Pedagogical Agents in Online Military Training:

The Effect of Voice, Image, and Animation on Social Agency and Learning

Jo E. MacDonald

University of West Florida

Pedagogical Agents in Online Military Training:

The Effect of Voice, Image, and Animation on Social Agency and Learning

The U. S. military is in the midst of a revolution—a revolution in training. Since the beginning of this century, a major cross-services initiative in military training has been the conversion of instructor-led classroom instruction to independent, asynchronous online learning. There are multiple reasons for this initiative: (a) to increase learners' access to course content, (b) to keep learners in the workforce while they are studying, and (c) to reduce travel costs. One of the shortcomings of products developed for this initiative has been the lack of a social context for learning. Learners study on their own without the support or cognitive input of an instructor or other learners. To help fill this void, many designers of instructional software have introduced pedagogical agents.

There is some disagreement as to what a pedagogical agent is. Some researchers prefer to limit the term to animated characters, human and otherwise (Shaw, Johnson, & Ganeshan, 1999), while others include still images or voice only (Atkinson, 2002). Despite these differences, researchers agree upon two points: a pedagogical agent possesses anthropomorphic qualities and serves an instructional function. For the purpose of this study, a pedagogical agent is defined as a media element that in some way suggests human presence for the purpose of learning. The agent may be as simple as audio that addresses the learner in a conversational tone using first and second person, or it may be as complex as an animated image that communicates nonverbally, poses questions, corrects and comments on learners' responses, and responds to learners' questions through the use of speech recognition technology and artificial intelligence.

There are four basic modalities for pedagogical agents that have been used for various military course conversion efforts: (a) voice only, (b) a still image used with text that suggests

spoken language, (c) a still image used with audio, and (d) an animated image used with audio. The problem is that this courseware has been developed without knowing which type of agent, if any, is most effective in the military training environment to simulate social interaction for learning.

The need for social interaction in learning has been well established. According to Vygotsky's theory of social cognitive development, proposed in the first part of the last century, instruction is most efficient when learners engage in activities within a supportive environment. Vygotsky further argued that learning takes place in a social context and interaction with others is an important part of the learning process (Roblyer, 2004). More recently, the Caine Learning Institute proposed twelve brain/mind learning principles, the second of which states that the brain/mind is social in nature and that learning is enhanced when the learner's social nature and need for relationship are engaged (Caine, Caine, McClintic, & Klimek, 2005). Because of the lack of human interaction in asynchronous online military courses, it is important to investigate the efficacy of different technologies to address this need. Pedagogical agents are one such technology.

In the last decade, researchers have investigated the use of pedagogical agents in several different roles, including instructor (Gilbert, 2002; Shaw et al., 1999), coach (Ting & Chong, 2003), mentor (Baylor & Kim, 2003; Moreno, Mayer, & Lester, 2000), tutor (Craig, Driscoll, & Gholson, 2004), and learning companion (Kim & Baylor, in press). However, for the agent to serve its intended educational purpose, it must be accepted by the learner (Baylor & Kim, 2003); and in order for an agent to be accepted, it must be believable. What makes it believable is its ability to reflect the persona, or defining characteristics, of its intended role. Baylor and Ryu (2003a, 2003b) list four key factors in agent persona: ability to engage the learner, possession of

human-like characteristics, credibility, and ability to facilitate learning. It follows that research is justified to determine the degree to which the different modalities for pedagogical agents evoke a persona that exhibits these critical factors.

Several studies have addressed the effects of audio, image, and/or animation on agent persona and learning. In a series of studies, Moreno and Mayer (n.d.) found that whereas voice promoted more meaningful learning, image or lack of image had no effect on learning. Baylor and Ryu (2003b), in their study contrasting image to no image and static image to animated image, found that animation significantly increased the persona effect but had no significant effect on learning. In 2002, Craig, Gholson, and Driscoll conducted an experiment using a 3 x 3 factorial design measuring the effect of agent properties (no agent, agent only, and agent with gesture) and picture features (static picture, sudden onset, and animation) on persona and learning. Whereas there was no significant persona effect, the researchers reported that sudden onset and animation had a significant positive effect on learning. Clearly, empirical investigations into this issue are inconclusive and warrant further investigation.

Although the proposed research will benefit other researchers by adding to the current body of knowledge about pedagogical agents, the intended audience for the proposed research is the instructional design community. Information about which, if any, agent modality is most effective for mitigating the lack of social interaction and positively affecting learning in asynchronous online military training will enable instructional designers to make informed decisions about the use of pedagogical agents in online instruction. It will also help those responsible for courseware acquisition to evaluate proposals involving the use of pedagogical agents. Perhaps the greatest benefit, however, will be for the learners themselves, if the findings suggest an application that will simulate social interaction to enhance learning.

Literature Review

Because pedagogical agents have the potential to engage learners and enhance learning, educational researchers have attempted to determine the relative effectiveness of different aspects of agent operationalization in multimedia, including voice, image, and animation—alone and in various combinations. This research has centered on the extent to which these features promote (a) perception of social interaction, (b) acceptance of agent persona, and (c) learning retention and transfer (see Appendix A for literature map). As can be seen in the following review of the literature, the results have not always been conclusive or similar.

Simulation of Social Interaction

One of the primary reasons for using a pedagogical agent in multimedia distance education is to simulate the social interaction that is missing in a solitary, self-paced learning environment. According to social agency theory, social cues in a multimedia message can influence learners to interpret their interaction with a computer as social. When this happens, "the rules of human-to-human communication come into play, so they try harder to make sense of what the computer is saying by engaging in deep cognitive processing" (Mayer, Sobko, & Mautone, 2003, p. 419). According to Reeves and Nass (1996), media elements do not even have to be realistic to elicit social responses.

To test the effects of social agency on learning, Moreno, Mayer, Spires, and Lester (2001) conducted a five-part study to examine whether learners make a stronger effort to understand content when self-paced computer-based instruction employs an animated pedagogical agent to simulate social interaction. The instructional program used in the study was a microworld called Design-A-Plant. Learners travel to an alien planet and must design a plant based on the type of leaves, stem, and roots that would best thrive in the environmental

conditions on the planet. Various aspects of the presentation were manipulated to determine the effect on three aspects of learning: (a) retention, (b) problem-solving transfer, and (c) expressed interest in the lesson.

Participants for four of the five experiments were recruited from the Psychology Subject Pool at the University of California, Santa Barbara. For the second experiment, Moreno et al. (2001) used 38 seventh-grade students from an urban middle school in the southeastern United States in order to see if the results from the first experiment would hold true for a different population. For all experiments, participants were randomly assigned to groups according to the design of the experiment. The sample sizes were adequate: 44, 48, 38, 64, and 79 respectively for the five experiments (Moreno et al., 2001). The higher numbers for the last two experiments reflected the need for a higher N with a factorial design.

Other than the variations in the computer-based presentation, Moreno et al. (2001) used the same materials and procedures for all five experiments. Participants were tested in groups of 1 to 4, each person in a separate cubicle. Headphones were provided for those assigned to a treatment with audio. A questionnaire was distributed to collect personal information on participants. After receiving oral instructions, the participants started the lesson. When all had finished, the experimenter administered a retention test consisting of three questions in which participants were asked to write down all the types of roots, stems, and leaves they could remember from the lesson. After 5 minutes, these tests were collected, and seven problem sheets were distributed, one at a time. Students had 3 minutes to complete each problem before the sheet was collected and the next one given out. Three of the problems were rated as *easy*, and four were rated as *hard*. Finally, participants were asked to complete a program-rating sheet that contained seven questions asking participants to rate their level of motivation, interest,

understanding, and the perceived difficulty of the material on 10-point scales. All instruments were scored by a scorer who was not aware of the treatment condition of each participant (Moreno et al., 2001).

The first two experiments tested the hypothesis that learning can be enhanced when learners interpret their relation with the computer as a social one involving reciprocal communication. In both experiments, one group of participants interacted with Herman the Bug, an animated pedagogical agent who spoke to them, while the other group received identical explanations as on-screen text with no agent or audio. The two experiments produced similar results. Although a two-tailed *t* test did not indicate a significant difference between the treatment and control groups in learning retention, students who learned with an animated pedagogical agent were significantly more interested in the program, $t(42) = 2.99$, $p < .01$ and $t(46) = 3.11$, $p < .005$ respectively for the first and second experiments. They also achieved significantly better transfer, particularly on the hard transfer items, than those who received text explanations, $t(42) = 3.38$, $p < .005$ and $t(46) = 3.55$, $p < .001$ respectively (Moreno et al., 2001). Although these experiments demonstrated that use of a pedagogical agent promotes learning by providing a sense of social agency, the researchers recognized that more research was needed to determine what aspect of the agent was most influential in producing this effect.

The third, fourth, and fifth experiments manipulated three aspects of the social agency environment to determine their effect on learning: (a) interaction, (b) image, and (c) voice. The third experiment tested the hypothesis that the increase in meaningful learning in an agent-based lesson is due to students' active participation in the learning environment. For this experiment, Moreno et al. (2001) compared the learning outcomes of students who were given the opportunity to design the plant structures with those of students who passively listened to and

viewed the instruction. Both groups received audio instructions with no visual image. The results from the two-tailed *t* tests showed that the students who participated in the design of the plant structure scored significantly higher on retention, $t(36) = 2.61$, $p = .01$, and on their ability to solve the difficult transfer problems, $t(36) = 2.10$, $p = .04$. Although these results supported the hypothesis that participation promotes deeper learning, the results may have been confounded by the effect of the contingent feedback received by the group in the participatory environment. Also, a two-tailed *t* test showed no significant difference in interest, which suggested that interactivity by itself is not sufficient to promote an emotional connection (Moreno et al., 2001).

The fourth experiment was a 2 x 2 between-subjects factorial design to determine the effect of voice and image on the three dependent variables. The two factors in the experiments were modality of the verbal information (narration or text) and image (image or no image). Four versions of the program were used: (a) animated image with narration, (b) animated image with text, (c) narration with no image, and (d) text with no image. A two-factor analysis of variance (ANOVA) failed to show an effect for image on any of the dependent variables. On the other hand, two-factor ANOVAs indicated a main effect for audio modality on retention, $F(1, 60) = 9.30$, $MSE = 17.30$, $p < .005$, and transfer, $F(1, 60) = 16.16$, $MSE = 1000.78$, $p < .0005$. Similarly, a two-way ANOVA for each of four program ratings—interest, understandability, perceived difficulty, and motivation—failed to reveal an image effect. However, the audio modality had a significant effect on interest $F(1, 60) = 9.51$, $MSE = 43.97$, $p < .005$ (Moreno et al., 2001).

