REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT DATE	3. DATES COVERED *(From - To)*
10-02-2004	Final Technical Report	10/1/01 - 9/30/03

4. TITLE AND SUBTITLE
Measurement and Evaluation of Animated Pedagogical Agents and Their Use in Training

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N00014-02-1-0120

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Robert K. Atkinson
Arizona State University

Mary Margaret Merrill
Louisiana State University- Shreveport

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Mississippi State University
Box 6156
Mississippi State, MS 39762

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
Ballston Centre Tower One
800 North Quincy Street
Arlington, VA 22217-5660

10. SPONSOR/MONITOR'S ACRONYM(S)

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release- Distribution is unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This project entailed a total of six experiments, five conducted in a laboratory and one implemented in a field setting. Across all six experiments, the multimedia computer-based instruction focused on fostering learners' proportional reasoning ability in the context of multi-step word problems. The experiments led to the recognition of four different effects: (a) a *voice effect*, which suggest that designers of multimedia learning environments should create life-like on-screen agents that speak in a human voice rather than a machine-synthesized voice; (b) an *image effect*, an agent's image fosters learning when it is programmed to explain complex visual information aurally; (c) an *embodiment effect*, which suggests that in a linear computer-based environment, the visual presence of an animated agent is a critical factor in optimizing learning outcomes whereas an agent's mobility is a less important factor; and (d) a *sequential effect*, which suggest that sequentially presented subgoals are superior to simultaneously presented subgoals in example-based instruction.

15. SUBJECT TERMS
Animated Pedagogical Agents, Social Agency Theory, Visual Cues, Multimedia Learning

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPONE NUMBER *(Include area code)*

FINAL REPORT

Award
ONR N00014-02-1-0120

Title
Measurement and Evaluation of Animated Pedagogical Agents and Their Use in Training

Original Principle Investigator
Robert K. Atkinson

Authors of Report
Robert K. Atkinson
Division of Psychology in Education
Arizona State University

Mary Margaret Merrill
Department of Psychology
Louisiana State University-Shreveport

Date of Submission
February 16, 2004

Current Principle Investigator
Thomas Hosie

Institution
Mississippi State University

Table of Contents

ABSTRACT

This project entailed a total of six experiments, five conducted in a laboratory and one implemented in a field setting. Across all six experiments, the multimedia computer-based instruction focused on fostering learners' proportional reasoning ability in the context of multi-step word problems. The experiments led to the recognition of four different effects: (a) a *voice effect*, which suggest that designers of multimedia learning environments should create life-like on-screen agents that speak in a human voice rather than a machine-synthesized voice; (b) an *image effect*, an agent's image fosters learning when it is programmed to explain complex visual information aurally; (c) an *embodiment effect*, which suggests that in a linear computer-based environment, the visual presence of an animated agent is a critical factor in optimizing learning outcomes whereas an agent's mobility is a less important factor; and (d) a *sequential effect*, which suggest that sequentially presented subgoals are superior to simultaneously presented subgoals in example-based instruction.

SUMMARY

Objectives

The goal of this project was to directly examine the positive learning gains associated with using animated pedagogical agents in early phases of cognitive skill acquisition and search for ways to optimize these gains during training. The research was also designed to provide fruitful guidance for the design of learning environments that deploy animated pedagogical agents.

Approach

The research included a total of six experiments, five conducted in a laboratory and one implemented in a field setting. Across all six experiments, the multimedia computer-based instruction focused on fostering learners' proportional reasoning ability in the context of multi-step word problems. These experiments were further grouped into three sets of experiments. The first set, Experiments 1 and 2, examined the impact of the agent's voice on performance. The second set, Experiments 3 and 4, examined the compared the effectiveness of agents to highlighting across low and high visual search environments. Finally, the third set, Experiments 5 and 6, examined the impact that an agent's degree of embodiment on performance. Additionally, this set also explored the effect of sequential presentation of problem states (low visual search complexity) versus simultaneous presentation of problem states (high visual search complexity) using computer-based worked examples. Across all of the experiments, rigorous empirical standards were applied. The following specific research questions were addressed:

In the context of a laboratory-based experiment (Experiment 1):

1. Does voice affect perceived example understanding?
2. Does voice affect perceived example difficulty?
3. Does voice affect performance on practice problems?
4. Does voice affect near transfer?
5. Does voice affect far transfer?
6. Does voice affect speaker rating?

In the context of a field-based experiment (Experiment 2):

7. Does voice affect perceived example understanding?
8. Does voice affect perceived example difficulty?
9. Does voice affect performance on practice problems?
10. Does voice affect near transfer?
11. Does voice affect far transfer?
12. Does voice affect speaker rating?

Under *low* visual search conditions (Experiment 3):

13. Does the visual presence of an animated agent foster learning more than voice alone?
14. Does the visual presence of an animated agent foster learning more than highlighting?
15. Does highlighting foster learning more than voice alone?

Under *high* visual search conditions (Experiment 4):

16. Does the visual presence of an animated agent foster learning more than voice alone?
17. Does the visual presence of an animated agent foster learning more than highlighting?
18. Does highlighting foster learning more than voice alone?

Under *low* visual search conditions (Experiment 5):

19. Does an agent's degree of embodiment affect learning?

Across both *low* and *high* visual search conditions (Experiment 6):

20. Does an agent's degree of embodiment affect learning?
21. Does visual search complexity affect learning?
22. Does an agent's degree of embodiment interact with visual search complexity?

Results

In *Experiments 1 and 2*, we obtained a voice effect in which students achieved better transfer performance when the on-screen agent spoke in a human voice than when the on-screen agent spoke in a machine synthesized voice. Importantly, learners also gave more positive ratings to the on-screen agent who spoke with a human voice rather than a machine voice on an instrument designed to capture the social characteristics of speakers.

In *Experiment 3*, the predicted advantage of the voice + agent condition over voice-only and voice + highlighting was not supported. Unlike Atkinson's (2002) published research, the present experiment did not document an *image effect* for an agent. In other words, we were not able to replicate the advantage Atkinson documented of an agent's visual presence over voice-only. There was also no evidence that the presence of agent improved learning more than highlighting. Finally, unlike Jueng et al. (1997) and Mautone and Mayer (2001), we were not able to document an effect of highlighting as a signal or cognitive aid, over voice-only. Upon reflection, we attributed the lack of differences between the conditions on the learning environment itself. In particular, we postulated that our

learning environments approach to presenting the worked examples, by successively presenting the problem states similar to an animation, contributed to creating a learning environment where the complexity of a learners' visual search can be characterized as low. According to Jeung et al. (1997), "if visual search is low then such indicators are less necessary and standard mixed mode presentations are likely to be superior to equivalent visual instructional formats" (p. 337). As a result, we elected to modify our learning environment by removing this successive presentation of problem states in an effort to increase the amount of visual search required by the learners before reexamining our three research questions.

In *Experiment 4*, after increasing the visual search complexity of our learning environment, we found partial support for our research questions. First, in a complex visual search environment, the visual presence of an image fostered learning more than voice-only. Participants assigned to the voice + agent outperformed their voice-only on practice problem solving (medium effect) and, more importantly, on far transfer (large effect). They also rated the voice more positively on one dimension (attractiveness) of the speaker rating evaluation instrument. Interesting, the voice + agent participants' dedicated significantly more time to solving the posttest items than their voice-only counterparts. Although descriptively speaking, the voice + agent participants produced higher transfer scores that their voice + highlighting peers on time, the only statistically significant difference between these two conditions was in terms of time spent solving the items on the posttest (medium-to-large effect). Finally, we found some evidence to support Jueng et al.'s (1997) documented advantage of voice + highlighting over voice-only is high visual search environments. Specifically, the participants in the voice + highlighting condition outperformed the participants in the voice-only condition in terms of practice problem-solving performance (medium effect).

Experiments 5 and 6 investigated whether various types of animated agents, designed to provide instructional elaborations during a computer tutorial involving proportional reasoning were able to increase participants' performances on learning measures. Although no differences were found between the conditions in Experiment 5 (low visual search environment), the findings from Experiment 6 indicated that students receiving instructions from a fully embodied agent outperformed their peers in no agent condition in terms of near and far transfer performance albeit the measurable effects of these difference was small.

Results from this Experiment 6 indicate that learning from worked examples is optimized when the examples are presented in subgoal-oriented fashion (i.e., simple learning environment). Determining which type of worked examples benefit student learning and understanding has direct implications for educators and instructional designers. Specifically, worked examples that are provided in textbooks and on mathematics worksheets to serve as expert models for solving mathematics problems should consist of sequentially presented problem states similar to those presented in the simple learning environment. Worked examples should be designed to encourage learners to process and encode each solution step of an example in an effort to increase the chances of recalling strategies when solving subsequent problem-solving tasks in particular domains. Although this study and numerous previous studies suggest the benefit of employing subgoal-oriented examples to model expert problem-solving steps and solutions (Catrambone, 1994, 1996, 1998; Renkl, 1997) scores of textbooks and classroom based instructional activities continue employing conventionally based examples that concurrently present an example's entire set of problem states as well as the final solution (i.e., such as the worked examples included in the complex learning environment).

Significance

On the practical side, our results from *Experiments 1 and 2* support a multimedia design principle, which can be called the *voice principle*: Designers of multimedia learning environments should create life-like on-screen agents that speak in a human voice rather than a machine-synthesized voice. The practical significance of our findings is reflected in the strong and consistent effect sizes: Across the two experiments, the effect sizes for near transfer measures were large and the effect sizes for far transfer measures were medium-to-large. Moreover, the magnitude of effects captured in the present study were comparable to the two voice effects reported by Mayer et al. (2003), namely the medium-to-large effect associated with improved transfer performance of learners exposed to native-born speaker versus one with a foreign accent and the large effect associated with enhanced transfer performance of learners exposed to a disembodied human voice versus a computer-generated voice. Importantly, we also obtained the same pattern of results in a laboratory experiment and a field experiment, suggesting the robustness of the voice effect.

