions. In summary, when the PAL expressed positive affect, PAL emotions and PAL gender interacted significantly to influence the learners’ social judgments, mirroring in part the interaction between gender and affect in the real world.

<Insert Figure 1 here>
Interest

First, MANCOVA revealed a significant main effect for PAL gender, Wilks’ Lambda = .893, $F(5, 127) = 3.05, p < .05$, partial $\eta^2 = .11$. Students who worked with the male PAL showed significantly higher interest in the task and in working with the PAL than did students working with the female PAL. There was no significant main effect for PAL emotional expressions on interest.

Second, MANCOVA indicated a significant interaction effect between emotional expressions and gender, Roy’s Largest Root = 1.0, $F(5, 128) = 2.51, p < .05$, partial $\eta^2 = .11$. The univariate analysis did not show significance in this interaction effect, however. A visual inspection of the data suggested the interactive relationship of the two: the learners’ interest in the male PAL conditions were differently patterned by the PAL emotion types. Figure 2 shows a graphical representation of the relationship. Simple Trend Analysis was conducted to test the statistical significance of this trend. The results indicated a significant linear relationship of PAL gender only in the positive emotion condition, $F(1, 136) = 5.21, p < .05$. When the PAL expressed positive emotions, students who worked with the male PAL showed higher interest than did those working with the female PAL. Again, this tendency is consistent with the results on social judgments.

<Insert Figure 2 here>
Self-Efficacy

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

Learning

MANOVA revealed a significant main effect for PAL gender on learning, Wilks’ Lambda = .83, $F(2, 65) = 6.66$, $p < .01$, partial $\eta^2 = .17$. The univariate results revealed a significant main effect for gender on recall, $F(1, 66) = 7.08$, $p < .01$. Students who worked with the male PAL ($M = 1.52$, $SD = 1.84$) achieved significantly higher recall scores than students working with the female PAL ($M = .65$, $SD = .89$). The standardized effect size for this difference was Cohen’s $d = 0.60$, a medium effect according to Cohen’s guidelines.

Experiment II

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

Participants

Participants were 56 pre-service teachers (11 males and 45 females) enrolled in a course in introductory educational technology. Sixty-eight percent of the participants identified

themselves as Caucasian; 12% as African-American; and 20% as other. The average age of participants was 20.71 ($SD = 2.92$) with a range of 18 through 34. The participants were homogeneous regarding their prior experience in the learning task of instructional planning. Convenient sampling was used with the consideration that, unlike Experiment I, the participants had to express their affect and thus needed to be more serious about 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. Their performances were reflected in the course grades.

**Instructional module**

In the web-based module, the learners’ task was to develop an instructional plan for sixth graders learning the economic concept of supply and demand. The module included five steps: Introduction, Case Study, Blueprints, Plan, and Assessment, in each step 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 and colleagues (2001): Interest, Boredom, Confidence, Anxiety, Satisfaction, and Frustration. Appendix B presents the screen excerpts of the module.

**Independent Variables**

*PAL empathetic responses*

Empathetic responses, i.e., whether or not the PAL responded with empathy to a learner’s affect, were categorized as responsive or non-responsive. In the responsive condition, the PAL verbally responded to a learner’s affect immediately after the learner clicked the emoticon expressing his/her affect. For instance, when a learner clicked an “Interested” button, the PAL 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 Non-responsive 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 amount of 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 I, the same male or female PAL, both named Chris, was included according to experimental conditions.

Dependent Variables

As in Experiment I, the same four dependent variables--learners’ social judgments, interest, self-efficacy, and learning--were included. The same measures were used, except for learning that 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 I. Participants were randomly assigned to one of the four conditions. The participants first logged on to 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 posttest questions.
Design and Analysis

The study employed a $2 \times 2$ factorial design, in which variables included PAL empathetic responses (Responsive vs. Non-responsive) and PAL gender (Male vs. Female). For social judgments, 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 $\alpha < .05$.

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 Judgments

MANOVA revealed a significant main effect for PAL gender, Wilks’ Lambda = .85, $F (3, 46) = 3.08, p < .05$, partial $\eta^2 = .15$. The univariate results revealed a significant effect on “facilitating learning,” $F (1, 48) = 3.8, p < .05$. Students who worked with the male PAL ($M = 3.56, SD = .64$) rated the PAL as significantly more “facilitating [to] their
learning” than did students working with the female PAL ($M = 3.14, SD = .82$). 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,” $F (1, 48) = 6.95$, $p < .05$. Students who worked with the male PAL ($M = 3.59, SD = .52$) rated the PAL as significantly more “human-like” than did students working with the female PAL ($M = 3.14, SD = .69$). 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,” $F (1, 48) = 4.11$, $p < .05$. Students who worked with the male PAL ($M = 3.79, SD = .52$) judged the PAL as significantly more “engaging” than did students working with the female PAL ($M = 3.51, SD = .43$). 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 = .53, $F (5, 20) = 3.54$, $p < .05$, partial $\eta^2 = .47$. Students who worked with the responsive PAL showed significantly higher interest in the task and working with the PAL than did students with the non-responsive PAL.

**Self-Efficacy**

MANCOVA revealed a significant main effect for PAL empathetic responses, Wilks’ Lambda = .71, $F (3, 31) = 4.29$, $p < .01$, partial $\eta^2 = .29$. Students who worked with the...
responsive PAL showed significantly higher self-efficacy than did students who worked with the non-responsive 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 PAL affect and gender on learners’ affective and cognitive gains measured with their social judgments, 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 PAL affect and gender, as is the case in human peer/peer relations. This indicated that the PALs played a social role, in that varying the emotion and gender produced the varying degrees of learners’ social perceptions and their motivation (interest and self-efficacy) as well as their learning. More important, the study revealed that the instructional impacts of differing PAL affective capabilities (i.e., emotional expressions and empathetic response) were clearly differentiated, implying that each capability might serve varying instructional goals.