The fifth experiment replicated the approach and conditions of the fourth with one difference: The image was a close-up video of an expressive actor. The findings from this experiment were similar to those of the previous one, "yielding further evidence for the

conclusion that a pedagogical agent’s image in a multimedia lesson does not hurt, nor does it provide any cognitive or motivational advantage for students’ learning" (Moreno et al., 2001, p. 207).

This exemplary study contributed considerably to knowledge regarding the effects of a social agency environment, as manifested by a pedagogical agent, on learning and student interest. The first two experiments yielded evidence that students prefer to learn within an agent-based environment and that they learn more deeply in a social agency environment than one in which information is presented as just text and graphics. The results of the third experiment indicate that students who actively participate in the lesson "are more actively involved in the processing of the materials of the lesson than students who learn identical materials in an environment based on a one-way transmission from computer to learner" (Moreno et al., 2001, p. 209). The fourth and fifth experiments yielded valuable information regarding the relative contribution of specific agent attributes to the social agency construct, demonstrating that whereas audio has a significant effect on learning, agent image "did not provide any cognitive or motivational advantage for students’ learning" (p. 209).

Although the study by Moreno et al. (2001) suggested that audio plays a primary role in establishing social agency using a pedagogical agent, additional research was needed to identify the attributes of audio that would be most beneficial. Mayer, Sobko, et al. (2003) conducted a study consisting of two experiments to further test the role of audio in promoting social agency and deep learning. They proposed that students try harder to understand narrated instruction in multimedia when they perceive social agency and that this sense-making effort results in deeper learning. They further proposed that students perform better on tests of problem-solving transfer and rate the speaker higher on socially desirable characteristics when narration is delivered using

a standard-accented human voice rather than either a foreign-accented human voice or a machine voice.

The first experiment sought to answer whether using an accented voice in multimedia narration affects retention, transfer, and the social rating of the speaker. The multimedia consisted of a narrated animation explaining the formation of lightning. In one version of the instruction, the narrator was a male native-English speaker, and in the other the narrator was a male speaker with a Russian accent. A supplemental study was conducted to ensure that learners could understand the words regardless of the speaker's voice. The hypothesis, which was not directly stated, was that based on social agency theory, students in the standard-accent group would score higher on tests of retention and transfer than students in the foreign-accent group and would rate the narrator more positively (Mayer, Sobko, et al., 2003).

To test the hypotheses, the researchers used a convenience sample of 68 college students from the Psychology Subject Pool at the University of California, Santa Barbara. Participants were randomly assigned to a treatment group. Although there was no specific mention of stratified random sampling, the reported percentages of women were similar for the two groups.

Data collection occurred in small groups. First, participants completed a questionnaire soliciting demographic information. After a brief explanation of the process, they took the multimedia instruction. One group received the standard-accent version, and the other received the version with the foreign accent. After they finished the instruction, they were given a retention test in which they were asked to write an explanation of how lightning formed based on what they could remember from the presentation. This test was followed by a sequential presentation of four transfer questions, which required participants to apply what they had learned to lightning-related problems. Scorers for the retention and transfer tests were not

informed of the participants' treatment condition. After the transfer test, participants were given a speaker-rating survey adapted from Zahn's and Hopper's (1985) Speech Evaluation Instrument. This instrument measures social characteristics of speakers in three areas: (a) superiority, (b) attractiveness, and (c) dynamism. The researchers selected five items from each area, each of which consisted of an adjective pair—for example, *literate-illiterate*. Participants used an 8-number scale to rate the narration on each attribute. Although the complete Speech Evaluation Instrument was not administered, the selected adjective pairs were appropriate, indicating that the research survey had face validity.

T tests performed on the data yielded mixed results. On the one hand, there was no significant difference in retention between the two groups. However, consistent with social agency theory, students who received the standard-accent version scored significantly better on problem-solving transfer than did those who received the foreign-accent version, $t(66) = 4.07$, $p < .001$, effect = 0.80. They also rated the narrator more positively, $t(66) = 3.27$, $p < .01$, effect = 0.71, indicating a social judgment based on accent (Mayer, Sobko, et al., 2003).

The second experiment sought to answer whether using a mechanized voice in multimedia narration affects retention, transfer, and the social rating of the speaker. For this experiment, another version of the instruction was created using a machine-synthesized male voice for the narration. This voice was likewise tested in the supplemental study mentioned earlier to ensure intelligibility. The hypothesis, which again was not directly stated, was that students in the human-voice group would score higher on tests of retention and transfer than students in the machine-voice group and would rate the narrator more positively (Mayer, Sobko, et al., 2003).

The researchers drew a convenience sample of 40 college students from the same subject pool, minus the students who participated in the first experiment, and randomly divided them into two groups. One group was given the human-voice version of the presentation (the same presentation used for the standard-accent in the first experiment), and the other was given the machine-voice version. The method and materials were the same except that the machine-voice version was used instead of the foreign-voice version and a two-item presentation survey was administered after the transfer test, on which participants used a 7-point scale to rate the difficulty of the presentation.

The results of multiple *t* tests showed significant support of the hypothesis. The human-voice group scored higher than the machine-voice group on both retention, $t(38) = 2.09$, $p < .05$, effect = 0.66, and transfer, $t(38) = 2.57$, $p < .02$, effect = 0.81. The human-voice group also rated the speaker significantly more favorably, $t(38) = 4.19$, $p < .001$, effect = 1.45. Finally, participants rated the human-voice version as easier to learn, $t(38) = 2.58$, $p < .02$, and easier to understand, $t(38) = 14.38$, $p < .001$ (Mayer, Sobko, et al., 2003).

Mayer, Sobko, et al. (2003) concluded that the results of the two experiments supported social agency theory. In both experiments, students who were presented a multimedia message in a human voice with a standard accent performed significantly better on problem-solving transfer than students who received the same message via a human voice with a foreign accent or a machine voice. The fact that the human voice with standard accent was also rated higher on social dimensions provided additional support for the theory.

The studies by Moreno et al. (2001) and Mayer, Sobko, et al. (2003) both provide empirical support for social agency theory. More importantly, with regard to the use of pedagogical agents to promote social agency, they also suggest that voice is a critical factor in

creating a sense of social presence that learners will respond to as they might to another human being. When this response is primed, "students try harder to understand and therefore learn more deeply from a multimedia message" (2003, p. 424).

Effect of Agent Voice, Image, and Animation on Agent Persona

Whereas Moreno et al. (2001) and Mayer, Sobko, et al. (2003) investigated the role of voice and image in promoting social agency, other researchers attempted to discover how different aspects and combinations of voice, image, and animation affect learners' perceptions of agent persona. *Persona* is the social front, the way of presenting oneself, that depicts one's role in a life situation. To be truly effective, a pedagogical agent must not only suggest a social presence but should also provide cues and evoke reactions appropriate to the role that it is intended to play in the instruction. Only then will learners accept the agent as a social entity.

Link, Kreuz, Grasser, and the Tutoring Research Group (TRG) at the University of Memphis (2001) performed a study to "evaluate the effects of manipulating linguistic expressions, intonation, and facial cues on individuals' perceptions of the evaluative feedback from a pedagogical agent" (p. 146). The study consisted of three experiments that varied verbal and nonverbal feedback features of an animated pedagogical agent to determine the effect on students' perception of feedback as positive, negative, or neutral.

The pedagogical agent in this study was *AutoTutor*, an intelligent tutoring system (ITS) created by TRG. AutoTutor is a *talking head*, a human-like image using facial expressions and gaze while delivering dialog as synthesized speech using inflection and intonation. AutoTutor uses latent semantic analysis (LSA) to analyze a student's textual inputs to an instructional interaction and subsequently "provide a pedagogically appropriate dialog move, such as a prompt, hint, or elaboration" (Link et al., 2001, p. 146).

There were three overall hypotheses for the study: (a) verbal and nonverbal feedback typically associated with positive emotions is associated with a perception of positive feedback; (b) verbal and nonverbal feedback typically associated with negative emotions is associated with a perception of negative feedback; and (c) "according to Massaro's (1998) theory of perception, neutral feedback terms paired with features associated with either positive or negative emotions will be associated with a perception of relatively positive or negative feedback, respectively" (Link et al., 2001, p. 146).

The series of experiments systematically manipulated feedback variables to determine their relative contribution to students' perception of feedback as positive, negative, or neutral. For each experiment participants viewed randomized pairings of artificially created student inputs and feedback illustrating one of the variables being studied. The "student" inputs for the experiments were sixteen advanced medical facts, which were chosen so that participants in the experiments would not be able to anticipate either positive or negative feedback. For each experiment, the researchers used a convenience sample of 30 undergraduate introductory psychology students from the University of Memphis, who participated for course credit. Different participants were used for each sample.

The first experiment concentrated on verbal features of feedback that previous research had shown to influence affective perceptions: (a) linguistic expression, (b) rate, (c) pitch, (d) pause, and (e) emphasis. The experiment did not consider pitch change or articulation, since Microsoft's text-to-speech engine cannot manipulate these variables. There were three expressions in each of the three feedback categories—one positive, one negative, and one neutral. These were randomly associated with the four other verbal features, resulting in 144 input-feedback combinations. The 30 participants viewed each combination and used a keyboard

to rate the feedback on a 6-point Likert scale ranging from 1 (*very negative*) to 6 (*very positive*). The researchers then conducted a multiple regression analysis using the five variables as multiple predictors and the participants' ratings as the criterion variable. Of the five variables, the only significant predictor was linguistic expression ($r^2 = .32$), indicating that "the participants relied entirely on the linguistic expression when they made their feedback judgments" (Link et al., 2001, p. 148).

The second experiment concentrated on the facial features that influence affective perceptions: (a) mouth curve, (b) eyebrow height, (c) eyebrow curve, and (d) eye size. Facial features such as forehead wrinkling, mouth openness, and nose wrinkling, which have also been shown to be important for the expression of emotion, were not manipulated because they express disgust or fear and were therefore considered "not relevant to tutorial feedback" (Link et al., 2001, p. 151). Three variations for each of the four variables resulted in 81 possible facial expressions, which were randomly associated with the 16 student inputs and displayed in random order. Although the same 6-point scale was used to rate the feedback, participants marked their ratings on a printed form instead of using a keyboard for electronic entry. Again the researchers conducted a multiple regression analysis, this time using the four facial features as the predictor variables and the subjects' ratings as the criterion variable. The only significant predictor was the mouth curve ($r^2 = .47$), indicating that "when the facial expression was the only form of feedback, the participants relied almost exclusively on the shape of the mouth in making their ratings" (p. 149).

