Martin Reisslein is an Associate Professor in the School of Electrical, Computer, and Energy Engineering at Arizona State University (ASU), Tempe. He received the Dipl.-Ing. (FH) degree from the Fachhochschule Dieburg, Germany, in 1994, and the M.S.E. degree from the University of Pennsylvania, Philadelphia, in 1996; both in electrical engineering. He received his Ph.D. in systems engineering from the University of Pennsylvania in 1998. During the academic year 1994-1995 he visited the University of Pennsylvania as a Fulbright scholar. From July 1998 through October 2000 he was a scientist with the German National Research Center for Information Technology (GMD FOKUS), Berlin and lecturer at the Technical University Berlin. From October 2000 through August 2005 he was an Assistant Professor at ASU. From January 2003 through February 2007, he was editor-in-chief of the IEEE Communications Surveys and Tutorials. He maintains an extensive library of video traces for network performance evaluation, including frame size traces of MPEG-4 and H.264 encoded video, at http://trace.eas.asu.edu. He is a member of the ASEE and a senior member of the ACM and the IEEE. Address: School of Electrical, Computer, and Energy Engineering, Goldwater Center, MC 5706, Arizona State University, Tempe, AZ 85287-5706; telephone: (+1)480-965-8593; fax (+1)480-965-8325; e-mail: reisslein@asu.edu.

Amy Marcelle Johnson, University of Memphis

Amy Johnson is an experimental psychology PhD student (cognitive track) working in the Institute for Intelligent Systems at the University of Memphis. Her research interests relate to Cognitive and Educational Psychology, including self-regulated learning, intelligent tutoring systems, cognitive load theory, and the cognitive processes underlying the integration of verbal and pictorial information in multimedia and hypermedia environments.
Effects of visual signaling on pre-college students’ engineering learning performance and attitudes: peer versus adult pedagogical agents versus arrow signaling

Introduction

Multiple-representation learning environments have the potential for improving learning in science, math, and engineering. Providing students with multiple representations of the same concept or procedure can facilitate their learning process. But these environments can also cause cognitive overload when novice students try to comprehend and integrate these representations. A challenge presented by these learning environments is that each representation needs to be understood and mentally integrated with the other representations, which may pose heavy cognitive demands on novice students\textsuperscript{1,2}. Despite this challenge, instruction rarely includes methods aimed at supporting the processes of selecting, organizing, and integrating multiple representations during learning. Signaling encourages learners to engage in productive cognitive processing during learning, which includes selecting relevant steps, organizing them into a coherent mental structure, and integrating them with already existing knowledge about the topic\textsuperscript{3}. Signaling that is used for focusing attention and for highlighting specific locations can be effectively used to guide the process of selecting relevant information\textsuperscript{4}. Several studies investigating the effects of signaling in multimedia learning environments reported positive learning outcomes for conditions that used signaling compared to conditions that did not employ any signaling\textsuperscript{5,6}. Signaling can be done in various forms such as voice, text, arrows, or animated pedagogical agents.

An animated pedagogical agent (APA) is a humanlike or otherwise animated on-screen character that provides learners with pedagogical assistance, such as directing attention, giving feedback, and presenting instruction, in computer-based learning environments\textsuperscript{5,7,8}. APAs have the intention of keeping students focused on important elements of the learning material, keeping them motivated, and providing learners with context-specific learning strategies\textsuperscript{9}. Also, APAs can be designed to simulate social interaction that may facilitate learners to engage in the learning task and to enhance learning in computer-based environments\textsuperscript{10}. Several studies have tested the persona hypothesis which posits that the visual presence of an APA in an interactive learning environment promotes students’ learning and positive perception of the learning experience\textsuperscript{11,12}. A few experimental studies found that the presence of an APA improves learning compared to no-agent conditions\textsuperscript{13,14,15,16,17}, and that an agent improves student attitudes\textsuperscript{18}.