Although we did not document an image effect in *Experiment 3*, the results of *Experiment 4* are consistent with an image effect, at least in terms of fostering far transfer. This replicates the Atkinson's (2002) findings under high visual search conditions. As suggested, the agent appeared to function as a visual indicator by using gesture and gaze to guide learners' attention to the relevant material. These non-verbal cues (e.g., gesture, gaze) apparently did not overburden the learners' limited cognitive resources (Sweller, 1999)—as indicated by improved learning when the agent's image was present. Perhaps the agent's use of non-verbal cues enabled the learners to dedicate their limited cognitive resources to the task of understanding the underling conceptual segments of the worked-out examples. Without the benefit of the agent's image, perhaps the voice-only participants were occupied with searching the learning environment in order to connect the audio and visual information, which prevented them from committing their restricted cognitive resources to the task of understanding the deep structure of the example at hand.

It also appears that there is a direct relationship between an agent's effectiveness and the complexity of the learning environment. Specifically, an agent's effectiveness appears to increase as the visual search complexity of the learning environment increases. Under the low visual search conditions used in Experiment 3, there was no advantage associated with an agent's image. On the other hand, under the high visual search conditions used in Experiment 4, the visual presence of an agent clearly fostered learning more than voice-only (i.e., *image effect*), particularly on far transfer (large effect).

In Experiment 4, we also found that the learners that interacted with the agent as opposed to voice + highlighting or voice-only also spent significantly more time solving the posttest items. One could argue that this provides additional evidence that animated agents assume the role of a human teacher and that the life-like characteristics and behaviors of an agent prompt the social engagement of the learner, thus allowing the learner to form a simulated human bond with the agent. In contrast, in an agent-less learning environment, a learner may identifying a computer interaction as being a case of information delivery due to the prevalence of weak social cues (e.g., disembodies voice), which leads to a failed attempt to foster an authentic social partnership with the learner. As a result, the learner does not rely on his or her sense-making processes, as in a case of social conversation, but merely attempts to learn by memorization. Due to the learner's inadequate cognitive processing (i.e., poor selection of information and ineffective organizational and integration strategies), his or her performance on subsequent tests of transfer suffered. This explanation offers one account for the advantage associated with the agent condition.

The results from *Experiments 5 and 6* contribute to the growing literature on animated pedagogical agents, worked examples as well as multimedia learning environments. Although no differences were found between the conditions in Experiment 5 (low visual search environment), the findings from Experiment 6 indicated that students receiving instructions from a fully embodied agent outperformed their peers in no agent condition in terms of near and far transfer performance albeit the measurable effects of these difference were small (evidence that supports an *image effect* of agents). The results of Experiment 6, however, replicate the results of Experiment 4, which suggests incorporating an animated agent into a computer-based learning environment enhances learning more than conditions in which agents are not included (i.e., voice-only).

The lack of a significant difference between the fully embodied condition and the minimally embodied condition in Experiment 6 is consistent with the findings of Experiment 5. The findings from the Experiment 6 in combination with those of the previous experiment suggest an *embodiment effect* in a linear computer-based environment, that is, the visual presence of an animated agent is a critical factor in optimizing learning outcomes whereas an agent's mobility is a less important factor.

Finally, Experiment 6 also empirically investigated the difference between sequentially presented worked examples (i.e., examples with sequentially presented subgoals) and simultaneous presented worked examples (i.e., examples in which the subgoals were simultaneously presented). Since the subgoal-oriented examples (i.e., low visual search environment) proved superior to simultaneous-oriented examples (i.e., high visual search environment), the current study suggests that a *sequential principle* exists, which suggests that sequentially presented subgoals are superior to simultaneously presented subgoals.

BACKGROUND

Presently, animated pedagogical agents—human-like characters that provide instruction through verbal and nonverbal modes of communication—are being used by multimedia instructional designers to create simulated human-to-human connections between learners and computers that are intended to help learners accept computers as social partners (Cassell, Sullivan, Prevost, & Churchill, 2000). The available empirical research suggests that animated pedagogical agents nested within multimedia learning environments can enhance the learners' ability to transfer what was learned to new situations as well as to increase their enjoyment of working with the learning tutorial (Atkinson, 2002; Johnson, Rickel, & Lester, 2000; Mayer, Sobko, & Mautone, 2003; Moreno, Mayer, Spires, & Lester, 2001) while not producing any split-attention effects (Craig, Gholson, & Driscoll, 2002) Despite the potential benefits of employing animated pedagogical agents to visually and aurally guide learners through computer based learning environments, almost no empirical research has examined the specific effect of (a) their voice—human or machine, (b) their advantages relative to other visual cues or indicators, such as highlighting, or (c) their degree of embodiment. Thus, the current set of experiments examine whether particular characteristics of an animated pedagogical agent can have an impact on the social connection between learners and computers, and ultimately, on the process and outcome of learning.

Social Agency Theory

One theoretical framework for considering the effectiveness and utility of fostering simulated human-to-human connections in multimedia learning environments is social agency theory (Mayer et al., 2003; Moreno et al., 2001). According to this theory, multimedia learning environments can be designed to encourage learners to operate under the assumption that their relationship with the computer is a social one, in which the conventions of human-to-human communication apply as described by Reeves and Naas (1996). Essentially, the theory posits that the use of verbal and visual social cues in computer-based environments can foster the development of a partnership by encouraging the learners to consider their interaction with the computer to be similar to what they would expect from a human-to-human conversation. For instance, the environment might rely on verbal social cues, such as a standard accented voice, or visual social cues, such as an animated agent that utilizes dynamic non-verbal signals (e.g., gaze, gesture, facial expressions), to encourage learners to approach this situation as if they are engaged in a human-to-human conversation.

Once this social partnership is established, learners can rely on several basic human-to-human social rules that guide their interaction with the multimedia learning environment (Mayer, et al., 2003). According to Grice (1975), these social rules include the cooperation principle and its four associated maxims. Specifically, Grice proposed that in human-to-human conversations, an individual listening to another person speaking will assume that he or she is making a concerted effort to make sense by being informative, accurate, relevant, and concise. Thus, the assumption of social agency theory is that learners will assume that the speaker in the multimedia learning environment—like a typical human speaker—is attempting to make sense.

According to social agency theory, priming the social interaction schema will cause the learner to try to understand and deeply process the computer's instructional message concerning academic subject matter. Mayer (1999, 2001) has posited that the cognitive processes that learners employ in order to understand an instructional message include: (a) selection of relevant information, (b) organization of

patterns of information, and (c) the integration of prior knowledge with newly presented information. The ability to process information with deep levels of understanding--that is, to engage in sense making processes--will affect whether the learner is able to transfer what was learned to related problem solving endeavors.

Additionally, social agency theory seeks to determine the conditions under which learners interpret their interaction with a computer based learning environment. Specifically, do learners perceive their computer experiences as an instance of social communication or information delivery? The difference between the two descriptions of human-computer interactions affects the learner's schema activation, levels of cognitive processing, and the quality of learning that takes place. Learners may perceive an interaction as social if they are able to receive the social cues necessary to form a simulated human-to-human conversation with the computer--cues that we posit are provided by friendly on-screen agents who speak in a human voice. Perceiving the computer as a social partner encourages the learner to engage in a sense making process that increases the probability of positive transfer (Mayer et al., 2003).

In contrast, a learner may identify a computer interaction as being a case of information delivery (Mayer et al., 2003). In this instance, the computer may incorporate weak social cues—perhaps by utilizing a computer-synthesized voice—that fail to foster an authentic social partnership with the learner. As a result, the learner does not rely on his or her sense-making processes, as in a case of social conversation, but merely attempts to learn by memorization. Due to the learners' inadequate cognitive processing (i.e., poor selection of information and ineffective organizational and integration strategies), their performance on subsequent tests of transfer will suffer.

Promoting Social Agency with Speaker's Voice

In addition to examining the role of animated agents as social cues in multimedia leaning environments, researchers have recently investigated the role of a speaker's voice as a social cue by varying the nature of the speaker's voice in a multimedia instructional program designed to convey information about lightning formation (Mayer et al., 2003). Specifically, Mayer et al. were interested in examining the relationship between the nature of a speaker's voice—whether it was socially appealing or not—and the learner's attribution of social agency. In the first of two experiments, the narration consisted of a male voice with either a standard accented speech—that is, a native speaker of standard American English—or a foreign accented speech—in this case, a non-native speaker of standard American English, one with a Russian accent. The participants in the standard accent condition scored better on a learning transfer test, which required them to solve new problems, than the participants in the foreign accent group, yielding a Cohen's d statistic of .90 (a large effect). Moreover, participants who listened to the standard accented voice rated the narrator more positively than the participants who listened to the foreign accented voice.

In a second experiment, Mayer et al. (2003) compared the social appeal of a human voice to that of a machine synthesized voice. Forty college students were randomly assigned to either a human voice (a male, native speaker of standard American English) group or a machine voice (a male, computer-generated voice) group. Results indicated the participants in the human voice group scored statistically significantly higher on learning performance tests than the machine voice group, yielding a Cohen's d statistic of .79 (a medium-to-large effect). Participants, as the research suggested, also ascribed more positive social characteristics to the human voice.

In sum, the Mayer et al. (2003) study supports several conclusions regarding the use of voice to support social agency in a multimedia learning environment. First, their research supports the prediction based on social agency theory that participants assigned to the standard accent group would outperform their peers in the foreign accented voice condition on measures of transfer and would rate the speaker's voice more positively. Second, Mayer et al.'s research supports the notion that a human voice can enhance the process and outcome of learning relative to a machine synthesized voice by providing strong social cues through the use of a familiar, socially appealing voice.

The Mayer et al. study can be criticized on the grounds that the human-machine voice effect is based on a single experiment involving an extremely short presentation (i.e., lasting approximately 2 minutes) in an artificial laboratory setting. The present study seeks to determine whether the voice effect will occur with a longer, more typical lesson in a realistic classroom setting with high school students as well as a laboratory setting with college students.