The third experiment manipulated combinations of verbal and facial features to determine whether one type of features was more influential than the other or whether the effect was additive and students used both types of information to interpret feedback. Four features were

manipulated: (a) linguistic expression, (b) mouth curve, (c) rate, and (d) pitch. All other parameters were varied randomly. Combinations of the four variables were randomly associated with the 16 student inputs to produce 96 possible combinations, which were presented in random order to the 30 participants. Because of Microsoft Agent's technical limitations, AutoTutor displayed the nonverbal feedback first and then provided the verbal feedback. "Participants rated the feedback holistically (i.e., considering both what was said and how it was said) on the six-point scale used in Experiment 2" (Link et al., 2001, p. 150). The researchers conducted a multiple regression analysis using the four variables as the predictors and the participants' ratings as the criterion variable. Once again, linguistic expression category and mouth curve were the only significant predictors. The r^2 values were calculated as change scores, "starting with the most robust predictor and then adding on the statistical contributions of each incremental predictor" (p. 150).

Link et al. (2001) concluded that when learners receive feedback from an animated pedagogical agent, they integrate information from both modalities—in particular, the linguistic expression and the mouth curve—to form their perception of its affective content. They also noted how these features affect each other. "For example, a neutral feedback term combined with an upturned mouth results in a positive rating. In addition, when the modalities conflict (e.g., a positive term combined with a downturned mouth), participants tended to provide neutral feedback ratings" (p. 151). Although it is possible that the participants' personal reactions to the feedback may have confounded the results, these experiments still suggest that even a simple, two-dimensional talking head, such as the one used in AutoTutor, has sufficient expressive range to establish persona and elicit an affective response.

Baylor and Ryu (2003b) furthered the research on persona with an exemplary study to investigate how image and animation affect agent persona. Citing a series of studies, they noted that whereas research has consistently demonstrated that voice is an important aspect of agent persona, "the role of agent image and animation is not clear" (p. 376). They predicted that "the presence of an agent image and particularly animation will lead the learner to perceive the agent as more person-like, engaging, credible, and instructor-like" (p. 377). The experiment was also designed to test the effect of image and animation on learner performance and perceived value of the agent.

For the study, the researchers used a convenience sample of 75 preservice teachers at a public university in the Southeast. The students were required to participate in the study as part of a course on instructional planning. Participants had little prior experience with instructional planning, as indicated by a mean score for a question on a demographic survey related to prior experience. For the first part of the course, all participants received the same instruction. Then they were randomly assigned to one of three treatment groups. By performing chi-square analyses on the demographic data, researchers determined that there were no significant differences in the groups regarding age, GPA, ethnicity, gender, and year in school (Baylor & Ryu, 2003b).

The experiment used MIMIC (Multiple Intelligent Mentors Instructing Collaboratively), a computer-based learning environment. Within MIMIC, participants were given an introduction and a case study to provide the context for the activity. They were then required to design an instructional plan to help students learn the concepts of supply and demand. Throughout the activity, they could ask questions from a mentor, Chris, who was operationalized in three different ways. For the *animated-image* group," Chris was manifested by the Microsoft Agent

character Merlin the Wizard. The built-in agent functionality enabled this version of Chris to use gestures to provide nonverbal communication and direct attention to another part of the screen. For the *static-image* group, only a static image of Merlin was used. For the *no-image* group, Chris was represented by a box with the words "Ask Chris." Learners interacted with Chris by clicking the image or the box, depending on the version assigned to their group. In addition to advising learners when requested, Chris provided an initial observation at the beginning of each planning phase and asked reflection questions to encourage self-evaluation. Three aspects of Chris were the same across all three groups: (a) the advice he gave, (b) his voice, and (c) a dialog bubble that presented redundant text for Chris's speech (Baylor & Ryu, 2003b).

After completion of the activity, participants completed a computer-based, research-designed questionnaire. A 1–5 Likert scale format was used to rate Chris on the extent to which he was engaging, person-like, and credible. Researchers established inter-item reliability by determining Chronbach's alpha for the items used to test each of the three characteristics. Participants were also asked three open-ended questions addressing the extent to which Chris was instructor-like and his value to the learning experience. Answers were coded according to preset criteria (Baylor & Ryu, 2003b).

To evaluate learner performance, the participants' instructional plans were scored according to a rubric with four sub-areas. To ensure inter-rater reliability, "two of the researchers met and together discussed what characterized a score of 1 through 5 (where 1 = *poor* and 5 = *excellent*) for each of the four sub-areas for five sample instructional plans" (Baylor & Ryu, 2003b, p. 383). Then, after each researcher scored an additional 15 instructional plans independently, a Pearson r was calculated, establishing a high degree of correlation between the raters' scores, $r > .90$. The scorers were not informed of the participants' treatment condition.

With inter-rater reliability established, one of the researchers then scored the remainder of the instructional plans using the same rubric (Baylor & Ryu, 2003b).

Data resulting from all the tests were analyzed by conducting a one-way analysis of variance (ANOVA) using an alpha of .05, "with the three agent conditions as treatment levels with the primary analysis consisting of two planned orthogonal contrast comparisons" (Baylor & Ryu, 2003b, p. 384). The first contrast was between image (both static and animated) and no image, and the second was between static image and animated image. The researchers performed a post-hoc pairwise comparison between the animated-image and no-image conditions if there was no significant difference for the first contrast and a significant difference for the second contrast.

The first contrast indicated no significant difference between the image and no-image conditions for *engaging*, *person-like,* or *instructor-like*. However the second contrast for these three characteristics showed a significant difference in favor of the animated image, with a medium effect for *engaging* and *person-like* and a large effect for *instructor-like.* This trend was reversed for *credibility*. Both image conditions were found to have more credibility than the no-image condition; however, there was no significant difference between the two image conditions. Post-hoc comparisons of the animated condition with the no-agent condition showed that the animated agent was significantly more engaging and more instructor-like than no agent but that there was no significant difference for the *person-like* attribute. In terms of agent value, there was a significant difference in only two dimensions. Participants in both of the agent-present conditions tended to state that they liked the information provided more frequently than did the participants in the no-agent condition. Also, participants in the no-agent condition were more

likely to say that they disliked nothing about the agent. There was no statistical difference among the various conditions for performance (Baylor & Ryu, 2003b).

Baylor and Ryu (2003b) concluded that overall "the presence of animation is either the most desirable or one of the most desirable instantiations of image and animation for agent persona perceptions in educational computer-based environments" (p. 388). However, if cost is a factor, then developers should base their decision on the criticality of the individual factors. For instance, if credibility is the most important issue, a static image will provide as good results as an animated one. On the other hand, because of the large effect sizes, animation is necessary for a more engaging and instructor-like agent. The researchers cautioned that because the study focused on perceptions of pedagogical agents, the results "may not be generalizable beyond educational agent-based environments" (p. 388).

Taken together, the studies by Link et al. (2001) and Baylor and Ryu (2003b) suggest that an animated image enhances the perception of persona. They also support the idea that an agent persona based on audio with no image is as likely to be interpreted as person-like as is an animated agent.

Effect of Agent Voice, Image, and Animation on Learning

If a pedagogical agent has an acceptable persona, it stands to reason that the resulting sense of social agency would increase learning, based on the findings of the studies on social agency by Moreno et al. (2001) and Mayer, Sobko, et al. (2003). However, the study by Baylor and Ryu (2003b) failed to establish a significant effect on learning. The following studies examined the effect of agent voice, image, and animation on learning.

Citing the small number of empirical studies of animated agents and the diversity of results, Moundridou and Virvou (2002) conducted a study of an educational system using an

animated pedagogical agent to evaluate the persona effect on learning. The system tested was a web-based authoring tool for the construction of intelligent tutoring systems (ITSs) in algebra-related domains, called WEAR. WEAR enables adaptive learning by associating each student with a level of knowledge based on past performance and then suggesting problems or study topics corresponding to that level. The agent in WEAR is an animated talking head that is used to communicate instructions to the learner and provide feedback. The audio for the experiment was created using a commercially available text-to-speech voice synthesizer.

The purpose of the experiment was to examine how the persona of an animated pedagogical agent affected students in three areas: (a) behavior while interacting with the system, (b) reactions to the experience, and (c) performance. Learners were divided into two groups, one with the agent present and one with the agent absent, in order to test the null hypothesis that there would be no difference between the means of the two groups in each of the three areas.

The sample, composed of 48 college students studying either informatics (applied information science) or economics at the University of Piraeus in Greece, was randomly divided into the two groups. It is assumed that the sample was a convenience sample, since no mention is made of random selection other than in the grouping of the participants. One group was given a version of the ITS in which a talking head was used to interact with learners. The other group was given a version with identical content and messages, the only difference being that the messages were delivered through onscreen text instead of the talking head.

Participants were first given a paper-and-pencil pretest consisting of five problems, for which each individual's score and completion time were recorded. After the pretest, participants worked with the ITS to solve word problems with the same level of difficulty. During this time, the ITS logged student data on (a) the amount of time spent on each problem, (b) response times,

(c) total time on the system, (d) the number of times the learner asked for the solution before solving a problem, (e) the number of times the learner dropped a problem before coming up with the solution, (f) the number of mistakes per problem, and (g) the number of problems solved correctly. Participants were allowed to spend as long as they desired and to solve as many problems as they desired. Afterwards, they were given a paper-and-pencil posttest with five similar problems, and each individual's score and completion time were recorded. Finally, students were asked to complete a seven-question research-designed questionnaire to record their affective experience with the program. The questions, which were listed in the article, included five questions scored on a 5-point Likert scale and two open-ended questions. Although the questions seemed to possess face validity, there was no discussion of the method of development or attempts to establish the reliability and validity of the instrument.

To test the null hypothesis, the data were subjected to a series of two-tailed *t* tests, using an alpha of .05. For student behavior, the researchers tested attentiveness, as indicated by total time with the system, provided that the time spent was not excessive—that is, it did not exceed a threshold value of 30 minutes. This threshold was established through a short empirical study involving 15 experts, most of whom were classroom instructors. Given a list of criteria from previous studies regarding attentiveness to computer-based instruction, the experts selected the time criterion that they would use to measure attentiveness. The difference between the mean times for the two groups was not statistically significant, $t(46) = 0.97$, $p(0.336) > .05$, indicating that the presence of the agent had no effect on students' attention.