Even though studies reported a positive effect of the presence of an APA on student learning and attitudes, the gender, age, and appearance of the APA remain open research questions. The APA’s gender, age, and appearance likely have an effect on student learning, as APAs become a model for the learners during the instructional process. Bandura\textsuperscript{19} proposed that observers acquire a cognitive representation of the model’s behavior that outlasts the modeling situation and enables learners to display observed behavior at later occasions. However, the learners do not always display the learned behavior; whether or not the behavior learned through observation will be displayed is influenced by motivational processes\textsuperscript{19,20}. Specific forms of model-observer similarity, such as similar image, gender, and age\textsuperscript{21}, or competence level of the
model\textsuperscript{22} may increase the probability of imitating the model’s behavior, and may have an effect on students learning and perceptions. Baylor\textsuperscript{23} noted that an effective social model is one that is similar to the observer, someone whom the observer aspires to be. Brumbaugh\textsuperscript{24} conducted a study on model's physical attractiveness and gender of the model, found that both physical appearance of the model and its interaction with gender of subject influenced the formation of personality inferences about the model. Similarly, Kim, Baylor, and Shen\textsuperscript{25} reported that learners’ perception and performance were influenced by the APA’s emotion and gender, as is the case in human peer–peer relations. Osman\textsuperscript{26} compared observation based-learning versus action-based learning and found that observational learning is as effective as action-based learning when the problem solving task is a non-specific goal, i.e., students learned about the entire system instead of a procedural-based goal, i.e., acquired specific knowledge for solution steps directly towards a goal.

The purpose of our two experiments that were conducted with a total of 339 middle school students was to investigate the comparative effects of signaling methods of visual information on students’ learning in multiple-representation learning environments. The study tested the hypothesis that animated pedagogical agents (APAs) can effectively support novice students’ learning when they serve as a way to signal visual information in these learning environments (through deictic gestures). Furthermore, the effects of a peer and an adult APA on students’ learning were explored in line with the social-cognitive research. Following the model-observer similarity premise, we hypothesized that peer-aged APAs would provide greater motivation toward the task, and increase learning outcomes.

Experiment 1

Method

Participants and design

The participants were a total of 190 7\textsuperscript{th} and 8\textsuperscript{th} grade precollege students in a middle school in the Southwestern U.S., 106 females and 84 males. The mean age of the participants was 13.0 years (SD = 0.62 years). One-hundred thirty five (71.1 \%) of the students reported that they were Caucasian, 18 (9.5 \%) students reported they had multiple ethnicities, 17 (8.9 \%) reported that they were Hispanic American, five (2.6 \%) reported being of other ethnicities, seven (3.7 \%) reported being African American, six (3.2 \%) reported being Asian American, and two (1.1 \%) reported their ethnicity as Native American. The students had completed the same school instruction in math and science, and had no school instruction on electrical circuits prior to participating in this study.

To determine the effect of different signaling methods, we manipulated the type of visual signaling students received in their program (APA signaling, arrow signaling, or no visual signaling). Dependent variables included performance on the posttest and student ratings of perceived difficulty and attitudes toward the instructional module. All participants were randomly assigned to the three conditions. There were 62 students in the APA signaling (P) condition, 61 in the arrow signaling (A) condition, and 67 in the no-signaling (C) condition. Comparisons were made among the groups on performance on posttest and program ratings.
Materials and apparatus

Each participant received computerized materials and paper-pencil materials. For each participant, the computerized materials consisted of an interactive program that included: (1) a demographic questionnaire; (2) an introduction to the objectives of the instructional program; (3) an instructional session providing a conceptual overview of a single-resistor electrical circuit; (4) a simulation session; and (5) a program rating questionnaire. The initial greeting and introduction to the objectives were presented by an animated pedagogical agent. All groups received narration throughout the modules.

The simulation session presented an electrical circuit with given default resistance and current values while the voice of a pedagogical agent explained how to obtain the voltage value by using Ohm’s Law equation, the circuit diagram, and the Cartesian graph of voltage as a function of current. Then, students were given three opportunities to select different current or voltage values and observe the outcome of their selection. For each of the selected current or voltage values, the pedagogical agent explained how to interpret the corresponding circuit diagram and Cartesian graph and how to obtain the missing voltage or current value using both Ohm’s Law equation and the Cartesian graph.