Additional evidence for the notion that human voice is associated with larger learning gains than machine synthesized voice can be inferred from a recently conducted study by Graesser and his colleagues (Graesser et al., in press). They examined the medium of presentation (i.e., text-only, voice-only, agent + voice, agent + voice + text) in the context of an intelligent tutoring system called AutoTutor designed to improve computer literacy among college students. The voice was machine synthesized using the same software for generating machine synthesized voice as we used in the present study. In contrast to the well-documented modality effect when the voice is human (Atkinson, 2002; Mayer & Moreno, 1998; Mousavi, Low, & Sweller, 1995), Graesser and his colleagues (Graesser et al., in press) found no modality effect when the voice in an intelligent tutoring system is machine synthesized. Apparently, the advantage of speech over text is lost when speech does not convey a human quality.

Promoting Social Agency with Animated Agents

In a typical educational setting, a social exchange—including verbal and nonverbal interaction—can naturally occur between a teacher and learner in conjunction with the presentation of academic material. However, when a learner is engaged in a computer-based learning episode, the opportunity for a social exchange between the learner and the learning environment is often times nonexistent (Mayer et al., 2003). Recently, Moreno et al. proposed a solution to this problem by incorporating animated pedagogical agents into multimedia learning environments in an effort to foster the development of a social relationship between learners and computers. According to the social agency theory, the combination of a multimedia learning environment and an animated agent elicits verbal and visual social cues that create virtual relationships between agents and learners as substitutes for authentic human-to-human interactions—interactions that possess the social properties employed in a human conversation. Moreover, animated agents assume the role of a human teacher giving instruction and feedback as the learner acquires and processes new information. Social agency theory stipulates that the life-like characteristics and behaviors of an animated agent prompt the social engagement of the learner, thus allowing the learner to form a simulated human bond with the agent.

Recent research focusing on the utility of animated agents has provided developers of computer based learning environments with a means of incorporating motivational and life-like characters to aid in the knowledge and skill acquisition of learners. In theory, an animated agent, with its humanistic communication capabilities, is able to direct a learner's attention to the appropriate

element of a problem-solving task using gestures, gaze, and locomotion. Moreover, multimedia learning environments incorporating animated pedagogical agents offer key features that traditional tutoring programs seem to lack. For instance, animated agents offer the potential to enrich and broaden the communicative relationship between learners and computers as well as provide computers with motivational and affective instructional features that actively engage students (Johnson, Rickel, & Lester, 2000). Additionally, simply having an animated agent present in a multimedia learning environment can positively influence the learner's perceptions of their educational experience (Lester, Converse, Kahler, Barlow, Stone & Bhogal, 1997). It has been proposed that the combination of an interesting animated agent and a well-structured learning environment can optimize a learner's active engagement with the task and increase the probability of future interactions with the instructional program (Johnson et al., 2000).

In a recent study, Atkinson (2002) examined whether the presence of an animated agent in a multimedia-based learning environment designed to teach learners how to solve word problems would enhance the process and outcome of leaning. Specifically, Atkinson examined whether the delivery method of instructional elaborations (i.e., aurally or textually) in conjunction with the presence or absence of an animated pedagogical agent had an affect on learning outcome measures. Findings indicated that the participants who were exposed to the agent in combination with narrated instructions achieved higher scores on both near and far transfer tests than the control participants who were not exposed to an animated agent (i.e., voice-only or text-only). Subsequently, Atkinson attempted to replicate the initial study by placing students in mixed (voice-plus-agent) or single (voice-only or text-only) modality conditions to receive instructional elaborations regarding mathematics word problems. Again, students receiving instructions verbally from an agent outscored their peers in the textual condition on near transfer, and outscored both the voice-only and text-only conditions in terms of far transfer performance. Presumably, an interactive relationship between a learner and a surrogate tutor was enabled by the presence of an animated agent with the capacity to narrate explanations of the instructions to the participant.

To explore whether learners will report an increased interest in learning and achieve better transfer performance if they experience a simulated human-to-human connection with the computer via an animated agent, Moreno and her colleagues (Moreno et al., 2001) conducted a series of experiments regarding the presence or absence of an animated agent in conjunction with the delivery of instructions through speech or on-screen text. Across five experiments, learners were asked to work with Design-A-Plant, a computer-based learning program in which they were expected to design a plant from a library of plant structures (e.g., roots, stems, leaves) that could thrive under specified environmental conditions. In the initial experiment, undergraduate college students who received instruction via an animated agent (i.e., Herman the Bug) scored significantly higher on complex transfer problems than did students who received the same verbal and visual instructional material without the agent. Moreover, participants in the agent group reported an increased interest in the material and a greater willingness to continue interactions with the program. Findings from additional experiments, including one with school-age children, supported the usage of an animated pedagogical agent in conjunction with spoken instruction as a tool for optimizing learning. Thus, this study capitalized on one of the chief premises of the social agency theory, that is, bringing together verbal and visual modalities of instruction with human-like features increases the likelihood that meaningful learning gains can occur through the mediation of a surrogate instructor. Consequently, Moreno et al.'s research provides evidence that a learner can capitalize on a social partnership with an on-screen animated agent, a partnership that can foster both the process and outcome of learning.

Visual Cues in Multimedia Learning

Jeung, Chandler, and Sweller (1997) conducted a study examining the impact of incorporating visual cues or indicators in a multimedia learning environment involving elementary geometry measurement. Specifically, they conducted three experiments in which they examined the use of visual indicators to direct learners' visual search under three basic conditions: (a) visual-visual, where the diagrams and associated statements were presented visually; (b) audio-visual, where the diagrams were presented visually and the associated statements were presented aurally; and (c) audio-visual-flashing, which was identical to the audio-visual condition with the exception that the relevant section of the diagram flashed when the associated statements were delivered aurally.

When high visual search material was used, they found that the learning gain attributed to the learners assigned to the audio-visual-flashing condition was significantly larger than those of the learners in the other two conditions, namely audio-visual and visual-visual. In contrast, when low visual search material was used, the learners assigned to the audio-visual-flashing and audio-visual conditions outperformed their peers in the visual-visual condition. Jeung et al. suggest that "...if visual search is likely to be high, then the inclusion of visual indicators such as flashing, color change, or simple animation is essential for audio-visual instruction to be an effective instructional teach technique [whereas] if visual search is low then such indicators are less necessary and standard mixed mode presentations are likely to be superior to equivalent visual instructional formats" (p. 337).

Using a short science lesson explaining how an airplane can achieve lift during flight, Mautone and Mayer (2001) performed a study that explored the use of visual indicators or signals as cognitive guides in a multimedia learning environment. After establishing that the signals used in text- and speech-based learning environments fostered understanding, the authors explored the use of signals in a narration-and-animation environment involving four multimedia presentations delivered via computer. In the context of multimedia messages, the authors incorporated two types of signals: (a) narration, where the salient spoken content was signaled by a shift in inflection followed by a noticeable pause, and (b) animation, where the relevant aspects of the animation was signaled by colored arrows. They found that incorporating both types of signals into a multimedia message improved learner understanding—as indicated by increased problem-solving transfer—in narration-and-animation environments.

Although their research was ostensibly about the effectiveness of animated agents in multimedia learning environments (a topic discussed later in the paper), Craig, Gholson, and Driscoll's (2002) research offers constructive insight into the use of visual cues during multimedia instruction. One of the factors they manipulated in their first experiments was the features of the pictures presented to the learners. There were three types of picture features: (a) static picture, (b) sudden onset, and (c) animation. In the static picture condition, all of the visual elements appeared simultaneously on the screen. In the sudden onset conditions, was identical to the static picture with one notable exception: the use of flashing. Essentially, as each element on the screen came under discussion by the spoken narration, the element flashed in the picture. Finally, the animation condition consisted of environment in which the salient elements were progressively added and removed during the course of the narration. According to their findings, the learners presented with either the sudden onset and animation conditions outperformed their counterparts exposed to the static pictures on a variety of measures, including a transfer test. Craig et al. conclude that it was noteworthy that "...the procedure of simply flashing appropriate parts of the pictorial information, when they were described in the

spoken narrative, was a effective as a full animation [and that] this finding may have practical implications, because creating flashing elements in a static picture can be less taxing with modern technology than creating full animations" (p. 433).

TECHNICAL APPROACH

While holding the learning content constant (i.e., proportional reasoning word problems), we conducted three sets of experiments. The first set examined the impact of the agent's voice on performance. The second set examined the compared the effectiveness of agents to highlighting across low and high visual search environments. Finally, the third set of experiments examined the impact that an agent's degree of embodiment on performance. Additionally, this set also explored the effect of sequential (low visual search complexity) versus static (high visual search complexity) worked examples.

Examining the Impact of an Animated Agent's Voice (Experiments 1 and 2)

The research on pedagogical agents supports the prediction based on social agency theory that animated on-screen agents are better able to promote social agency in multimedia leaning environments than a text-only or voice-only environment. Moreover, research on the role of a narrator's voice supports the hypothesis that the type of voice can have an impact on social agency. In the present experiments, we examined the impact of an agent's voice in a realistic mathematics lesson.

Across two experiments, participants received a narrated set of worked-out examples for proportional reasoning word problems spoken by a female native-English speaker (human voice condition) or by a female machine-synthesized voice (machine voice condition). Experiment 1 was conducted in a university-based computer laboratory with college undergraduates. Experiment 2 was conducted in a computer classroom with high school students. Both learning process and learning outcome measures were collected. The learning process measures included perceived example understanding, perceived example difficulty, and performance on practice problems. The learning outcome measures included a posttest, which contained both near and far transfer items, and a speaker-rating questionnaire designed to detect the social characteristics attributed to speakers.