<Insert Table 1 here>

---

The Effect of PAL emotional expressions

The learners’ social judgments about PAL instructional functionality varied with the PAL’s emotional states. The PAL’s expressions of positive emotions had a constructive impact on the learners’ judgments of the PAL. 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 favoring 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’ motivation (measured by interest and self-efficacy) or their learning. A possible reason can be the lack of variations 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 affect 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, 2001). This speculation also seems plausible when we consider that positive and neutral PALs were not judged differently but that, rather, both PALs were perceived favorably. Another explanation might be that the students were performing an optional task and that their motivation to learn the topic might be less-than-desired in general. At any rate, 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 & 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, but be not
sufficient to increase learning and motivation.

The Effect of PAL empathetic responses

Students showed both higher interest and higher self-efficacy when the PAL 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 a PAL
showed that he/she 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
PAL empathetic responses implies that, to be effective, a PAL’s affect should be tied to a
learner’s affect when possible. Remember that the PALs’ 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 PAL 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, the use of PAL empathetic responses to address

instructional needs should be judicious. Typically, in instructional settings, there are different goals and emphases, focusing on cognitive skill acquisitions or on affective gains such as behavioral or attitude changes. Given the results, PAL empathetic responses might be more appropriate for motivation than for knowledge and skill acquisition.

The Effect of PAL gender

The results of Experiment I indicated the positive impact of the male PAL on social judgments and interest. Students who worked with the male PAL perceived the PAL more favorably and showed significantly higher interest in the task and in working with the PAL than did students with the female PAL. Moreover, students who worked with the male PAL recalled more ideas provided by the PAL 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, a caution should be noted. The students’ recall scores across the conditions were overall very low, ranging from 1.52 (highest) through .65 (lowest) out of the total of 15 points. So even with the statistical significance, the generalizability of the results seems to be in question.

Experiment II supported the greater impact of the male PAL on learners’ social judgments, but neither on their interest nor on their recall. This differential result in Experiment I and II might be attributed to the differing ratios of learner gender in the two experiments. Note that Experiment I included 40% males and 60% females and Experiment II included 20% males and 80% females. The inclusion of learner gender in the analysis may provide another implication, but, in the current study, the number of
participants in each cell was too limited to validate the three-way procedures (i.e. PAL affect × PAL gender × learner gender). Subsequent research may pursue this issue, overcoming the limitation.

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

Conclusion

This study confirmed that the PALs were not just a type of multimedia -- a combination of texts, images, and animation. Rather, the PALs were perceived as social entities on which the learners projected social conventions and stereotypes in human-to-human relationships. The study highlights the differential instructional impacts of PAL emotional expressions and empathic responses. Given the cost and technological difficulties of creating even one facet of agent affect, integrating the affect of expression, response, and recognition will be exceptionally challenging. So we should be sure that expending the effort is justified by the educational impact.

There are several limitations in the study, however. 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 II -- seem to suggest cautious interpretation. Most notably, the technologies were limited in fully featuring natural affective interactions. The technologies for PAL affective 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 affect and gender. Subsequent research is invited to deploy the variables of the study more fully and to confirm or disconfirm these findings.

Acknowledgments

This study was sponsored by National Science Foundation Grant # IIS-0218692. The authors thank all the students who participated in the study.

References


Klein, J., Moon, Y., & Picard, R. W. (2002). This computer responds to user frustration; theory, design, and results. *Interacting with Computers, 14*, 119-140.


---


Fig. 1. Interaction of emotional expressions and gender on social judgments

Fig. 2. Interaction trend of emotional expressions and gender on interest
Table 1. Summary of the results

<table>
<thead>
<tr>
<th></th>
<th>Emotional expressions (Experiment I)</th>
<th>Empathetic responses (Experiment II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social judgments</td>
<td>• Positive or Neutral PAL &gt; Negative PAL</td>
<td>• Male PAL &gt; Female PAL</td>
</tr>
<tr>
<td></td>
<td>• Positive male PAL &gt; Positive female PAL</td>
<td></td>
</tr>
<tr>
<td>Interest</td>
<td>• Positive male PAL &gt; Positive female PAL</td>
<td>• Responsive PAL &gt; Non-responsive PAL</td>
</tr>
<tr>
<td></td>
<td>• Responsive PAL &gt; Non-responsive PAL</td>
<td></td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>--</td>
<td>• Responsive PAL &gt; Non-responsive PAL</td>
</tr>
<tr>
<td>Learning (Recall)</td>
<td>• Male PAL &gt; Female PAL</td>
<td>--</td>
</tr>
</tbody>
</table>

APPENDIX A:
Application Excerpts of Experiment I

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

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

<table>
<thead>
<tr>
<th>Neutral Male</th>
<th>Neutral Female</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image5" alt="Image" /></td>
<td><img src="image6" alt="Image" /></td>
</tr>
</tbody>
</table>
APPENDIX B:
Application Excerpts of Experiment II

<table>
<thead>
<tr>
<th>Case Study</th>
<th>Stage 1: Blueprints</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
</tr>
</tbody>
</table>

**Stage 2: Plan**

**Stage 3: Assessment**

<table>
<thead>
<tr>
<th>Emoticons</th>
<th>Affective Response</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
</tbody>
</table>