The instructional program had three different visual signaling versions, all of which contained an identical introduction including the presence of the pedagogical agent (step 2), identical narrated explanations and the same two visual representations of the circuit problems in the instructional session (step 3). The versions differed only during the simulation session (step 4): in the pedagogical agent-signaling (P) condition, an agent appeared on the screen to point to the key variables, symbols, and visual elements of the display as the explanation progressed; the arrow (A) condition used an arrow instead of the agent to direct students’ visual attention; and the no-signaling (C) condition did not include any visual signaling. In this experiment, the pedagogical agent was a young-male, approximately of the same age as the student participants, and was dressed casually, similarly to the students. His design was inspired by several similar avatars found in games that are popular among precollege students.

The last section in the computer program included a program rating questionnaire, which was a 12-item Likert instrument asking participants to rate their learning perceptions on a 5-point scale ranging from 0--strongly disagree to 4--strongly agree. The questionnaire was a revised version of a 16-item survey that the authors had developed in collaboration with experts in computer-based engineering education\(^\text{27,28}\). For the present study 12 items relating to perceptions towards the program and content matter and two items relating to the perceived cognitive load (a scale previously developed by Paas and Van Merrienboer\(^\text{29}\)) were retained. The construct validity of the revised survey was assessed with the judgment of subject matter experts in electrical engineering instruction.

To examine the reliability of the program rating instrument in the present study, we conducted a factor analysis using principal axis estimation, with all 12 items from the program rating instrument. Results demonstrated that two factors accounted for 64.5% of the variance for student ratings. Extraction of two factors was based on a threshold of one eigenvalue. The two factors identified related to 1) positive evaluations of the program or content matter (ten items,
such as “I would recommend this program to other students” and “I would like to learn more about electrical circuits”, with factor loadings ranging from .71 to .87) and 2) difficulty ratings (two items, "The lesson was difficult" and "Learning the material in the lesson required a lot of effort", with factor loadings .87 and .87). The internal reliability of the positive evaluation scale and difficulty rating scales was .92 and .72, respectively, as measured by Cronbach's alpha. A perceived difficulty score was computed by averaging the ratings produced in the cognitive-load questionnaire. A program ratings score was computed by averaging the ratings from the ten questions which loaded highly on this factor.

The paper and pencil materials consisted of a pretest and a posttest. The pretest was an 11-item multiple-choice test on students’ prior knowledge in the domain, and the posttest included 13 novel electrical circuit problems to be solved both with the symbolic approach using Ohm’s Law equation and graphically with the Cartesian graph. Both pretest and posttest were designed and printed using the same color and layout scheme as the computer program. Two independent scorers who were blind to the conditions of the participants scored the pretest and posttest (inter-rater reliability of .99).

The computer program used in the study was developed using Adobe Flash CS5 software, an authoring tool for creating web-based and standalone multimedia programs. The apparatus consisted of a set of laptop computer systems, each with a screen size of 1680 x 1050 pixels, and headphones.

Results

Table 1 shows means and standard deviations for the three treatment groups on the pretest and the posttest as well as their difficulty ratings. Initial analyses indicated no significant differences between groups on pretest scores, \( F(2, 187) = 1.28, MSE = 4.26, p = .33 \), or overall time-on-task (recorded by the computer system), \( F(2, 187) = .05, MSE = 5719.85, p = .96 \).

Analyses of variance (ANOVA) were conducted on students’ posttest scores, difficulty ratings, and overall program ratings using treatment condition as between-subject factor. The ANOVA on the posttest scores indicated a main effect for signaling condition, \( F(2, 187) = 3.80, MSE = 75.0, p = .02, \eta^2 = .04 \). Post-hoc tests demonstrated that the P group produced significantly higher posttest scores than the C group \( (p = .01) \), and that the A group produced marginally significantly higher posttest scores than the than the C group \( (p = .08) \). No other significant differences were observed.