Performing a two-tailed *t* test on the mean difference for each of the five affective measures resulted in a significant difference for four of the five, including (a) enjoyment, $t(46) = 6.74$, $p = 0.000$; (b) perceived ease of use $t(46) = 5.02$, $p = 0.000$; perceived difficulty of

problems, $t(46) = 6.18$, $p = 0.000$; and the system's effectiveness in improving their problem-solving skills, $t(46) = 6.74$, $p = 0.027$. In fact, these findings were significant even at the most stringent alpha level. However, a *t* test failed to reject the null hypothesis for usefulness of the system in comparison to a human tutor $t(46) = 0.66$, $p = 0.515$. The researchers reported being surprised by this finding because they felt that "the presence of the speech-driven anthropomorphic agent could give students the impression of a real tutor" (Moundridou & Virvou, 2002, p. 8). Although the researchers did not devise a way to statistically measure the answers given to the open-ended questions, they noted that many of the comments centered on the agent's voice. "This is probably an indication that the most important aspect of the agent is the speech rather than the animation" (p. 8).

To test the effect of the agent on learning outcomes, Moundridou and Virvou (2002) compared data from the pretests, the posttests, and interactions with the system. There was no significant difference between the groups in pretest scores and completion times, indicating a level baseline for comparison. Both groups showed improvement in their scores and completion times, indicating that the ITS improved performance in either condition. However, although the group with the agent showed greater improvement in posttest scores and completion times than the control group, the difference between the groups was not statistically significant: time improvement, $t(46) = -0.86$, $p = 0.394$; grade improvement, $t(46) = 1.22$, $p = 0.227$. Also, although the group with the agent had a higher mean grade for solving problems with the system than the control group, the difference again lacked statistical significance, $t(46) = 1.69$, $p = 0.098$. The researchers concluded that although the results favored the agent version of the system, there was no support for the claim that either the presence or absence of an animated agent affects learning outcomes.

The researchers suggested that the main advantages of using an animated agent center on learner motivation. Students enjoyed interacting with the agent, perceived that problems were less difficult, and felt that the agent helped them to improve their learning skills. Although the study did not indicate a significant agent effect on learning outcomes, Moundridou and Virvou (2002) hypothesized that increased motivation "may show positive results in learning outcomes in the long run" and suggested that this would be their focus in future research (p. 10).

Whereas Moundridou's and Virvou's (2002) study contrasted the effect of an animated image with audio to that of text only, Atkinson's (2002) study added a third condition, audio only, which separated the effect of audio from that of image. The purpose of the study was "to examine empirically how the efficacy of a computer-based learning environment could be improved through the use of an animated agent" (p. 417). Specifically it focused on whether explanation provided by an animated pedagogical agent would be more effective than text-only or audio-only explanations of worked examples in helping participants learn how to solve word problems in mathematics. The study consisted of three phases: a pilot and two experiments. The animated agent used in all three phases was Microsoft Agent's Peedy the Parrot. The pilot used audio generated by a text-to-speech engine. The two experiments used audio recordings of a human tutor. The research questions varied somewhat from one phase to the next. However, they all centered on the effectiveness of voice and an animated image at promoting learning when compared to each other and to text alone.

Each phase used a convenience sample consisting of volunteer undergraduates at Mississippi State University. The researcher then used random sampling to identify participants for each group. Participants filled out a demographic questionnaire and then studied an eight-page review on how to identify and solve proportional relationships in word problems. The

purpose of this was to control for prior learning. In the pilot and the first experiment, the review was followed by a pretest to determine a score prior to treatment. The pretest was dropped for the third experiment in order to condense the procedure into a single session, thereby increasing participation; however, doing so prevented calculation of a difference score, which might have helped to clarify the effect of each treatment on performance (Atkinson, 2002).

The instructional component was a multimedia program that consisted of four stepped worked examples and four practice problems. Different versions of the program were developed for each of the treatments tested. In each version, participants were asked after each example to use a 5-point Likert scale to rate the relative difficulty of the problem. The instructional component was followed by a posttest with an equal number of near-transfer and far-transfer problems. The near-transfer problems were structured similarly to the practice ones. The far-transfer problems differed in structure and gradually increased in difficulty. For each new phase Atkinson increased the number of problems by two questions, one near-transfer and one far-transfer, "to bolster the reliability of the posttest" (Atkinson, 2002, p. 424). Participants in the two experiments were asked to fill out a brief research-designed affective questionnaire after the posttest, in which they rated the effectiveness of the instructional program on a 5-point Likert-type scale. As with the posttest, the number of questions on the questionnaire was increased to improve reliability.

The three conditions used for the pilot were *voice plus agent*, *voice only*, and *text only* (the control). Atkinson used analyses of covariance (ANCOVAs), with the pretest score as the covariate, to analyze the effect of each condition on the learning-process measures (example difficulty and practice problem performance) and on the learning-outcome measures (near transfer and far transfer). The results failed to establish an effect for the animated agent; in fact,

participants in the text-only condition scored significantly higher on the practice problems. These results were at odds with those from a similar study by Moreno, Mayer, and Lester in 2000, in which participants in an agent-based environment significantly outperformed their peers in a text-based environment. Atkinson attributed this discrepancy to the fact that he had used a text-to-speech engine rather than a human voice, as had been used in study by Moreno et al. He revised the materials to incorporate audio recordings of a human tutor for the first experiment.

The first experiment used a 2 (modality of explanation) x 2 (presence or absence of an animated agent) factorial design, with a control condition that had no instructional explanations or agent. Dunnett's multiple comparison procedure was used to compare each of the treatment conditions with the control condition. Then a 2 x 2 (agent x voice) ANCOVA was performed, using the pretest as a covariate. If a condition had a significant main effect, the sequential Hayter's procedure was used to examine pairwise comparisons among the various treatment conditions. Testing each measure for homogeneity of regression yielded no significant results. This time the results were more in keeping with the modality effect:

> Participants in the voice-plus-agent condition reported lower levels of perceived difficulty with regard to the examples with which they were presented than did their counterparts in the control condition. Moreover, unlike their voice-only counterparts, the voice-plus-agent participants also outperformed their control peers on both near and far transfer. Thus, although the measurable effects were not as dramatic or as pervasive as might be expected, the findings provide a modicum of support for the conclusion that animated pedagogical agents are effective at promoting learning from examples. (Atkinson, 2002, p. 423)

Noting that the less-than-anticipated effects may have resulted from the small sample sizes, Atkinson used a non-factorial design for the second experiment, which enabled him to decrease the number of groups from five to three—(a) voice plus animated agent, (b) voice only, and (c) text only—and increase the number of participants per group. In this experiment, participants in the voice-only condition outscored their counterparts in the text-only condition in all learning-process and learning-outcome measures. However, the participants in the voice-plus-agent group had the largest gains, outperforming their counterparts in both the text-only and voice-only groups. "Cohen's f statistic for these data yields effect size estimates of 0.30 for near transfer (medium effect), 0.31 for far transfer (medium effect), and 0.36 for the affective questionnaire (medium-to-large effect)" (Atkinson, 2002, p. 425).

Atkinson concluded that an animated pedagogical agent capable of delivering aural instruction and using nonverbal communication to draw attention to important information is more effective than text alone or audio alone to foster learning (Atkinson, 2002). This conclusion contradicts the findings from the Baylor and Ryu (2003b) study, which showed no statistical difference in the performance of participants in the audio-only condition and the animated-image condition.

In a similar study, Mayer, Dow, and Mayer (2003) conducted four experiments to "examine four features of self-explaining environments that might affect how well learners understand a scientific explanation" (p. 806). The four features were modality of instruction, interactivity, use of a prequestion to engage learners, and presence of agent image. Because interactivity and use of a prequestion are not relevant to the proposed research, this review will address only the first and fourth experiments, which dealt with modality of instruction and the presence of agent image respectively.

The first experiment examined the effect of agent voice on learning. The hypothesis, based on the modality principle of multimedia instruction, was that "students who learn from graphics and narration will learn more deeply and therefore perform better on problem-solving transfer tests than will students who learn from graphics and on-screen text" (Mayer, Dow, et al., 2003, p. 806). The researchers noted that because prior research had evaluated the modality effect in noninteractive multimedia environments, this experiment was designed to extend the theory to interactive multimedia environments.

The researchers used a convenience sample of 52 college students at the University of California, Santa Barbara. The participants, who all received course credit for their participation, were randomly divided into two groups (Mayer, Dow, et al., 2003). Participants were first administered a research-designed questionnaire to gather data regarding age, gender, SAT scores, knowledge of household repair, and experience in nine electrical-type activities. Next, each group received one of "two versions of an interactive multimedia program designed to teach students how an electric motor works" (Mayer, Dow, et al., 2003, p. 807). In both versions, the program presented a color illustration of an electric motor. When learners clicked on a part of the illustration, the program displayed a list of questions related to the selected part. Clicking a question caused an animated cartoon character named Dr. Phyz to appear and give the answer. One of the treatment groups received the answer as audio, and the other received it as on-screen text. After participants had completed the instruction, they were given seven paper-based problem-solving questions designed to test near transfer. These questions had previously been pilot tested to ensure their reliability and validity.

Participants in the audio group "generated significantly more answers on the problem-solving transfer test ($M = 8.43$, $SD = 2.56$) than did students in the text group ($M = 6.54$, $SD =$

2.22), $t(54) = 2.96$, $p = .0046$. The effect size was 0.85" (Mayer, Dow, et al., 2003, p. 809). The researchers concluded that the results supported the modality principle. Unfortunately, the results have a degree of ambiguity because *generated significantly more answers* can be interpreted as simply having answered more questions instead of answering more questions correctly, which would provide better evidence of learning.

As mentioned before, the second and third experiments were not relevant to the proposed study. However, experiment four was designed to examine the effect of image on learning. The researchers used a convenience sample of 39 college students and randomly assigned them to two groups. As with the first experiment, participants received credit for participation. Noting that Dr. Phyz's image provided "little or no instructional content" (Mayer, Dow, et al., 2003, p. 810), the researchers suggested that it may serve as a *seductive detail*, "an interesting but irrelevant part of the presentation" (p. 810) that distracts the learner. Thus they hypothesized that displaying Dr. Phyz's image would *not* result in students learning more deeply.

The researchers used the same materials and method in this experiment as they used in the first one, except that one group received a version of the multimedia program with narration but no agent image. The other group received the agent-plus-narration version used in the first experiment. Analysis of the results of the problem-solving transfer test supported the hypothesis: Participants in the agent-present condition "did not generate significantly more answers on the problem-solving transfer test ($M = 6.60$, $SD = 3.28$) than did students who learned with no agent present on the screen ($M = 5.95$, $SD = 3.52$), $t(37) = .55$, $p = .5835$" (Mayer, Dow, et al., 2003, p. 811).