Experiment 1

In this experiment, students received a computer-based mathematics lesson that provided four worked-out examples along with step-by-step descriptions of how to solve them. Narration accompanying the on-screen examples was presented in a human voice or a computer-synthesized voice (machine voice). According to social agency theory, students in the human voice group should produce higher scores than students in the machine voice group on the practice problems, the near and far transfer tests designed to measure the depth of learner understanding, and rate the speaker more positively while, at the same time, not rating the examples as any more difficult or reporting any differences in understanding than their machine voice counterparts.

Sample and Design

The participants were 50 undergraduate college students recruited from educational psychology courses at Mississippi State University. They were randomly assigned in equal numbers to one of two conditions, with 25 serving in the human voice group and 25 serving in the machine voice group. The percentage of females was 80% in the human voice group and 84% in the machine voice group; the percentage of juniors and seniors was 80% in the human voice group and 76% in the machine voice group; the percentage of students majoring in education or educational psychology was 80% in the human voice group and 72% in the machine voice group; and the mean GPA was 3.00 for the human voice group and 2.93 for the machine voice group.

Computer-Based Learning Environment

The computer-based materials consisted of two versions (i.e., human voice and machine voice) of a multimedia training program on how to solve proportional reasoning word problems. The training program was created using Director 8.0 (Macromedia, 2000) coupled with Microsoft Agent and XtrAgent 2.0 for deployment within a Windows-based operating system, and was based on an earlier program (Atkinson, 2002; Atkinson & Derry, 2000).

The learning environment, which was 800 by 600 pixels in size, included an instruction pane—for displaying the instructions for the current problem (see top left of Figure 1), a problem text pane—for displaying the problem on which the worked example was based (see middle left of Figure 1), a control panel—allowing the user to proceed through the instructional sequence at his/her own pace (see bottom left of Figure 1), a workspace—for displaying the solution to the example's problem (see right side of Figure 1), a calculator (see middle right of Figure 1), and an animated agent in the form of a parrot named Peedy—an agent capable of 75 animated behaviors, including behaviors specifically designed to direct attention to objects on the screen, such as gesturing and/or looking in specific directions (e.g., up, down, left, right). The agent was created from several off-the-shelf pieces of software, including Microsoft Agent, a collection of programmable pieces of software designed to support the presentation of the animated agent and XtrAgent 2.0, used to animate the agent within a Director-based learning environment.

The instructional materials presented in the program consisted of four example/practice problem pairs, where each worked example was followed by an isomorphic practice problem. For example, one of the worked examples was the "Bill's Hometown Furniture Store" problem:

> Bill's Hometown Furniture Store creates custom-ordered furniture. Bill, the owner, received an order for 12 identical kitchen cabinets last week. Bill hired four carpenters to work for five days, and they made 7 cabinets in that time. However, one of the carpenters broke his arm over the weekend and, as a result, will be unable to help finish the order. If Bill has the three healthy carpenters complete the remaining cabinets, how long will it take them to finish the job?

The worked examples were structured to consist of a sequential presentation of problem states and to emphasize problem subgoals. Unlike examples that simultaneously display all of the solution components (i.e., simultaneous examples), the sequential examples used in the present experiment appeared initially unsolved. The learning environment was structured to permit the learner to proceed through each example and watch as problem states were successively added over a series of pages

until the final page in the series presented the solution in its entirety. Each solution step was coupled with instructional explanations delivered orally that were designed to underscore what was occurring in that step (e.g., "First, we need to set up a proportional relationship to determine the production rate"). The examples also relied on two explicit cues—the visual isolation and labeling of each subgoal (e.g., "Total Amount 1")—to clearly demarcate a problem's subgoals.

Moreover, the learning environment was configurable to run in one of two instructional modes that reflected the two conditions of the present experiment:

Human Voice Condition - Since the animation service provided by Microsoft Agent permits audio files to be used for a character's spoken output, Peedy was programmed in the human voice condition to deliver recoded audio files consisting of instructional elaboration created by a human tutor designed to highlight what is occurring in each of the example's sequentially-presented solution step (see Figure 1). These audio files were created by a 29-year-old female graduate student who spoke with a standard North American English accent. The software automatically synchronized Peedy's mouth to the human tutor's voice by using the characteristics of the audio file.

Machine Voice Condition - The machine voice condition was identical in every respect to the human voice condition with one exception: Instead of using voice files containing a human voice to deliver instructional explanations in the examples, the Lernout & Hauspie® TruVoice TTS text-to-speech engine (http://www.microsoft.com/msagent/downloads/user.asp#tts)—a computer-based system able to read text aloud—provided by Microsoft delivered the instructional elaborations orally in North American English. Specifically, the learner in this condition listened to "Mary", a machine-generated voice based on a 30-year-old female (Model ID # c77c5170-2867-11d0-847B-44455354000) delivered, along with the presence of an agent's image, the exact same instructional explanations that were used to highlight the solution steps in the human voice condition. Regardless of which instructional mode the learning environment was configured to employ, Peedy was programmed to move around the workspace, using gesture and gaze to highlight the example's solution (see right side of Figure 1).

Following each worked example, two questions—one focused on perceived example understanding and the other addressing perceived example difficulty—were presented to the learners on the computer screen. First, they were asked to respond to the statement "I understood the worked example just presented to me" by selecting a reaction on a balanced five-point rating scale that ranged from "very much agree" (1) to "very much disagree" (5). Second, they were presented with an item adapted from an instrument used by Paas and Van Merrienboer (1993) designed to measure participants' perceived cognitive load. Specifically, they were asked "please rate the difficulty of the worked example just presented" by selecting a response on a balanced five-point rating scale that ranged from "very easy" (1) to "very difficult" (5).

After rating their understanding of the example and how difficult they perceived it to be, a practice problem was presented on the computer screen that was parallel in structure to the example itself. For example, the practice problem coupled with the "Bill's Hometown Furniture Store" problem is the following:

> A local high school needs 120 classrooms painted over the summer. They hired 5 painters who worked for six days and completed 49 classrooms. Due to a conflict with management,

however, 3 painters quit after 6 days of work. If the 2 remaining painters finish the job, how long will it take them to finish painting the classrooms?

The learner was required to enter a response to the practice problem before he or she was given the final answer to the problem. The answers to the practice problems did not include solutions to problem steps or any explanation about the solution. During the presentation of each practice problem, Peedy disappeared and only returned when the subsequent example was presented.

The computer-based environment was deployed on a total of eight Gateway E-1200 computer systems (600mhz, 256 RAM), each equipped with 15-in color monitors and Optimus Nova 80 headphones.

Pencil-Paper Materials

The paper materials consisted of a participant questionnaire, an 8-page mathematics review booklet, a 15-item speaker survey, and a posttest consisting of four near transfer items and four far transfer items. The review booklet and the transfer tests were adopted from Atkinson (2002). The 15-item speaker survey was adopted from Mayer et al. (2003). The participant questionnaire solicited information concerning the participant's demographic characteristics including gender and academic major. The mathematics review booklet provided a brief review of how to solve simple one-step proportional reasoning word problems; it included three problems that the participants were encouraged to try, followed by step-by-step descriptions of the correct solution procedure.

The speaker rating survey was a 15-item instrument adapted from Zahn and Hopper's (1985) Speech Evaluation Instrument. We adapted Zahn and Hopper's (1985) speech evaluation instrument because of its effectiveness in detecting the social characteristics attributed to speakers. Instructions at the top of the page asked the participant to circle a number from 1 to 8 indicating how the speaker sounded along each of 15 dimensions. For each dimension, the numbers 1 through 8 were printed along a line with one adjective above the "1" and an opposite adjective above the "8". The 15 adjective pairs were: literate-illiterate, unkind-kind, active-passive, intelligent-unintelligent, cold-warm, talkative-shy, uneducated-educated, friendly-unfriendly, unaggressive-aggressive, fluent-not fluent, unpleasant-pleasant, confident-unsure, inexperienced-experienced, unlikable-likeable, energetic-lazy. There were 5 items from each of 3 subscales--superiority, attractiveness, and dynamism. According to Zahn and Hopper, superiority "...combines intellectual status and competence, social status items, and speaking competency items", attractiveness captures the social and aesthetic appeal of a speaker's voice, and dynamism characterizes a speaker's "...social power, activity level, and the self-presentational aspects of [his or her] speech" (p. 119).

The near transfer items on the posttest consisted of four proportional reasoning word problems that were structurally identical to one of the problems presented during instruction albeit they had different surface stories. For example, the following near transfer item is structurally isomorphic to the "Bill's Hometown Furniture Store" problem used during instruction:

> Mike, a wheat farmer, has to plow 2100 acres. He rented six tractors with people to drive for 3.75 days, and they completed 1200 acres. If he rents four tractor/drivers, how long will it take them to complete the plowing?

The far transfer items on the posttest consisted of four proportional reasoning word problems that were not structurally identical to any of the problems presented during instruction. For example, the following far transfer problem is not isomorphic to the "Bill's Hometown Furniture Store" problem or any other problem presented during instruction:

> Brian is selling newspapers at a rate of 3 newspapers every 10 minutes on one side of a downtown street, while Sheila at her newsstand across the street is selling papers at the rate of 8 newspapers every 20 minutes. If they decide to go into business together, how many newspapers will they sell in 40 minutes at these rates?

To help control for a possible order effect, four versions of the posttest test were constructed. Within each version, the near and far transfer problems were randomly ordered.

Procedure

In this experiment, participants learned in a single session by working independently in a laboratory containing eight workstations, each located in its own cubicle. During this session, the participants filled out a demographic questionnaire, and then read through the eight-page review on solving proportion problems. When participants completed the review on proportion problems, they began the computer-based lesson in which they studied the four example/practice problem pairs. Based on random assignment, some participants received a program that had a human voice to explain the worked examples whereas others received a program that had a machine voice. Each of the four example/practice problem pairs consisted of a condition-specific worked example along with a paired isomorphic practice problem presented on the computer monitor. The learners were asked to solve the practice problem on paper and then check the accuracy of their solutions using the solution presented in learning environment. After instruction, the participants were administered the eight-item pencil-paper posttest, which took approximately fifty minutes to complete. The speaker survey was administered after the posttest, which took approximately five minutes to complete.