The ANOVA on the perceived difficulty ratings of the students indicated a significant main effect for signaling condition, \( F(2, 187) = 4.60, MSE = 5.59, p = .01, \eta^2 = .05 \). Follow-up post-hoc tests demonstrated that the C condition students rated the program significantly more difficult than the P condition students \( (p < .01) \) and the A condition students \( (p < .05) \). There were no other significant differences for the difficulty ratings. Further ANOVAs indicated no other significant differences among the signaling conditions on students’ overall attitudes towards the program.
Experiment 2

Participants and design

The participants were a total of 149 7th and 8th grade students in a middle school in the Southwestern U.S., 85 females and 64 males. The mean age of the participants was 13.5 years ($SD = 0.52$ years). One-hundred one (67.8 %) of the students reported that they were Caucasian, 19 (12.8 %) students reported they were Hispanic American, nine (6.0 %) reported being of other ethnicities, eight (5.4 %) reported having multiple ethnicities, five (3.4 %) reported their ethnicity as Native American, four (2.7 %) reported being Asian American, and three (2.0 %) reported being African American. The students were from the same school and grade level as the students in Experiment 1; thus, they had completed the same school instruction in math and science, and had no school instruction on electrical circuits prior to participating in this study.

Experiment 2 had the identical experimental design and materials as Experiment 1. The only difference between the two experiments was that in Experiment 2, the APA was an older male, wearing an outfit that resembled a school teacher’s outfit. The same voice of the agent as in Experiment 1 was used for the older agent.

As in Experiment 1, participants were randomly assigned to the three conditions. There were 54 students in the APA signaling (P) condition, 50 students in the arrow signaling (A) condition, and 45 students in the no visual signaling (C) condition. Comparisons were made among the groups on performance on post-test, difficulty ratings, and program ratings. The difficulty ratings and program ratings were again composite scores averaged from the two cognitive load questions and 10 program ratings questions, respectively.

Results

Table 2 shows the means and standard deviations for each of the three treatment groups on the pretest, posttest, and difficulty ratings of experiment 2. An initial ANOVA on pretest scores showed no significant differences between groups, $F(2, 146) = 2.01, MSE = 1.22, p = .82$. An ANOVA on time-on-task (recorded by the computer system) indicated a significant effect for

<table>
<thead>
<tr>
<th>Signaling type</th>
<th>Pretest (max 11)</th>
<th>Posttest (max 13)</th>
<th>Difficulty ratings (max 4)</th>
<th>Program Rating (max 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedagogical agent (P) ($N = 62$)</td>
<td>$M = 5.08$</td>
<td>$SD = 1.98$</td>
<td>$M = 1.31$</td>
<td>$SD = 0.81$</td>
</tr>
<tr>
<td>Arrow (A) ($N = 61$)</td>
<td>$M = 4.98$</td>
<td>$SD = 2.16$</td>
<td>$M = 1.34$</td>
<td>$SD = 1.01$</td>
</tr>
<tr>
<td>No visual signaling (C) ($N = 67$)</td>
<td>$M = 4.60$</td>
<td>$SD = 1.69$</td>
<td>$M = 1.83$</td>
<td>$SD = 2.32$</td>
</tr>
<tr>
<td>Total ($N = 190$)</td>
<td>$M = 4.88$</td>
<td>$SD = 1.94$</td>
<td>$M = 1.50$</td>
<td>$SD = 2.28$</td>
</tr>
</tbody>
</table>
signaling condition, $F(2, 146) = 3.52, MSE = 4564.29, p = .03$. Follow-up post-hoc tests revealed that the A group spent significantly less time (1032.8 seconds) to complete the module compared to the P (1063.8 seconds) and C (1064.0 seconds) groups.

Analyses of variance (ANOVAs) were conducted on students’ posttest, difficulty ratings, and overall program ratings using treatment condition as between-subject factor. The ANOVA on the posttest scores showed no significant difference among conditions, $F(2, 146) = 0.5, \, MSE = 20.64, \, p = .95, \, \eta^2 = .00$. Similarly, the ANOVA on the perceived difficulty ratings of the students showed no significant difference among conditions, $F(2, 146) = 2.03, \, MSE = .99, \, p = .14, \, \eta^2 = .03$. Moreover, there were no significant differences on students’ overall attitudes towards the program.