Mayer, Dow, et al. (2003) concluded that "although the agent's voice is important for improving learning, the agent's physical image is not" (p. 811). If, as in this case, image is not

needed to present or clarify content, it may act as a seductive detail and actually distract the learner.

Craig, Driscoll, and Gholson (2004) took a somewhat different slant on the learning question, studying the relative learning gains of learners who actively engaged with an animated pedagogical agent compared with those of vicarious learners who observed rather than participated in the learning process. Although the study involved two experiments, this review will address only the first. The reason is that the second experiment replicated the first in procedures and results, with the exception of the addition of a collaborative vicarious learning condition, which is not relevant to the current research proposal unless it is simulated through an agent, which it was not.

The pedagogical agent in this study was AutoTutor, the same intelligent tutoring system (ITS) used by Link et al. (2001). As mentioned previously, AutoTutor is a talking head, a human-like image using facial expressions, gestures, and gaze while delivering dialog as synthesized speech using inflection and intonation. Whereas Link et al. asked participants only to evaluate preselected feedback segments, Craig et al. (2004) used AutoTutor to deliver instruction. AutoTutor was programmed to fill in pieces of information missing from expected answers and to try to explain any misconceptions that were detected. This study also differed from the one conducted by Moreno et al. (2001), in that "nearly all the learning that takes place results from dialog between the learner and the ITS" (Craig et al., 2004, p. 166).

The research questions were never expressly stated. However, they may be surmised from the descriptions of the data collection procedures and results:

1. Does active engagement during online instruction result in greater learning gains than vicarious learning through observation of the instructional interactions?

2. Does incorporation of a talking head result in greater learning gains than narration alone?
3. Does providing redundant text in addition to the audio result in greater learning gains than audio alone?

To answer these questions, the researchers selected a 2 x 2 x 2 factorial design, the factors being *instruction type* (interactive vs. vicarious), *talking head* (present vs. absent), and *printed text* (present vs. absent). These factors resulted in eight conditions, four of which were interactive and four of which were vicarious (Craig et al., 2004).

The sample for the first experiment consisted of 120 students drawn from an introductory psychology class at the University of Memphis. Craig et al. (2004) randomly assigned participants to one of eight cells, each of which represented one of the conditions being tested. Participants in each of the four interactive cells were *yoked* to participants in the corresponding vicarious cell so that the videotaped lesson from the former was shown to the latter.

After participants signed an informed consent form, a pretest was administered. The pretest was one of two multiple-choice tests, each of which consisted of 24 questions, 2 on each of the twelve topics covered in the computer-based instruction. Use of each test as either pretest or posttest was alternated for each participant in a cell, except for the 15th participant, who received a randomly selected pretest. The data from 33 participants were replaced, because their pretest scores were greater than 9, which had been set as the domain-knowledge criterion (Craig et al., 2004).

After the pretest, the participants took the computer-based lesson. Audio with intonation and inflection, created using Microsoft Agent's text-to-voice engine, was used in all four interactive conditions. In the talking-head-present condition, participants interacted with a

talking head that communicated through audio, animated facial expressions, gestures, and gaze. In the text-present condition, redundant on-screen text appeared in a dialog bubble. Each topic began with a brief aural information delivery followed by a question for the learner to answer. Participants contributed to the conversation by selecting options presented in a dialog box or by typing their contributions in the box. The program responded by presenting an appropriate follow-on script, based on recognition of key concepts present in the learner's input. Each session was saved to an AVI file and presented to a yoked participant in the corresponding vicarious-tutoring condition, who could only watch and listen to the session. After the computerized session, all participants took their assigned posttest (Craig et al., 2004).

A 2 x 2 x 2 ANOVA performed on the pretest data produced no significant effects for the before-treatment conditions. However, analyses on each condition revealed significant differences between pretest and posttest for the interactive conditions. "Contrasting the interactive tutoring condition with the vicarious condition revealed a significant effect, $F(1,111) = 4.28$, $p < .05$, in favor of interactive tutoring" (Craig et al., 2004, p. 172). Using Cohen's *d* statistic with the pooled standard deviation for each condition to compute the pretest-to-posttest effect size resulted in an effect size of 1.73 for the interactive condition and 1.19 for the vicarious condition. "This latter score, while smaller than the effect size in the interactive condition, is more than one standard deviation, which is usually considered a full letter grade (Cohen et al., 1982; Graesser & Person, 1994)" (p. 172).

The experiment failed to obtain a significant effect for redundancy (text with narration). Moreover, an ANOVA showed no effect for the presence of a talking head displaying facial expressions, gestures, and gaze when compared to spoken narration alone. "Collapsing over

other variables, the mean learning gain for talking-head present was 3.76 and the mean for talking-head absent was 3.73" (Craig et al., 2004, p. 173).

Contributions of the Literature to the Proposed Study

Taken together, the research tends to show that pedagogical agents in both the animated-image and audio-only modalities can be used to simulate social agency in self-paced distance instruction (Mayer, Sobko, et al., 2003; Moreno et al., 2001). Also, the research shows that the use of animated agents tends to support the perception of persona (Baylor & Ryu, 2003b; Link et al., 2001). However, neither of these studies isolated for audio.

The effect of agent modality on learning is the most contentious aspect, with research providing conflicting results as to the relative effectiveness of audio alone and narrated animation in promoting retention and transfer. It should be noted that of all the studies, only the one by Baylor and Ryu (2003b) isolated for static image, and no condition in their study had a significant effect on learning. Of the four studies on the effects of agent voice, image, and animation on learning, two (Craig et al., 2004; Mayer, Dow, et al., 2003) showed no significant effect for an animated image. Mayer, Dow, et al. concluded from their study that audio was more important for enhancing learning; Craig et al. found no significant difference between the animated-image condition and the audio-only condition. Moundridou and Virvou (2002) concluded that the main advantage to using animated agents is motivational.

On the other hand, Atkinson (2002) found that an animated pedagogical agent is more effective than text alone or audio alone to foster learning. This corroborated one set of findings from the study by Moreno et al. (2001), which indicated that use of an animated pedagogical agent achieved significantly better transfer than text alone. However, Moreno et al. concluded that whereas audio provided a significant advantage for learning, the animated agent did not.

Although it may be tempting to say that Atkinson's results were a fluke, a closer look reveals that they may have resulted from the way that the animated agent was used in the instruction. Of the studies that compared the effect of an animated agent with voice to the effect of audio alone, most did not make maximum use of the agent's nonverbal communication capabilities for instructional purposes. For instance, the animated agent in two of the studies (Baylor & Ryu, 2003b; Mayer, Dow, et al., 2003) merely spoke to the learner. In Atkinson's study, the agent gestures and glances toward a step in the problem while explaining it.

Clearly, the research indicates the need for more research into the effects of voice, image, and animation on agent persona and learning. Not only is there disagreement in the findings, but also only one study isolated for static image, a common agent modality in lower-end online courseware. The proposed study will isolate for each of these variables to determine the efficacy of using a pedagogical agent to mitigate the lack of social interaction and positively affect learning. It will also expand agent research to a new learning environment: online military training—specifically, online Navy "A" school courses. "A" school provides the basic skills required to perform a specific job in the Navy, such as aviation electronics. Starting in the late 1990s, the Navy began converting instructor-led "A" school courses to online asynchronous instruction. Hosted on Navy e-Learning, these courses can be taken anywhere, anytime.

Research Questions

The purpose of this study is to relate pedagogical agent voice and image to learning and the perception of social agency for sailors enrolled in an online "A" school course distributed via Navy e-Learning. The study will address the following research questions:

1. How do agent voice and image, alone or in combination, relate to the perception of social agency for sailors taking an online "A" school course distributed via Navy e-Learning?
2. How do agent voice and image, alone or in combination, relate to learning for sailors taking an online "A" school course distributed via Navy e-Learning?
3. How does perception of social agency vary under different combinations of agent voice and image for sailors taking an online "A" school course distributed via Navy e-Learning?
4. How does learning, as evidenced by scores on a posttest, vary under different combinations of agent voice and image for sailors taking an online "A" school course distributed via Navy e-Learning?

The null hypothesis for the first research question is that there is no significant difference between online instruction with no agent audio or image and online instruction with a pedagogical agent operationalized using audio and/or image in terms of social agency for sailors enrolled in an "A" school course distributed via Navy e-Learning. The alternative hypothesis is that a pedagogical agent operationalized using audio and/or image in an online "A" school course delivered via Navy e-Learning increases learners' perception of social agency.

The null hypothesis for the second research question is that there is no significant difference between online instruction with no agent audio or image and online instruction with a pedagogical agent operationalized using audio and/or image in terms of learning, as evidenced by scores on a posttest, for sailors enrolled in an "A" school course distributed via Navy e-Learning. The alternative hypothesis is that a pedagogical agent operationalized using audio

and/or image in an online "A" school course delivered via Navy e-Learning increases learning, as evidenced by scores on a posttest.

The null hypothesis for the third research question is that there is no significant variation in perception of social agency under different combinations of agent voice and image for sailors taking an online "A" school course delivered via Navy e-Learning. The alternate hypothesis is that there is significant variation in perception of social agency under different combinations of agent voice and image for sailors taking an online "A" school course delivered via Navy e-Learning.

The null hypothesis for the fourth research question is that there is no significant variation in learning, as evidenced by scores on a posttest, under different combinations of agent voice and image for sailors taking an online "A" school course delivered via Navy e-Learning. The alternate hypothesis is that there is significant variation in learning, as evidenced by scores on a posttest, under different combinations of agent voice and image for sailors taking an online "A" school course delivered via Navy e-Learning.

Methods

The study will use a 2 x 3 factorial design to examine the independent and simultaneous effects of pedagogical agent voice and image on learning and the perception of social agency. The primary advantages in using a factorial design, according to Harris (1998), are as follows:

1. It is a more efficient way to assess the effect of two independent variables on two dependent variables than conducting four separate studies.
2. It permits assessment of the interactions among variables.
3. It permits greater confidence in the effect of either independent variable because the analysis conclusions would be valid for all levels of the other independent variable.

4. It permits a more accurate assessment of the effect of an independent variable than could be made by testing the effect of the variable by itself because the effect of the other variable is taken into consideration in the analysis.

The disadvantage of using a factorial design is that it requires a larger sample size. However, use of a sample size table will enable estimation of an adequate number of participants per group to ensure sufficient power.

The first factor, agent voice, will have two levels: *absent* and *present*. The second factor, agent image, will have three levels: *absent*, *still*, and *animated*. Combining factors by level will result in six groups, as shown in Table 1.