Scoring

The two measures collected after each example was presented—perceived example understanding and perceived example difficulty—were scored in the same fashion. The participants' responses to each of these queries were summed across all four examples and divided by four, thereby generating a measure of average perceived example understanding and a measure of average perceived example difficulty, both with values ranging from 1 to 5.

The protocols generated during practice problem solving as well as the near and far transfer tests were coded for conceptual accuracy according to a set of guidelines for analyzing the written problem-solving protocols designed to help determine where the learner fell along a problem-comprehension continuum. According to these guidelines, each item—the four practice problems, the four near transfer, and the four far transfer—was awarded a conceptual score, ranging from 0 to 3 depending upon the degree to which the participant's solution was conceptually accurate (e.g., 0 = no evidence of the student understanding the problem; 3 = there is perfect understanding of the problem, ignoring minor computational/copying errors, and the student used a complete and correct strategy to arrive at an answer). For all three measures (i.e., performance on practice problems, near transfer, and far transfer), 12 was the maximum score that a learner could achieve (e.g., 3 points-per-problem x 4 items). To create an average conceptual score, with values ranging from 0 to 3, the conceptual

scores awarded on each measure were summed across all four items and divided by 4. Internal consistency reliabilities (Cronbach's Alpha) for the practice problem, near transfer, and far transfer measures were .82, .77, and .76, respectively.

One research assistant who was unaware of the participants' treatment conditions independently scored each problem-solving protocol. To validate the scoring system, a second rater also unaware of the participants' treatment conditions independently scored a random sample of 20% of the problem-solving protocols. The scores assigned by the two raters to reflect the conceptual accuracy of the participants' responses across all three measures were consistent 96% of the time. Discussion and common consent were used to resolve any disagreement between coders.

Finally, an overall speaker rating (from 1 to 8) was constructed. This was accomplished by averaging the scores from the three subscales (i.e., superiority, attractiveness, and dynamism); with 1 indicating the most positive rating and 8 indicating the most negative rating. Internal consistency reliability (Cronbach's Alpha) for this measure was .90.

Results and Discussion

The major research question addressed in this experiment concerned whether learners in the human voice condition reported increased interest in learning and achieved better transfer than learners in the machine voice condition. Table 1 shows the mean score (and standard deviation) for each group in Experiment 1 on the perceived example understanding, perceived example difficulty, performance on practice problems, near transfer test, far transfer test, the speaker rating survey, and instructional time. Separate two-tailed t-tests were conducted on these measures, each at alpha = .05. Cohen's d statistic was used as an effect size index where d values of .2, .5, and .8 correspond to small, medium, and large values, respectively (Cohen, 1988).

Does voice affect perceived example understanding? As revealed in the first row of Table 1, the human voice group ($M = 1.32$, $SD = 0.45$) and the machine voice group ($M = 1.31$, $SD = 0.35$) did not statistically differ in terms of perceived example understanding, $t(48) = 0.87$, p = ns. In sum, the results show that learners reported that the examples were relatively easy to understand, regardless of which voice—human or machine—accompanied the examples.

Does voice affect perceived example difficulty? As illustrated in the second row of Table 1, the perceived example difficulty (i.e., cognitive load) reported by the participants assigned to the human voice condition ($M = 2.21$ $SD = 0.75$) did not differ significantly from the mean ratings of the machine voice group ($M = 1.99$, $SD = 0.58$), $t(48) = 1.16$, p = ns. In general, the results reveal that participants presented with either the human voice or the machine voice perceived the examples to be moderately difficult, with no statistically significant difference between the two conditions.

Does voice affect performance on practice problems? As shown in the third row of Table 1, the scores associated with the practice problems for participants in the human voice group ($M = 2.67$, $SD = .63$) were significantly higher than those of their peers in the machine voice group ($M = 2.09$, $SD = .85$) on, $t(48) = 2.74$, $p < .01$. Cohen's d statistic for these data yields an effect size estimate of 0.79 for practice problem-solving performance, which corresponds to a medium-to-large effect. Overall, the results show that human voice fostered better understanding of how to solve the practice problems that accompanied the examples during instruction than did machine voice.

Does voice affect near transfer? As shown in the fourth row of Table 1, the human voice group (M = 2.23, SD = .71) scored significantly higher than the machine voice group (M = 1.62, SD = ..77) on the near transfer test, $t(48) = 2.91$, $p < .01$. Cohen's d statistic for these data yields an effect size estimate of 0.84 for near transfer, which corresponds to a large effect. In general, the results show that human voice fostered better understanding of how to solve problems like those presented during instruction than did machine voice.

Does voice affect far transfer? As shown in the fifth row of Table 1, the human voice group (M = 1.32, SD = .90) scored significantly higher than the machine voice group (M = .77, SD = .67) on the far transfer test, $t(48) = 2.46$, $p < .05$. Cohen's d statistic for these data yields an effect size estimate of 0.71 for far transfer, which corresponds to a medium-to-large effect. Overall, the results show that human voice fostered deeper understanding of how to solve problems that were not like those presented during learning than did machine voice.

Does voice affect speaker rating? As shown in the sixth row of Table 1, the human voice group (M = 2.29, SD = .84) rated the speaker significantly more favorably than did the machine voice group (M = 3.10, SD = 1.30) on the speaker rating survey, $t(48) = 2.64$, $p = .01$. Cohen's d statistic for these data yields an effect size estimate of 0.76 for speaker rating, which corresponds to a medium-to-large effect. Overall, students in the human voice condition reported a more positive evaluation of the speaker's attractiveness, dynamism, and superiority than did the machine voice group.

Related issues. To determine whether the results could be attributed to differences in the intelligibility of the human and machine voices, we conducted a supplemental study in which participants listened to a word problem spoken in machine voice and wrote down the words and also listened to a question spoken in human voice and wrote down the words. Specifically, ten undergraduate students (3 males, 7 females; average GPA = 2.98; 6 educational psychology majors, 4 education majors) were presented with the first two worked-out examples of the previously described computer-based learning environment. Using a counterbalanced procedure, the participants listened to the narration accompanying an on-screen worked example spoken in machine voice and wrote down the words and also listened to an example spoken in human voice and wrote down the words. One worked example consisted of 255 words while the other consisted of 192 words. The participants were not expected to solve the accompanying practice problems. The participants correctly recorded an average of 94.5% (SD = 5.9%) of the example's narration when it was a human voice and an average of 93.4% (SD = 5.0%) of the example's narration when it was a machine voice. According to a paired-sample t-test (alpha = .05), the percentage of words recorded by the participants did not statistically differ across examples, $t(9) = .34$, p = ns. Thus, the pattern of results cannot be attributed to the human voice being substantially easier to discern than the machine voice.

Moreover, the pattern of results cannot be attributed to the human voice group spending more time during learning than the machine voice group, because we found that the machine voice group averaged 40.4 minutes (SD = 17.1) on the instructional program whereas the human voice group averaged 39.2 minutes (SD = 9.2).

To explore the possibility that the pattern of results could be attributed to a "novelty effect", we examined the performance on the practice problems in the first half versus the second half of training. According to a novelty effect, there should be large differences between machine and human voices for early practice problems but not later practice problems. That is, at the outset of instruction, the learners might be distracted by the machine voice to the point of decreasing attention

to the content (resulting in lower practice problem performance) before adjusting to the machine voice, and thereafter the machine and human voices become equivalent. To test for this possibility, we calculated an average score on the first two and on the last two practice problems for each group. The averages associated with the first two practice problems for participants in the human voice group ($M = 2.66$, $SD = .67$) were significantly higher than those of their peers in machine voice group ($M = 2.10$, $SD = .97$) on, $t(48) = 2.38$, $p < .05$ (Cohen's $d = .67$). The same pattern emerged for the last two practice problems where the participants in the human voice group ($M = 2.68$, $SD = .64$) were significantly higher than those of their peers in machine voice group ($M = 2.08$, $SD = .85$) on, $t(48) = 2.81$, $p < .01$ (Cohen's $d = .81$). Thus, the observed differences do not appear to result from a novelty effect.

In summary, the human voice condition produced statistically and practically significant differences in terms of practice problem solving (medium-to-large effect), near transfer (large effect), and far transfer (medium-to-large effect) as well as perception of the speaker's voice (medium-to-large effect). Interestingly, despite these performance differences, there did not appear to be any difference in perceived example understanding or perceived example difficulty. Moreover, we conclude that the observed differences should not be attributed to variation in the intelligibility of the voices, time on task, or novelty effect.

Experiment 2

In an effort to replicate and extend these findings, we decided to conduct a small-scale field experiment at an area high school with students enrolled in one of several sections of the same college-preparatory mathematics courses. Furthermore, to help ensure the authenticity of the task, the experiment was run in the computer laboratory at the high school with the entire class—including the instructor—present at the time of the experiment. As with the previous experiment, we hypothesized that students in the human voice group should produce higher scores than students in the machine voice group on the practice problems, the near and far transfer tests designed to measure the depth of learner understanding, and rate the speaker more positively while, at the same time, not rating the examples as any more difficult or reporting any differences in understanding than their machine voice counterparts.

Sample and Design

The participants were 40 high school students recruited from several mathematics courses taught by the same instructor at Starkville High School (in Starkville, Mississippi). They were randomly assigned to condition, with 20 serving in the human voice group and 20 serving in the machine voice group. The percentage of males was 75% in the human voice group and 25% in the machine voice group; the percentage of juniors and seniors was 70% in the human voice group and 55% in the machine voice group; and the mean GPA was 3.55 for the human voice group and 3.58 for the machine voice group.

Computer-Based Learning Environment

The computer-based learning environment was identical to Experiment 1. The apparatus consisted of 25 PC computer systems (750mhz, 256 RAM) with 15-in color monitors and headphones.