<table>
<thead>
<tr>
<th>Signaling type</th>
<th>Pretest (max 11)</th>
<th>Posttest (max 13)</th>
<th>Difficulty ratings (max 4)</th>
<th>Program ratings (max 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedagogical agent (P)</td>
<td>$M$ 5.11</td>
<td>7.46</td>
<td>1.51</td>
<td>2.06</td>
</tr>
<tr>
<td>($N = 54$)</td>
<td>$SD$ 2.16</td>
<td>4.41</td>
<td>1.05</td>
<td>1.00</td>
</tr>
<tr>
<td>Arrow (A)</td>
<td>$M$ 5.32</td>
<td>7.23</td>
<td>1.80</td>
<td>1.97</td>
</tr>
<tr>
<td>($N = 50$)</td>
<td>$SD$ 2.39</td>
<td>4.99</td>
<td>1.01</td>
<td>0.82</td>
</tr>
<tr>
<td>No visual signaling (C)</td>
<td>$M$ 5.02</td>
<td>7.19</td>
<td>1.41</td>
<td>2.00</td>
</tr>
<tr>
<td>($N = 45$)</td>
<td>$SD$ 2.58</td>
<td>4.16</td>
<td>0.90</td>
<td>0.91</td>
</tr>
<tr>
<td>Total</td>
<td>$M$ 5.15</td>
<td>7.30</td>
<td>1.57</td>
<td>2.01</td>
</tr>
<tr>
<td>($N = 149$)</td>
<td>$SD$ 2.36</td>
<td>4.51</td>
<td>1.00</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Further analysis: Comparative analysis between peer-pedagogical agent and adult-pedagogical agent.

In this further analysis, comparisons were made between Experiments 1 and 2, but only for the agent signaling condition. The purpose of this further analysis was to investigate, in isolation, the differential efficacy of a young agent versus an old agent in teaching electrical circuits to pre-college students. The independent variable of this contrast was the age of the pedagogical agent (peer vs. adult), and dependent variables were the performance on the posttest, difficulty ratings, and student attitudes.

Participants were a total of 116 7th and 8th grade precollege students in a middle school in the Southwestern U.S. There were 62 students in the peer-pedagogical agent (PPA) condition (from Experiment 1), namely 35 females and 27 males. There were 54 students in the adult-pedagogical agent (APA) condition (from Experiment 2), namely 28 females and 26 males.

Results

Table 3 shows the means and standard deviations for each of the two treatment groups on the pretest, posttest, and difficulty ratings. Initial independent samples t-tests were conducted on pretest scores and time spent on the program. There were no significant differences on the pretest
between PPA and APA conditions, \( t(114) = .08, p = .94 \). A separate t-test conducted on the time spent on the program also showed no significant differences between the PPA and APA conditions, \( t(114) = .31, p = .76 \).

An independent samples t-test showed that there was a marginally significant difference between the PPA and APA groups on the posttest; \( t(114) = 1.84, p = .07 \). The PPA group scored marginally significantly higher on the than the APA group. There were no significant differences on the difficulty ratings between the two groups; \( t(114) = .98, p = .33 \). There were also no significant differences on students’ overall attitudes towards the program.

Further analysis was conducted on the amount of improvements within the PPA and APA conditions, and between the amounts of improvement from pretest to posttest. There was a substantial main effect for pretest to posttest gains in both PPA and APA groups, Wilks’ Lambda = .59, \( F(1, 112) = 78.14, p < .001, \eta^2 = .41 \). A mixed between-within subjects analysis of variance was conducted to assess if there were any significant differences between the PPA and APA groups, in terms of the amount of improvement they had from pretest to posttest. Analyses showed that there was a significant interaction between the groups and the amount of improvement from pretest to posttest, Wilks’ Lambda = .96, \( F(1, 112) = 4.5, p = .04; \eta^2 = .04 \). The PPA group improved significantly more than the APA group from pretest to posttest.