Table 1

Combinations of Factors by Level

	Image		
Voice	Absent	Still	Animated
Absent	No voice and no image	No voice and still image	No voice and animated image
Present	Voice and no image	Voice and still image	Voice and animated image

Setting

The proposed setting for the study is the Navy's online distributed learning environment, Navy e-Learning. The study will likely be conducted as part of a contract either to convert an instructor-led course to an online course or to revise a current online course. There are several avenues for making this setting possible. One is a contract through the Office of Naval Research (ONR), which funds basic and applied research in areas of knowledge and technology and has

cooperative and collaborative research and development agreements with various non-profit institutions and commercial firms. The study could also be pursued as part of an ongoing contract between the researcher's company and the Naval Education and Training Command (NETC). A third avenue would be the Department of Defense's Small Business Innovation Research (SBIR) program, which is designed to stimulate technological innovation as well as use small businesses to meet the Department's research and development needs. One of the areas of SBIR research is agent technology.

Because this research setting involves a contract, the specific course cannot be identified before contract negotiations begin. There will be two requirements for any course selected for the study: (a) It must be an "A" school course, and (b) its throughput must be sufficiently high to ensure that the minimum number of participants required for the study will have completed the course within the proposed maximum data collection period of six months. The specific course to be used in the study will be determined by the Navy at the time of contract, based on the Navy's needs at that time and the required parameters. After the Navy has identified the course, a suitable module will be selected for development of the treatment. Six versions of this module will be developed, one for each combination of agent voice and image.

Participants

The population of interest consists of naval personnel who are enrolled in online "A" schools and who may be located anywhere in the world, either aboard ships or at shore stations. This population was selected because it consists primarily of 17- to 20-year-old high school graduates who have had little or no experience with independent, self-paced online instruction and are therefore more likely to feel a lack of social context for learning more acutely. The target population will be students who register for the "A" school course identified for use in the study.

Sampling. Because this study will likely be conducted as part of a Navy contract with a limited time period for performance, it will employ a convenience sample of all the students who register for the course within a contract-specified time period. This time period will be calculated based on the average course throughput, time to complete training, and a minimum sample size of 65 participants per group (N = 390). Obviously, the higher the throughput is, the more students who will be available to participate at any one time, and the shorter the data collection period.

The sample size was calculated using Lipsey's (as cited in Creswell, 2005) power analysis table, using a significance level of alpha = .05, a power level of .80, and an effect size of .50. Students who consent to participate in the study will be assigned to one of the six treatment groups using an automated systematic sampling plan, such that every sixth student receives treatment 1, the next student to register after that student receives treatment 2, the next student after that student receives treatment 3, and so on.

Informed consent. Although the contract will grant permission to conduct the study, permission will also be obtained from each participant. After students register for and log onto the course, they will be presented the online consent form (see Appendix B). This form not only provides prospective participants an overview of the study but also enables informed consent. The overview includes the purpose of the study, a brief description of it, how the learner will contribute to it, and the importance of the learner's participation. To encourage participation and reduce sampling error, prospective participants will be offered a simple, in-house-developed Web page builder (EISE®) for free download after they complete their input to the study.

In respect of participants, the consent form states participants' rights, including their right not to participate and their right to withdraw from the study at any time. It also assures potential

participants that no identifying information will be associated with any of the findings or used in the final report. The form includes an e-mail address so that students can ask questions of a research associate before making their decision to participate. This address will be included in the help section of the courseware so that participants will always have access to someone who can answer questions about the study.

Learners who opt out of the study will take a version of the course that has no agent voice or image. If learners elect to participate, the program will register their consent and automatically assign them to one of the six treatments. Learners who later withdraw from the study will still need to complete the course if they want credit toward "A" school completion, but they will not need to provide any inputs related strictly to the study. All prior inputs of this nature will be deleted, and course data that has been collected will not be considered in the analysis.

Instruments

Data to be collected during the study will consist of background information, performance measures, and attitudinal measures. Instruments will include a questionnaire developed specifically for this study, pre-existing or contract-developed online tests, and a slightly modified version of Baylor's and Ryu's (2003a) Agent Persona Instrument.

Background data. Background information data will be collected using an online questionnaire (see Appendix C), which will consist of a participant identifier that can be cross-referenced against the standard student identifier in the Navy learning management system (LMS) and predominantly closed-ended items to elicit the following information: age, gender, educational background, learning environment, computer experience, and experience/comfort level with online classes. The purpose of collecting this background data will be to identify possible mediating variables that may influence the results of the study. To determine the

questionnaire's face validity, the instrument will be submitted for review by a senior instructional designer with extensive experience in questionnaire development. After incorporating any comments, permission will be sought from the Director of Training at Naval Technical Training Center Corry Station, Pensacola, to administer the instrument to a convenience sample of "A" school students (N = 30) who volunteer to participate in a reliability test of the instrument. The instrument will be administered twice to this group, one month apart, to determine test-retest reliability.

Performance measure. The performance measures will be students' mean scores on test items associated with the treatment module on an online pretest and posttest. Depending on the "A" school course used for the study, the pretest and posttest may be part of an existing online course or may be developed as part of a contract to convert a course to online instruction. If a test is developed, it will conform to the test specifications set by the contract. Traditionally, specifications for online tests call for a single item bank with a minimum of four questions per objective. The pretest and posttest are dynamically generated when the learner elects to take the test through the use of automated random sampling stratified by objective.

Test items may be multiple-choice, matching, or short answer. The items will be submitted for review by Navy subject matter experts, who will determine content validity. It is assumed that any existing assessment will already have undergone such a review. Test statistics generated from use of both existing or newly developed instruments will be used to determine alternate forms reliability, and individual item analysis will be conducted to determine item reliability.

Attitudinal measure. Data for the attitudinal measure will be collected on a slightly modified version of the Agent Persona Instrument (API), developed by Baylor and Ryu (2003a)

to measure learners' perceptions of pedagogical agent persona (see Appendix D). This instrument differs from instruments developed by other researchers in that it gives consideration to the computer-based aspects of the agent. Baylor and Ryu developed the instrument by collecting and collating items from other instruments and deleting duplicate items and items that did not specifically measure agent persona. Initially, the instrument had 38 items. To score the items, they used a five-point Likert scale (1 = *strongly disagree*, 2 = *disagree*, 3 = *neutral*, 4 = *agree*, and 5 = *strongly agree*). They validated the instrument in a two-phase study using undergraduate students working within the Multiple Intelligent Mentors Instructing Collaboratively (MIMIC) agent-based research environment.

For the first phase, the researchers used a convenience sample of 80 students. Baylor and Ryu (2003a) administered the instrument after the students completed the MIMIC assignment. Then the researchers conducted an exploratory factor analysis to identify the factorial structure for the instrument. Test data supported a five-factor model. Examination of the strength of association between the factors and items resulted in the deletion of eight items. Baylor and Ryu then performed a second factor analysis on the revised list of items, which resulted in an improvement of the explained variance from 68.76% to 72.66%. In addition, the researchers established a high reliability for the internal consistency of each factor based on Cronbach's alpha. They then labeled the five factors "to represent the agent persona features of Credible, Facilitating Learning, Mentor-like, Engaging, and Human-like" (p. 2).

To validate their findings, Baylor and Ryu (2003a) conducted a second experiment using a convenience sample of 133 students and MIMIC. A review of the correlation coefficients and examination of the contextual meaning of the items resulted in the researchers dropping *mentor-like* and its associated items from the instrument. "It was determined that these items were more

appropriate for a human teacher or a more adaptive agent system rather than general pedagogical agents" (p. 3). A confirmatory factor analysis on the results indicated a much improved model fit, validating the current four-factor instrument. Operational definitions of the four factors, according to Baylor and Ryu, are as follows:

- Facilitating learning: "how well the agent helps the student through the learning process, e.g., by promoting reflection and learning support" (p. 3)
- Credible: "the credibility and believability of the agent and its advice for helping the learner understand the learning content" (p. 3)
- Human-like: "the agent's behaviour [sic] and emotional expression in terms of its naturalness and personality" (p. 3)
- Engaging: "how entertaining and enjoyable it was for the learner to work with the agent" (p. 3)

Baylor, Ryu, & Shen (2003) used three of the API sub-scales (*credible*, *engaging*, and *human-like*) in a subsequent study of perceptions of agent persona. Using Cronbach's alpha, they assessed the overall reliability of the instrument at 0.97.

Repurposing the API for use with this study involved formatting the instrument for online delivery and adding appropriate directions for completing it. In addition, two changes were made to the items themselves. First, the items under *human-like* were rearranged to a more logical sequence, so that the second item (The agent's emotion was natural) appears after the fifth item (The agent showed emotion). Second, the fourth item under *human-like* was deleted because it deals with agent movement, and four versions of the treatment content will not involve animation. A total score of all the subscales will be used to measure social agency because the

subscales are not sufficiently discrete to provide useful contrasts. There is too much overlap between *engaging* and *facilitating learning*, and between *human-like* and *credible*.

Data Collection

The experiment will use six versions of one module of the online "A" school course identified for the study, each identical except for agent condition: (a) no voice and no image, (b) no voice and still image, (c) no voice and animated image, (d) voice and no image, (e) voice and still image, and (f) voice and animated image. The only difference between the voice-present and voice-absent conditions will be the media used for instructional guidance: The voice-absent conditions will present the same instructional guidance as the voice-present conditions but will use text instead. The audio used in all voice conditions will be a recorded human voice, as research has indicated that compared to audio created by text-to-speech engines, a human voice is superior for promoting perception of agent persona (Atkinson, 2002; Baylor et al., 2003). The animated image will be a cartoon version of a Navy instructor with the fictitious name of Petty Officer Cantora, which will be developed using e frontier® Poser®. The animated image will use facial expressions, including eye movement to draw attention to content. For the voice-plus-image condition, the animated image will also use lip synchronization. The still image will be a static frontal view of Petty Officer Cantora. The no-voice/no-image condition will serve as a control for the five agent conditions.

After participants complete the Background Information Questionnaire, they will take the online course, beginning with the pretest to assess their current level of knowledge. They will continue with the course at their own pace, in accordance with Navy Integrated Learning System policy. An analysis of covariance will control for the effects of course completion time. When participants access the treatment module, Petty Officer Cantora will introduce himself in each of

the five agent conditions. After participants complete the module, they will be asked to complete the API and submit it. When they complete the course, they will take the posttest. As soon as the LMS registers their score on the posttest, they will receive an appreciation message containing the Web address for the free download of EISE®.