Pencil-Paper Materials

The pencil-paper materials were identical to Experiment 1.

Procedure

The procedure was similar to Experiment 1. Instead of arriving individually in a laboratory equipped with 8 workstations, in the present experiment, an intact class arrived during two consecutive class periods at a lab containing 25 workstations (25 PC computer systems with 750mhz, 256 RAM, 15-in color monitors, and headphones) where they were asked to locate a workstation where they would work independently. All other aspects of the procedure were identical to Experiment 1.

Scoring

The scoring was identical to Experiment 1. As with the previous experiment, a research assistant who was unaware of the participants' treatment conditions independently scored each problem-solving protocol while a second rater independently scored a random sample of 20% of the protocols. They agreed on scoring 98% of the time. Discussion and consensus were used to resolve any disagreement between raters. Internal consistency reliabilities (Cronbach's Alpha) for the practice problem, near transfer, and far transfer measures were .74, .80, and .79, respectively. Internal consistency reliability for the speaker rating survey was .87.

Results and Discussion

The major research question addressed in this experiment concerned whether the results from Experiment 1, in which learners in the human voice reported increased interest in learning and achieved better transfer performance than learners in the machine voice condition, could be replicated with high school students. Table 2 shows the mean score (and standard deviation) for each group in Experiment 2 on the perceived example understanding, perceived example difficulty, performance on practice problems, near transfer test, far transfer test, speaker rating survey, and instructional time. Separate two-tailed t-tests were conducted on these measures, each at alpha = .05.

Does voice affect perceived example understanding? As indicated in the first row of Table 2, there was no significant difference in perceived example understanding between the participants assigned to the human voice condition ($M = 1.65$, $SD = 0.45$) and those assigned to the machine voice condition ($M = 1.51$, $SD = 0.43$), $t(38) = 0.98$, p = ns. In general, across both voice conditions, the results indicate that learners reported that the examples were relatively easy to understand.

Does voice affect perceived example difficulty? As illustrated in the second row of Table 2, the human voice group ($M = 2.44$, $SD = 0.75$) and the machine voice group ($M = 2.40$, $SD = 0.79$) reported similar levels of perceived example difficulty, $t(38) = 0.15$, p = ns. In sum, there was no difference in the perceived difficulty of the examples across conditions.

Does voice affect performance on practice problems? As shown in the third row of Table 2, the human voice group ($M = 2.33$, $SD = .64$) scored significantly higher than the machine voice group ($M = 1.80$, $SD = .86$) on solving practice problems, $t(38) = 2.20$, $p < .05$. Cohen's d statistic for these data yields an effect size estimate of 0.63, which corresponds to a medium effect. In general, the

results show that human voice fostered better understanding of how to solve the practice problems than did machine voice.

Does voice affect near transfer? As shown in the fourth row of Table 2, the human voice group (M = 2.51, SD = .59) scored significantly higher than the machine voice group (M = 1.84, SD = .86) on the near transfer test, $t(38) = 2.89$, $p < .01$. Cohen's d statistic for these data yields an effect size estimate of 0.83 for near transfer, which corresponds to a large effect. As with Experiment 1, the results demonstrate that human voice fostered better understanding of how to solve problems like those presented during instruction than did machine voice.

Does voice affect far transfer? As shown in the fifth row of Table 2, the human voice group (M = 1.74, SD = .70) scored significantly higher than the machine voice group (M = 1.15, SD = .82) on the far transfer test, $t(38) = 2.42$, $p < .05$. Cohen's d statistic for these data yields an effect size estimate of 0.70 for far transfer, which corresponds to a medium-to-large effect. As with Experiment 1, the results demonstrate that human voice fostered increased understanding of how to solve problems that were not like those presented during learning than did machine voice.

Does voice affect speaker rating? As shown in the sixth row Table 2, the human voice group (M = 3.19, SD = 1.05) rated the speaker significantly more favorably than did the machine voice group (M = 4.23, SD = 1.30) on the speaker rating test, $t(38) = 2.78$, $p = .008$. Cohen's d statistic for these data yields an effect size estimate of 0.83 for speaker rating, which corresponds to a large effect. As with Experiment 1, students in the human voice condition reported a more positive evaluation of the speaker's attractiveness, dynamism, and superiority than did the machine voice group.

Related issues. As with Experiment 1, the pattern of results cannot be attributed to the human voice group spending more time during learning than the machine voice group, because we found that the machine voice group averaged 42.1 m (SD = 10.3) on the instructional program whereas the human voice group averaged 40.7 m (SD = 16.4).

In sum, the human voice condition once again produced statistically and practically significant difference in terms of practice problem solving (medium effect), near transfer (large effect), and far transfer (medium-to-large effect) as well as perception of the speaker's voice (large effect). In spite of these performance differences, there did not appear to be any difference in perceived example understanding or perceived example difficulty. This finding represents a slight departure from the Experiment 1 where there did not appear to be any difference in perceived example understanding as opposed to perceived example difficulty. Finally, as with Experiment 1, these differences could not be attributed to time on task.

Conclusions Regarding the Impact of an Animated Agent's Voice

In our computer-based learning environment designed to teach mathematics, we attempted to foster a sense of social presence in which learners would be more likely to interpret the computer-based narrator as a social partner. Overall, across two different experiments, we obtained a voice effect in which students achieved better transfer performance when the on-screen agent spoke in a human voice than when the on-screen agent spoke in a machine synthesized voice. Importantly, learners also gave more positive ratings to the on-screen agent who spoke with a human voice rather than a machine voice on an instrument designed to capture the social characteristics of speakers.

Overall, this study provides an important extension of preliminary research conducted by Mayer and his colleagues (Mayer et al., 2003) by (a) using a new learning environment (i.e., one that relies on an animated agent as the source of verbal support rather than a voice over), (b) teaching a new type of material (i.e., procedural knowledge rather than conceptual knowledge), (c) incorporating a new domain (i.e., math rather than science), (d) using a new population (i.e., high school students rather than college students), (e) extending the length of the instructional episode (i.e., 40 min rather than 2 min), (f) incorporating new dependent measures (i.e., performance on practice problems, near and far transfer problem solving, ratings, and instructional time), (g) relying on a new independent variable (i.e., the voices used are different), and most importantly, (h) using an authentic educational context (in Experiment 2 of the present study) rather than a lab setting. Thus, the present study shows that voice effects occur across at least two settings including an authentic classroom environment.

Implications

On the theoretical side, the results are consistent with social agency theory, which posits that social cues in multimedia messages can encourage learners to interpret human-computer interactions as similar to human-to-human conversation. Although the results are tentative, we found little evidence that our attempts to promote social agency (by using a human voice) increased cognitive load—that is, there were no differences between the two voice conditions in terms of perceived example understanding or in terms of perceived example difficulty (i.e., cognitive load). In particular, the learners who received the human voice showed substantial advantages in solving practice problems during instruction and on solving near and far transfer problems after instruction, as well as reporting a more positive rating of the on-screen agent in terms of at least two speaker dimensions, namely attractiveness and dynamism. The voice effects we found in these two experiments replicate and extend the voice effect reported by Mayer et al. (2003) by employing more differentiated dependent measures and new instructional materials, and by showing that the same effects occur in both laboratory and school settings.

On the practical side, our results support a multimedia design principle, which can be called the *voice principle*: Designers of multimedia learning environments should create life-like on-screen agents that speak in a human voice rather than a machine-synthesized voice. The practical significance of our findings is reflected in the strong and consistent effect sizes: Across the two experiments, the effect sizes for near transfer measures were large and the effect sizes for far transfer measures were medium-to-large. Moreover, the magnitude of effects captured in the present study were comparable to the two voice effects reported by Mayer et al. (2003), namely the medium-to-large effect associated with improved transfer performance of learners exposed to native-born speaker versus one with a foreign accent and the large effect associated with enhanced transfer performance of learners exposed to a disembodied human voice versus a computer-generated voice. Importantly, we also obtained the same pattern of results in a laboratory experiment and a field experiment, suggesting the robustness of the voice effect.

Future Directions

First, it appears worthwhile to examine cognitive load more closely, by incorporating other measures of mental effort beyond the item used in the present study (i.e., perceived example difficulty) adapted from an instrument developed by Paas and Van Merrienboer (1993) to measure participants' perceived cognitive load. One promising lead is an assessment of cognitive load using a dual task methodology (Brünken, Plass, & Leutner, 2003; Brünken, Steinbacher, Plass, & Leutner, 2002).

Second, as machine synthesized voices improve, it would be important to test their effectiveness to see if they can close the performance gap that this study highlights between human voice and machine voice. In addition, it would be worthwhile to examine whether the effects are diminished if learners receive more practice with the machine voice.

Third, the present study relies on indirect measures of the degree to which students experience a social relation with the agent. Future research could explore the creation of more direct measures of this phenomenon, such as the level of facial expression or gesture they display during learning.

Comparing Agents to Highlighting across Low and High Visual Search Environments (Experiments 3 and 4)

As previously mentioned, the research on pedagogical agents supports the prediction based on social agency theory that animated on-screen agents are better able to promote social agency in multimedia leaning environments than a text-only or voice-only environment. Specifically, at least two studies support the use of agents capable of speech over text-based environments (Atkinson, 2002; Moreno et al., 2001). There is also a modicum of evidence that visual presence of an agent can foster social agency beyond a voice-only environments, at least in terms of far transfer performance (Atkinson, 2002). This latter effect, however, has not been replicated in any other published experiment. Thus, one purpose of this set of experiments is to attempt to reproduce this effect, that is, the superiority of agent + voice over voice-only.