<table>
<thead>
<tr>
<th>Signaling type</th>
<th>Pretest (max 11)</th>
<th>Posttest (max 13)</th>
<th>Difficulty ratings (max 4)</th>
<th>Program ratings (max 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer agent (PPA) ( (N = 62) )</td>
<td>( M ) 5.08</td>
<td>( SD ) 1.98</td>
<td>1.31</td>
<td>2.36</td>
</tr>
<tr>
<td>Adult agent (APA) ( (N = 54) )</td>
<td>( M ) 5.11</td>
<td>( SD ) 2.16</td>
<td>1.12</td>
<td>0.81</td>
</tr>
<tr>
<td>Total ( (N = 116) )</td>
<td>( M ) 5.10</td>
<td>( SD ) 2.07</td>
<td>1.28</td>
<td>2.22</td>
</tr>
</tbody>
</table>

Discussion

The experiments examined the animated pedagogical agent (APA) signalling hypothesis, which states that APAs can effectively promote learning by supporting students’ selection of relevant visual information in multiple-representation environments. Additionally, the experiments investigated the effects of age (peer versus adult) of the APA on students’ learning and attitudes. The findings of this current research are in line with the animated pedagogical agent hypothesis; both experiments indicated that the agent (P) condition scored the highest compared to the arrow signalling (A) condition, and compared to the no signalling (C) condition on the posttest measure. This reveals that the deictic movements of a visual animated agent promote better learning than using no visual signalling. Moreover, deictic gestures of an animated agent is a more effective means of visual signalling than is using arrows to direct attention.
In Experiment 1, significant differences were observed for the signalling condition regardless of the type of signalling over the control group. This is consistent with the findings of previous studies that report positive learning outcomes for conditions that used signaling compared to conditions that did not employ any signaling\textsuperscript{5,6,13}. Additional support for the signaling hypothesis comes from students’ difficulty ratings, in that students rated the signaling conditions significantly less difficult than the no-signaling condition. Due to the added cognitive support from the arrow and agent students perceived their learning experiences significantly less difficult compared to no visual signaling.

In Experiment 2, the results had the same trends as in Experiment 1, in that the agent signalling group achieved the highest learning outcomes compared to the arrow signalling and control groups, but the effect of the agent was smaller such that it did not promote any significant differences in performance scores among the conditions. Also, there were no significant differences among conditions for the perceived difficulty ratings. Even though students didn’t perceive significantly less difficulty in their learning within the three signalling conditions, from the time data they seem to learn more and spend significantly less time when they learned from the arrow condition.

When only the two agent conditions were compared (PPA vs. APA), we gained clarification on the lack of a significant treatment effect in Experiment 2. According to our further analysis, the young agent had a greater effect on students’ learning compared to the old agent. This finding is in line with the findings of the previous studies stating that students are more likely to model and imitate people that are more similar to themselves\textsuperscript{21,23,24}. As the students were middle school students, they were more similar to the young agent than the older agent. The effect of the more salient peer agent was observed in the marginally significant differences on learning outcomes between the peer-agent and adult-agent groups. These findings suggest that learning about an engineering topic, which is perceived as difficult, can be promoted by using an animated agent similar to the target learner.

Practical applications and future research:

When teachers are providing new (or technical) information on math and engineering topics using multiple representations, signalling helps novice students to select, process and retain relevant information. Our results indicate that computer-based learning environments, using any type of visual signalling promote learning better than environment with only auditory guidance. Additionally, for middle-schoolers, the use of a peer-animated agent that is similar to the students may improve learning. More research is needed to determine whether the peer agent effect demonstrated for this population generalizes to older or younger students.

This initial comparative examination of the peer versus adult agents only considered male agents. Future research may investigate matching the agent to the learner on other characteristics, such as gender. Furthermore, it is yet to be determined whether the peer agent effect would extend to female animated pedagogical agents, regardless of the gender of the learner.
Bibliography