Data Analysis

Data will be collected electronically and stored in a SQL database. To eliminate problems with missing data, the program will be designed so that participants will not be able to submit an instrument until they have provided all required responses. The research team will review the data to find any data entry errors or anomalies and will then clean it. Data for any participant who did not submit all four instruments (the questionnaire, the API, the pretest, and the posttest) will be deleted from the database. Researchers will attempt to match sets of incomplete data resulting from errors in entering the participant identifier. For instance, if participant 3147 completed the questionnaire but not the API, and participant 3157 completed the API but not the questionnaire, it will be assumed that 3147 and 3157 are the same person.

After cleaning the data, the research team will code the data from the two open-ended questions on the questionnaire. Similar answers to these questions will be grouped and assigned a code and numerical identifier.

The research team will use SPSS for Windows® to import the data from the SQL database and analyze it. Because the study is intended to examine the effect of two independent categorical factors (voice and image) on two continuous dependent variables (learning and perception of social agency), a multiple analysis of variance (MANOVA) is indicated to determine the main effects and interaction effects for voice and image on the dependent variables. The measure of the learning variable will be the posttest score on test items for the

treatment module, and the measure of the perception of social agency variable will be the total of all subscales on the API. This analysis meets the following assumptions for MANOVAs:

1. The factors are independent; voice and image are unrelated with the exception of lip synchronization, which will be controlled by using only facial expressions in the voice-absent/animated-image condition.
2. The independent variables are categorical.
3. The dependent variables are continuous and interval level.
4. Participants will be randomly assigned to treatment groups.
5. The groups will be equal.

There are several covariates that may intervene with these effects. To determine the degree to which these variables account for the total variation in the factors of voice and image, the research team will run a series of multiple analyses of covariance (MANCOVAs) using the pretest score, course completion data, and data from the Background Information Questionnaire as covariates. Covariates from the questionnaire will include the following: gender, level of education, trade school experience, location, environment, computer access, computer use, experience with online instruction, and comfort level with online instruction.

The only descriptive statistics anticipated are the mean and standard deviation for the dependent variables for each treatment group. Examination of the means will permit the research team to reach preliminary conclusions regarding the relative effectiveness of the various treatments. The standard deviation is used to calculate the F statistic, which enables a determination to be made regarding the statistical significance of the differences among the means. Use of the SPSS for Windows® general linear models, multivariate procedure will yield an F statistic and significance level for each main effect and interaction. If the main effect for

image is significant at the .05 level with 2 degrees of freedom (2 levels of voice minus 1 X 3 levels of image minus 1), a post hoc Tukey HSD test will be used to determine which image level contributes most to the explanation, collapsing across the levels for voice. The Tukey test is preferred when the number of groups is large. If the *trade school* covariate, *experience with online instruction* covariate, or c*omfort with online instruction* covariate accounts for a significant amount of the variance, then the research team will examine the explanatory data (area of trade school specification, specific online course experiences, and reasons for discomfort with online instruction) to pinpoint the reasons.

Discussion

Careful consideration has been given to control for threats to internal and external validity in the proposed study. For instance, the design recognizes several intervening variables that might affect the results and has included them as covariates. The danger in doing so is that the main effects may be so small that the only option is to retain the null hypotheses.

Even if there are conclusive results in favor of using one of the pedagogical agent modalities, further study will be needed to verify the results with other online military audiences and to determine other aspects of agent use. For instance, what is the effect of agent gender and color on learners' perceptions of the agent? Would the ability to select agent characteristics increase learners' motivation and engagement? What is the relationship between agent modality and content? Regardless, the proposed research will add to what is now known about pedagogical agents and may provide a design rationale for instructional designers involved in designing independent, self-paced online courseware for the military.

References

Atkinson, R. K. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology, 94*, 416–427.

Baylor, A. L., & Kim, Y. (2003). Validating pedagogical agent roles: Expert, motivator, and mentor. *World Conference on Educational Multimedia, Hypermedia and Telecommunications (2003)*, 463-466. Retrieved September 15, 2005, from http://www.aace.org/dl/files/EDMEDIA2003/paper_3012_5460.pdf

Baylor, A. L., & Ryu, J. (2003a). *The API (agent persona instrument) for assessing pedagogical agent persona*. Paper presented at the 2003 World Conference on Educational Multimedia, Hypermedia, and Telecommunications. Retrieved September 15, 2005, from the AACE Digital Library Web site: http://www.aace.org/dl/files/EDMEDIA2003/paper_3012_5457.pdf

Baylor, A. L., & Ryu, J. (2003b). The effects of image and animation in enhancing pedagogical agent persona. *Journal of Educational Computing Research, 28*, 373–394.

Baylor, A. L., Ryu, J., & Shen, E. (2003). *The effects of pedagogical agent voice and animation on learning, motivation and perceived persona.* Paper presented at the 2003 World Conference on Educational Multimedia, Hypermedia, and Telecommunications. Retrieved September 15, 2005, from the AACE Digital Library Web site: http://www.aace.org/dl/files/EDMEDIA2003/paper_3012_5657.pdf

Caine, R. N., Caine, G., McClintic, C., & Klimek, K. (2005). *12 brain/mind learning principles in action: The fieldbook for making connections, teaching, and the human brain*. Thousand Oaks, CA: Corwin Press.

Craig, S. D., Driscoll, D. M., & Gholson, B. (2004). Constructing knowledge from dialog in an intelligent tutoring system: Interactive learning, vicarious learning, and pedagogical agents. *Journal of Educational Multimedia and Hypermedia, 13*, 163-183.

Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology, 94*, 428–434.

Creswell, J. W. (2005). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (2nd ed.). Upper Saddle River, NJ: Pearson Education.

Gilbert, J. (2002). Playing teacher. *World Conference on E-Learning in Corporations, Government, Health, and Higher Education, 2002*, 325-330. Retrieved September 17, 2005, from http://www.aace.org/dl/files/ELEARN2002/paper_3009_2521.pdf

Harris, M. B. (1998). *Basic statistics for behavioral science research* (2nd ed.). Upper Saddle River, NJ: Pearson Education.

Kim, Y., & Baylor, A. L. (in press). A social-cognitive framework for pedagogical agents as learning companions. *Educational Technology Research and Development.* Retrieved September 11, 2005, from http://it.usu.edu/~yangheekim/documents/Socio-CognitiveFrameworkForPALs.pdf

Link, K. E., Kreuz, R. J., Graesser, A. C., and the Tutoring Research Group. (2001). Factors that influence the perception of feedback delivered by a pedagogical agent. *International Journal of Speech Technology, 4*, 145–153.

Mayer, R. E., Dow, G. T., & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: What works in the design of agent-based microworlds? *Journal of Educational Psychology, 95*, 806–813.

Mayer, R. E., Sobko, K., & Mautone, P. D. (2003). Social cues in multimedia learning: Role of speaker's voice. *Journal of Educational Psychology, 95*, 419–425.

Moreno, R., & Mayer, R. E. (n.d.). *Pedagogical agents in constructivist multimedia environments: The role of image and language in instructional communication*. Retrieved September 9, 2005, from http://www.unm.edu/~moreno/PDFS/Roundtble.pdf

Moreno, R., Mayer, R. E., & Lester, J. C. (2000). Life-like pedagogical agents in constructivist multimedia environments: Cognitive consequences of their interaction. *World Conference on Educational Multimedia, Hypermedia and Telecommunications 2000*, 776-781. Retrieved September 18, 2005, from http://www.aace.org/dl/files/ EDMEDIA2000/EdMedia2000p776.pdf

Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction, 19*, 177–213.

Moundridou, M., & Virvou, M. (2002). Evaluating the persona effect of an interface agent in an intelligent tutoring system. Draft version of a paper published in *Journal of Computer Assisted Learning, 18*. Retrieved September 16, 2005, from http://citeseer.ist.psu.edu/ cache/papers/cs/25746/http:zSzzSzthalis.cs.unipi.grzSz~mariamzSzJCAL.pdf/ moundridou02evaluating.pdf

Reeves, B., & Nast, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places.* New York: CSLI.

Roblyer, M. D. (2004). *Integrating educational technology into teaching* (3rd ed.). Upper Saddle River, NJ: Pearson Education.

Shaw, E., Johnson, W. L., & Ganeshan, R. (1999). *Pedagogical agents on the web*. Retrieved September 16, 2005, from the Information Sciences Institute web site: http://www.isi.edu/isd/ADE/papers/its98/ITS98-WW.htm

Ting, C., & Chong, Y. (2003). *Enhancing conceptual change through cognitive tools: An animated pedagogical agent approach*. Paper presented at the Third IEEE International Conference on Advanced Learning Technologies. Retrieved September 11, 2005, from the IEEE Computer Society Digital Library site: http://csdl2.computer.org/comp//proceedings/icalt/2003/1967/00/19670314.pdf