Although social agency theory is one possible theoretical framework one can use to account for the superiority of deploying an agent capable of speech compared to equivalent voice-only environments, it is also plausible to suggest that an agent is simply functioning as a visual indicator, signal, or cognitive guide (Jueng et al., 1997; Mautone & Mayer, 2001). According to the research by Jueng et al. (1997) and Mautone and Mayer (2001), incorporating a simple animation (e.g., flashing, colored arrows) that is coordinated with the voice over to direct attention to the relevant aspects of the screen, like the visual presence of an agent, effectively fosters transfer. In other words, perhaps the agent functions as a visual indicator—akin to the electronic flashing employed by Jeung et al. (1997)—by using gesture and gaze to guide learners' attention to the relevant material. Thus, it seems that one potential explanation for an agents' effectiveness is that it functions as a visual indicator. One way to test whether this is the case is to compare the effectiveness of presenting learners with agent + voice versus signal + voice.

In sum, we were interested in examining the impact of an agent's image versus highlighting as a visual cue in a realistic mathematics lesson. Across two experiments (Experiments 3 and 4), participants received a narrated set of worked-out examples for proportional reasoning word problems spoken by a female native-English speaker in one of three conditions: voice + agent, voice + highlighting or voice-only. In Experiment 3, the three conditions were presented in a *low* visual search environment (i.e., sequential presentation or problem states), one in which the examples were unfolded one subgoal at a time. In Experiment 4, the three conditions were presented in a *high* visual search environment (i.e., static presentation of problem states), one in which the entire solution was presented at the onset of the worked example as opposed to unfolding over time. As with the previous experiments (Experiment 1 and 2), both learning process and learning outcome measures were collected. The learning process measures included perceived example understanding, perceived example difficulty, and performance on practice problems. The learning outcome measures

included a posttest, which contained both near and far transfer items, and a speaker-rating questionnaire designed to detect the social characteristics attributed to speakers.

Experiment 3

Experiment 3 was designed to address three questions. Specifically, under *low* visual search conditions: (a) does the visual presence of an animated agent foster learning more than voice alone; (b) does the visual presence of an animated agent foster learning more than highlighting; and (c) does highlighting foster learning more than voice alone?

According to social agency theory, students in the voice + agent condition should produce higher scores than students in the voice-only and voice + highlighting group on the practice problems, the near and far transfer tests designed to measure the depth of learner understanding, and rate the speaker more positively while, at the same time, not rating the examples as any more difficult or reporting any differences in understanding than their voice-only and voice + highlighting counterparts.

Sample and Design

Seventy-five undergraduate students (2 Freshman, 15 Sophomores, 24 Juniors, and 34 Seniors) from the educational psychology and psychology departments at a large, southeastern university volunteered to participate in the study. The sample consisted of 15 males and 60 females (mean GPA = 3.08, mean, ACT = 20.87). The participants were randomly assigned in equal proportions ($n = 25$) to one of the three conditions: voice + agent, voice + highlighting, or voice-only.

Computer-Based Learning Environment

The learning environment used in Experiments 1 and 2 was modified to accommodate the present experiment. Essentially, the human voice condition was used as the foundation for all three conditions in this experiment. As with previous experiments, the worked examples provided in this learning environment consisted of the sequential presentation of problem states in order to highlight problem subgoals—which we characterize as a *low* visual search condition for purposes of this experiment. Specifically, the sequential presentations were presented as follows: Initially the examples appeared unsolved. Then the learner proceeded through each example while the problem states were gradually added on the screen until the example was presented in its entirety. This type of worked example focuses the student's attention on the practice of creating a solution to the problem. This practice allows students to study each component of the example's solution in isolation from the one preceding it, because learners can progress through each example, examining each problem state and the transformation required to accomplish the following state. For each example, a control panel was provided thus allowing learners to move throughout each example at their own pace. Throughout each solution step, instructional elaborations were orally provided to highlight the activity in each solution step (i.e., "First, we need to set up a proportional relationship to determine the cost of the travel package without the discount"). The subgoals nested within each example were labeled (i.e., "Initial Amount") in order to distinguish the problem's subgoals from one another.

Moreover, the learning environment was configurable to run in one of three instructional modes that reflected the three conditions of the present experiment:

Voice-Only Condition - In the voice-only condition (see Figure 2), learners listened to a human tutor's voice reading the textual explanations designed to highlight what was occurring in that step (e.g., "Second, we need to set up another proportional relationship to determine the production time").

Voice + Highlighting Condition - The voice + highlighting condition was indistinguishable from the voice-only condition with one notable exception: the presence of a box highlighting the portion of the problem under discussion (see Figure 3). In each worked example, a bright flashing box enclosed each newly introduced subgoal. At the onset of each subgoal, the box flashed once as it outlined the problem state then remained present during the aural instructions that corresponded with the subgoal. Once the instructional elaborations related to the subgoal concluded, the highlighting box disappeared and only returned during the presentation of the subsequent subgoal. The function of the highlighting box was identical to that of the animated agent: direct learner attention to the appropriate problem state of the worked example.

Voice + Agent Condition - The voice + agent condition was also identical to the voice-only condition with one notable exception: the presence of an agent. In this condition, an animated agent maintained a visual presence throughout instruction while explanations—the same explanations found in the voice-only and voice + highlighting conditions—were delivered aurally (see Figure 1). Additionally, the agent integrated aural information (i.e., instructional elaborations) with visual information (i.e., solution steps) by using nonverbal modes of communication throughout the instruction to encourage learners to attend to the current problem state. For instance, in Figure 2, the agent is gesturing and glancing toward an example's solution step while using a word balloon to deliver the instruction explanation ("So, the travel package for John's group will cost $18,947.09."). This condition was identical to the human voice condition used in Experiments 1 and 2.

Pencil-Paper Materials

The pencil-paper materials were identical to Experiment 1.

Procedure

The procedure was identical to Experiment 1.

Scoring

The scoring was similar to Experiment 1 with one notable exception: Instead of a single scoring representing the overall speaker rating, three scores corresponding to the three subscales of the speaker rating survey were constructed. This was accomplished by averaging the scores within each of the three subscales (i.e., superiority, attractiveness, and dynamism) with 1 indicating the most positive rating and 8 indicating the most negative rating.

Results and Discussion

Table 3 presents the means scores and standard deviations for each group on each of the dependent measures. An analysis of variance (ANOVA) was conducted on each learning process measure and performance measure (alpha = .05). Significant main effects were followed up with Fisher's LSD test, based on a familywise alpha of .05 (Kirk, 1995).

Does the visual presence of an animated agent foster learning more than voice alone? On the learning process measures, there was no significant differences between the voice + agent and voice-only conditions on practice problem-solving performance, $F(2, 74) = .07$, $MSE = .37$, $p > .05$, perceived example understanding, $F(2, 74) = .10$, $MSE = .35$, $p > .05$, perceived example difficulty, $F(2, 74) = 1.39$, $MSE = .67$, $p > .05$, or instructional time, $F(2, 74) = .29$, $MSE = 114.56$, $p > .05$.

Similarly, on the learning outcome measures, there was no significant differences between the voice + agent and voice-only conditions on near transfer, $F(2, 74) = .79$, $MSE = .37$, $p > .05$, far transfer, $F(2, 74) = .01$, $MSE = .07$, $p > .05$, superiority, $F(2, 74) = 1.39$, $MSE = .67$, $p > .05$, attractiveness, $F(2, 74) = 1.39$, $MSE = .67$, $p > .05$, dynamism, $F(2, 74) = 1.39$, $MSE = .67$, $p > .05$, or time to completion of posttest, $F(2, 74) = .29$, $MSE = 114.56$, $p > .05$.

Does the visual presence of an animated agent foster learning more than highlighting? On both the learning process and learning outcome measures, there were no significant differences between voice + agent and voice + highlighting conditions under low visual search conditions.

Does highlighting foster learning more than voice alone? On both the learning process and learning outcome measures, there were no significant differences between voice + highlighting and voice-only conditions under low visual search conditions.

Overall, the predicted advantage of the voice + agent condition over voice-only and voice + highlighting was not supported. Unlike Atkinson's (2002) published research, the present experiment did not document an *image effect* for an agent. In other words, we were not able to replicate the advantage Atkinson documented of an agent's visual presence over voice-only. There was also no evidence that the presence of agent improved learning more than highlighting. Finally, unlike Jueng et al. (1997) and Mautone and Mayer (2001), we were not able to document an effect of highlighting as a signal or cognitive aid, over voice-only.

Upon reflection, we attributed the lack of differences between the conditions on the learning environment itself. In particular, we postulated that our learning environments approach to presenting the worked examples, by successively presenting the problem states similar to an animation, contributed to creating a learning environment where the complexity of a learners' visual search can be characterized as low. According to Jeung et al. (1997), "if visual search is low then such indicators are less necessary and standard mixed mode presentations are likely to be superior to equivalent visual instructional formats" (p. 337). As a result, we elected to modify our learning environment by removing this successive presentation of problem states in an effort to increase the amount of visual search required by the learners before reexamining our three research questions.

Experiment 4

An open question is whether we would find differences between the conditions used in Experiment 3 (voice + agent, voice + highlighting, and voice-only) using *high* visual search material. There is some empirical evidence to support this contention. As noted previously, Jueng et al. (1997) found that when high visual search material was used, learners assigned to the audio-visual-flashing condition demonstrated significantly larger learning gains than those of the learners in the other two conditions, namely audio-visual and visual-visual. On the other hand, when low visual search material was used, the learners assigned to the audio-visual-flashing and audio-visual conditions outperformed their peers in the visual-visual condition. They concluded that "...if visual search is likely to be high, then the inclusion of visual indicators such as flashing, color change, or simple animation is essential for audio-visual instruction to be an effective instructional teach technique."

Similar to the previous experiment, Experiment 4 was designed to address three questions. Specifically, under *high* visual search conditions: (a) does the visual presence of an animated agent foster learning more than voice alone; (b) does the visual presence of an animated agent foster learning more than highlighting; and (c) does highlighting foster learning more than voice alone?