Appendix A

Literature Map

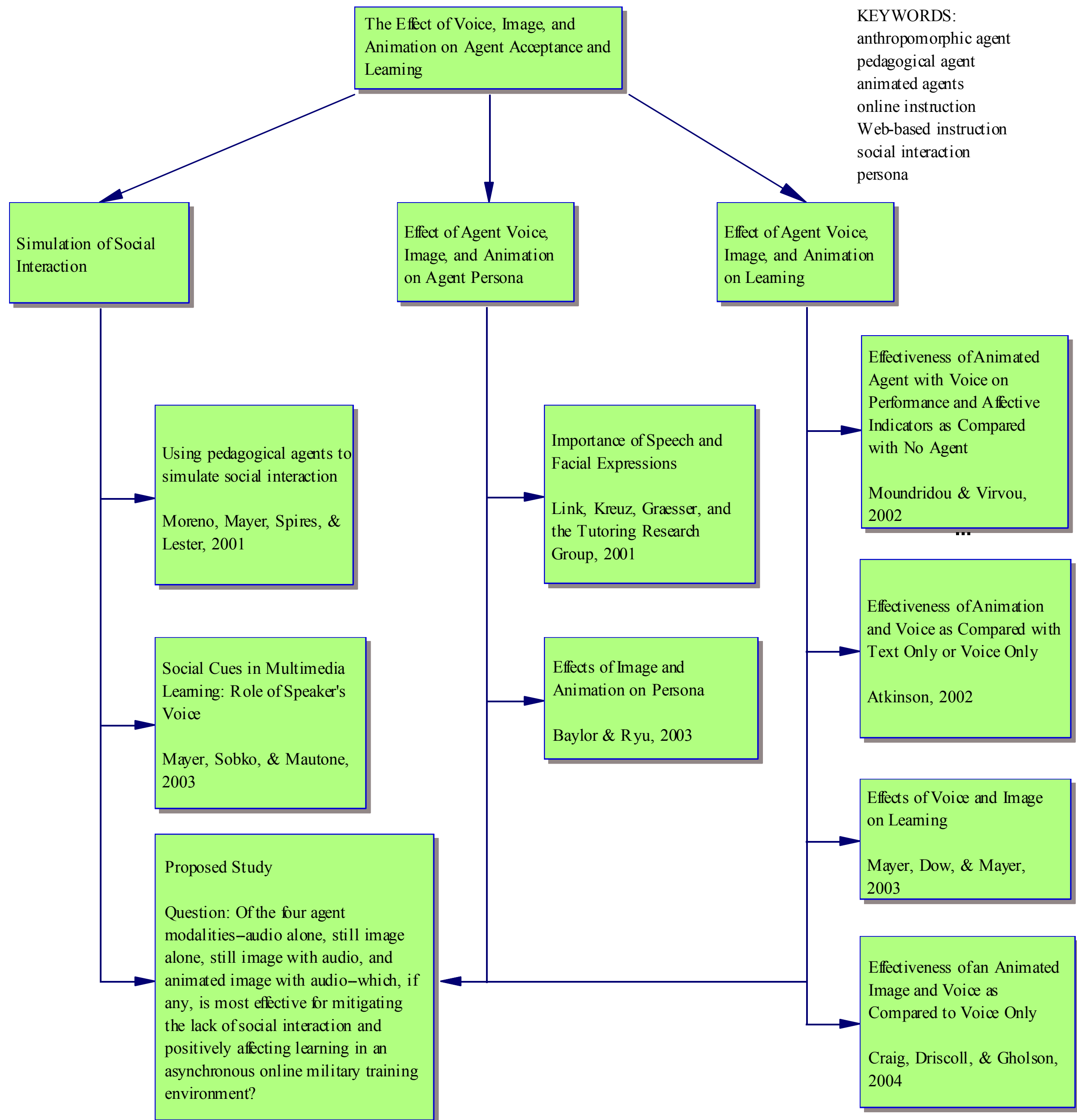

Figure 1. A thematic map showing the relationship of the proposed study to other research on the ability of pedagogical agents to simulate social interaction, promote a perception of social agency, and enhance learning.

Appendix B

Consent to Participate in a Study

You are being asked to participate in a study authorized by the U. S. Navy to assess the effectiveness of using instructional agents in online "A" school courses. An instructional agent is a media element (audio, image, or a combination of both) that suggests the presence of another human being and is intended to help you learn. The course for which you just registered may utilize an instructional agent for one of its modules to provide instructional presentations, ask and answer questions, introduce exercises, and provide feedback. You have been randomly assigned to one of six different versions of this module, five of which have an instructional agent. The only difference in the six versions is the presence/absence of an agent and the form of the agent when one is used. The content provided is the same.

If you agree to participate in this study, you will provide input by taking the course and associated assessments as well as completing the following:

- A brief background questionnaire asking for your age, gender, educational background, and experience with computers and online classes
- A second questionnaire (for those who receive one of the modules with an instructional agent) that you complete after the module, asking you to rate the agent in four different areas

Your input is important!

However, as important as your input is to the success of this study, your participation is entirely voluntary. Be assured that your confidentiality and anonymity will be maintained. With the exception of your course data, which is automatically incorporated in your five-vector model for professional development, the input you provide will be used only for this research effort. Your results will be compiled with those of other participants so that no identifying information will be associated with any of the findings in the resulting report.

There is no penalty if you decide later to withdraw.

If you agree to participate in the study, you still have the option of withdrawing at any time. You may still be presented material using a pedagogical agent; however, you will not be required to answer any questions regarding it, and your course data will not be used in the study. There is no penalty of any kind for withdrawal.

You get a cool reward for participating!

Even if you do not get a module with an instructional agent, if you complete the study you will be rewarded for your participation with a free, personal copy of EISE® Web-site development

software. Even people who have no experience with html can use this easy-to-use, self-explanatory program to create dynamic personal Web sites.

Do you have any questions?

If you have any questions now or in the future, please contact the research team at agent@idsi.com. This address will always be available to you during the study by looking in the course help menu. You are guaranteed to get an answer in 12 hours or less.

Are you willing to help? If so, please select the *consent* button below.

I consent to participate in the study.

I want to opt out of the study.

Appendix C

Background Information Questionnaire

You have consented to participate in a study to determine whether an instructional agent will enhance the learning experience in online "A" school instruction and, if so, what type of agent is most effective. Before you begin the instruction, please answer the following ten questions.

1. What are the last 4 digits of your social security number?

2. How old are you?

3. What is your gender?
   - Female
   - Male

4. What level of education have you achieved? (Select one.)
   - High school graduate
   - Some college
   - A.A. or A.S. degree
   - B.A. or B.S. degree

5. Have you taken courses at a trade or vocational school? (Select one.)
   - None
   - Some
   - Graduate

   If you attended or graduated from a trade school, please enter your area of specialization:

6. Where are you taking this course? (Select one location and one environment.)

   Location:
   - Aboard a ship
   - At a schoolhouse
   - At a shore-based station, but not in a schoolhouse

   Environment:
   - In a classroom
   - In a Library Multimedia Resource Center
   - In a work center or quarters

7. In the last two years before joining the Navy, what access did you have to a computer? (Check all applicable answers.)
   - ☐ I had my own computer.
   - ☐ I used my family's computer.
   - ☐ I used a friend's computer.
   - ☐ I used a computer at school or in a library.
   - ☐ I used a computer at work.
   - ☐ I had no access or very infrequent access to a computer.

8. In the last two years before joining the Navy, how frequently did you use a computer for each of the following? (Select one answer for each use type.)

Use Types	Daily	At least once weekly	Every other week	Infrequently	Never
Communication (e.g., e-mail, IM)	○	○	○	○	○
Web-surfing	○	○	○	○	○
School papers/research	○	○	○	○	○
Luck/Skill games (e.g., Free Cell, Tetris)	○	○	○	○	○
Fantasy/Action-adventure games (e.g., Diablo, Sims)	○	○	○	○	○

9. What is your experience with online instruction? (Select one of the first two options.)
   - ○ This will be my first experience.
   - ○ I have taken at least one class online. (If so, select as many of the following as are applicable.)
     - ☐ I have taken online instruction in a classroom with an instructor present.
     - ☐ I have taken an online class that was facilitated by an online instructor.
     - ☐ I have taken an online class in which I had no contact with an instrutor.
     - ☐ I was usually able to get help when I needed it.
     - ☐ I communicated in person with other students who were taking the class.
     - ☐ I communicated electronically with other students who were taking the class.

10. How comfortable do you feel with using a computer for taking a class? (Select one.)

- Very comfortable
- Somewhat comfortable
- Somewhat uncomfortable
- Very uncomfortable

If you marked *somewhat uncomfortable* or *very uncomfortable* above, please describe why.

Submit answers and begin course

Opt out of study

Note. Learners who select *Opt out of study* will be presented a pop-up note box with the following:

> You have indicated that you no longer wish to participate in the study. Please reconsider your decision. Your input will provide critical data needed to make recommendations regarding the use of pedagogical agents to enhance online learning. The time you take now will not only benefit you (remember the free Web site development program you will receive when you complete your input) but will also benefit thousands of future online "A" school students.

This note box will include two button options: (a) Return to questionnaire, and (b) Opt out and start course.

Appendix D

Agent Persona Instrument (API)

Baylor and Ryu
2003

Rate the Agent

As you learned in the introduction to the study, an instructional agent is a media element that suggests the presence of another human being and is intended to help you learn. The instruction that you just completed was facilitated by just such an agent: PO1 Cantora. Now we'd like to get your reactions. You will be rating PO1 Cantora in four different areas: (a) facilitating learning, (b) credible, (c) human-like, and (d) engaging. Your opinion is very important in helping us find out whether an agent is actually helpful. Also, even though you experienced only one type of agent, your answers will help us determine what type of agent, if any, is most helpful.

First, please enter the last 4 digits of your social security number. [ ]

Now, please rate PO1 Cantora. There are ten statements for *facilitating learning*, five for *credible*, four for *human-like*, and five for *engaging*. For each area, please rate the extent to which you agree with each statement by selecting a number from 1 to 5 according to the following scale:

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

1	2	3	4	5	Area 1: Facilitating Learning
☐	☐	☐	☐	☐	1. The agent led me to think more deeply about the presentation.
☐	☐	☐	☐	☐	2. The agent made the instruction interesting.
☐	☐	☐	☐	☐	3. The agent encouraged me to reflect what I was learning.
☐	☐	☐	☐	☐	4. The agent kept my attention.
☐	☐	☐	☐	☐	5. The agent presented the material effectively.
☐	☐	☐	☐	☐	6. The agent helped me to concentrate on the presentation.
☐	☐	☐	☐	☐	7. The agent focused me on the relevant information.
☐	☐	☐	☐	☐	8. The agent improved my knowledge of the content.
☐	☐	☐	☐	☐	9. The agent was interesting.
☐	☐	☐	☐	☐	10. The agent was enjoyable.

1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

1	2	3	4	5	Area 2: Credible
☐	☐	☐	☐	☐	1. The agent was knowledgeable.
☐	☐	☐	☐	☐	2. The agent was intelligent.
☐	☐	☐	☐	☐	3. The agent was useful.
☐	☐	☐	☐	☐	4. The agent was helpful.
☐	☐	☐	☐	☐	5. The agent was instructor-like.

1	2	3	4	5	Area 3: Human-Like
☐	☐	☐	☐	☐	1. The agent has a personality.
☐	☐	☐	☐	☐	2. The agent was human-like.
☐	☐	☐	☐	☐	3. The agent showed emotion.
☐	☐	☐	☐	☐	4. The agent's emotion was natural.

1	2	3	4	5	Area 4: Engaging
☐	☐	☐	☐	☐	1. The agent was expressive.
☐	☐	☐	☐	☐	2. The agent was enthusiastic.
☐	☐	☐	☐	☐	3. The agent was entertaining.
☐	☐	☐	☐	☐	4. The agent was motivating.
☐	☐	☐	☐	☐	5. The agent was friendly.

© 2003 by Amy L. Baylor and Jeeheon Ryu. Permission to copy, transmit, or use this instrument must be granted by the copyright holder.

[Submit answers] [Opt out of study]

Note. Learners who select *Opt out of study* will be presented a pop-up note box with the following:

> You have indicated that you no longer wish to participate in the study. Please reconsider your decision. Your input will provide critical data needed to make recommendations regarding the use of pedagogical agents to enhance online learning. The time you take now will not only benefit you (remember the free

Web site development program you will receive when you complete your input) but will also benefit thousands of future online "A" school students.

This note box will include two button options: (1) Return to questionnaire, and (2) Opt out and continue course.