Once again, according to social agency theory, we predicted that students in the voice + agent condition should produce higher scores than students in the voice-only and voice + highlighting group on the practice problems, the near and far transfer tests designed to measure the depth of learner understanding, and rate the speaker more positively while, at the same time, not rating the examples as any more difficult or reporting any differences in understanding than their voice-only and voice + highlighting counterparts. We also predicted that voice + highlighting will outperform their voice-only counterparts on the learning process and learning outcome measures.

Sample and Design

Seventy-eight undergraduate students (5 Sophomores, 29 Juniors, and 48 Seniors) from the educational psychology and psychology departments at a large, southeastern university volunteered to participate in the study. The sample consisted of 16 males and 59 females (mean GPA = 3.04, mean, ACT = 20.94). The participants were randomly assigned in equal proportions (n = 26) to one of the three conditions: voice + agent, voice + highlighting, or voice-only.

Computer-Based Learning Environment

The learning environment used in Experiments 3 was modified to accommodate the present experiment. Essentially, the only difference between the learning environments, including the three instructional modes, was the manner in which the examples were presented. In other words, the worked examples presented in the *high* visual search environment were identical to those presented in the simple learning environment with one notable exception: the problem states were presented simultaneously. That is, the worked examples simultaneously displayed all of the solution components in their entirety. Identical to the simple learning environment, instructional elaborations were orally provided to emphasize the activity in each solution step. Additionally, the subgoals were labeled in order to distinguish the problem's subgoals from one another.

As with Experiment 3, the learning environment was configurable to run in one of three instructional modes that reflected the three conditions of the present experiment:

Voice-Only Condition - In the voice-only condition (see Figure 4), learners listened to a human tutor's voice reading the textual explanations designed to highlight what was occurring in that step (e.g., "Second, we need to set up another proportional relationship to determine the production time").

Voice + Highlighting Condition - The voice + highlighting condition was indistinguishable from the voice-only condition with one notable exception: the presence of a box highlighting the portion of the problem under discussion (see Figure 5). In each worked example, a bright flashing box enclosed each newly introduced subgoal. At the onset of each subgoal, the box flashed once as it outlined the problem state then remained present during the aural instructions that corresponded with the subgoal. Once the instructional elaborations related to the subgoal concluded, the highlighting box disappeared and only returned during the presentation of the subsequent subgoal. The function of the highlighting box was identical to that of the animated agent: direct learner attention to the appropriate problem state of the worked example.

Voice + Agent Condition - The voice + agent condition was also identical to the voice-only condition with one notable exception: the presence of an agent. In this condition, an animated agent maintained a visual presence throughout instruction while explanations—the same explanations found in the voice-only and voice + highlighting conditions—were delivered aurally (see Figure 6). Additionally, the agent integrated aural information (i.e., instructional elaborations) with visual information (i.e., solution steps) by using nonverbal modes of communication throughout the instruction to encourage learners to attend to the current problem state. For instance, in Figure 2, the agent is gesturing and glancing toward an example's solution step while using a word balloon to deliver the instruction explanation ("So, the travel package for John's group will cost $18,947.09.").

Pencil-Paper Materials

The pencil-paper materials were identical to Experiment 1.

Procedure

The procedure was identical to Experiment 1.

Scoring

The scoring was identical to Experiment 3.

Results and Discussion

Table 4 presents the means scores and standard deviations for each group on each of the dependent measures. An analysis of variance (ANOVA) was conducted on each learning process measure and performance measure (alpha = .05). Significant main effects were followed up with Fisher's LSD test, based on a familywise alpha of .05 (Kirk, 1995). Cohen's d statistic was used as an effect size

index where d values of .2, .5, and .8 correspond to small, medium, and large values, respectively (Cohen, 1988).

Does the visual presence of an animated agent foster learning more than voice alone? There was a significant main effect for condition on practice problem-solving performance, $F(2, 75) = 3.71$, $MSE = 1.09$, $p < .05$. According to Fisher's LSD test, participants in the voice + agent condition outperformed the participants in the voice-only condition. For this measure, Cohen's d statistic for pairwise comparison between voice + agent and voice-only conditions yields an effect size estimate of .62 (medium effect). With regard to the other learning process measures, there was no significant main effects on perceived example understanding, $F(2, 75) = .51$, $MSE = .77$, $p > .05$, perceived example difficulty, $F(2, 75) = 1.41$, $MSE = .62$, $p > .05$, or instructional time, $F(2, 75) = .69$, $MSE = 106.26$, $p > .05$.

Although there was no significant main effect on near transfer, $F(2, 75) = .31$, $MSE = .11$, $p > .05$, there was a significant main effect for condition on far transfer, $F(2, 75) = 3.87$, $MSE = .09$, $p < .05$. According to Fisher's LSD test, participants in the voice + agent condition outperformed their counterparts in the voice-only condition. Cohen's d statistic for pairwise comparison yields an effect size estimate of .80 for far transfer, which corresponds to a large effect.

Moreover, the was a significant main effect for condition on the attractiveness dimension of the speaker rating scale, $F(2, 75) = 3.20$, $MSE = 2.00$, $p < .05$. According to Fisher's LSD test, participants in the voice + agent condition outperformed their counterparts in the voice-only condition. Cohen's d statistic for pairwise comparison yields an effect size estimate of .71 for far transfer, which corresponds to a medium-to-large effect. There was no significant main effect for condition on the superiority dimension, $F(2, 75) = .03$, $MSE = 1.56$, $p > .05$, or dynamism dimension, $F(2, 75) = .33$, $MSE = 1.62$, $p > .05$, of the speaker rating survey.

Finally, there was a significant main effect for condition in time spent on posttest, $F(2, 75) = 3.87$, $MSE = .09$, $p < .05$. According to Fisher's LSD test, participants in the voice + agent condition spent significantly more time than their counterparts in the voice-only condition solving the problems—both near and far—on the posttest. Cohen's d statistic for pairwise comparison between voice + agent and voice-only yields an effect size estimate of 1.02 for time on test, which corresponds to a large effect.

Does the visual presence of an animated agent foster learning more than highlighting? Across all of the learning process and learning outcome measures, there was only one significant omnibus test with implications for this research question, namely the ANOVA used to analyze time spent on posttest, $F(2, 75) = 3.87$, $MSE = .09$, $p < .05$. According to Fisher's LSD test, participants in the voice + agent condition spent significantly more time than their counterparts in the voice + highlighting condition solving the problems—both near and far—on the posttest. Cohen's d statistic for pairwise comparison between voice + agent and voice + highlighting yields an effect size estimate of .68 for time on test, which corresponds to a medium-to-large effect.

Does highlighting foster learning more than voice alone? Across all of the learning process and learning outcome measures, there was only one significant omnibus test with implications for this research question, namely the ANOVA used to analyze practice problem-solving performance, $F(2, 75) = 3.71$, $MSE = 1.09$, $p < .05$. According to Fisher's LSD test, participants in the voice + highlighting condition outperformed the participants in the voice-only condition. For this measure,

Cohen's d statistic for pairwise comparison between voice + highlighting and voice-only conditions yields an effect size estimate of .65 (medium effect).

Conclusions Regarding the Comparison of Agents to Highlighting across Low and High Visual Search Environments

After increasing the visual search complexity of our learning environment, we found partial support for each of the research questions addressed across these two experiments. First, in a complex visual search environment, the visual presence of an image fosters learning more than voice-only. Participants assigned to the voice + agent outperformed their voice-only on practice problem solving (medium effect) and, more importantly, on far transfer (large effect). They also rated the voice more positively on one dimension (attractiveness) of the speaker rating evaluation instrument. Interesting, the voice + agent participants' dedicated significantly more time to solving the posttest items than their voice-only counterparts. Although descriptively speaking, the voice + agent participants produced higher transfer scores that their voice + highlighting peers on time, the only statistically significant difference between these two conditions was in terms of time spent solving the items on the posttest (medium-to-large effect). Finally, we found some evidence to support Jueng et al.'s (1997) documented advantage of voice + highlighting over voice-only is high visual search environments. Specifically, the participants in the voice + highlighting condition outperformed the participants in the voice-only condition in terms of practice problem-solving performance (medium effect).

Implications

The results of Experiment 4 are consistent with an image effect, at least in terms of fostering far transfer. This replicates the Atkinson's (2002) findings under high visual search conditions. As suggested, the agent appeared to function as a visual indicator by using gesture and gaze to guide learners' attention to the relevant material. These non-verbal cues (e.g., gesture, gaze) apparently did not overburden the learners' limited cognitive resources (Sweller, 1999)—as indicated by improved learning when the agent's image was present. Perhaps the agent's use of non-verbal cues enabled the learners to dedicate their limited cognitive resources to the task of understanding the underling conceptual segments of the worked-out examples. Without the benefit of the agent's image, perhaps the voice-only participants were occupied with searching the learning environment in order to connect the audio and visual information, which prevented them from committing their restricted cognitive resources to the task of understanding the deep structure of the example at hand.

It also appears that there is a direct relationship between an agent's effectiveness and the complexity of the learning environment. Specifically, an agent's effectiveness appears to increase as the visual search complexity of the learning environment increases. Under the low visual search conditions used in Experiment 3, there was no advantage associated with an agent's image. On the other hand, under the high visual search conditions used in Experiment 4, the visual presence of an agent clearly fostered learning more than voice-only (i.e., *image effect*), particularly on far transfer (large effect).

In Experiment 4, we also found that the learners that interacted with the agent as opposed to voice + highlighting or voice-only also spent significantly more time solving the posttest items. One could argue that this provides additional evidence that animated agents assume the role of a human teacher and that the life-like characteristics and behaviors of an agent prompt the social engagement of the

learner, thus allowing the learner to form a simulated human bond with the agent. In contrast, in an agent-less learning environment, a learner may identifying a computer interaction as being a case of information delivery due to the prevalence of weak social cues (e.g., disembodies voice), which leads to a failed attempt to foster an authentic social partnership with the learner. As a result, the learner does not rely on his or her sense-making processes, as in a case of social conversation, but merely attempts to learn by memorization. Due to the learner's inadequate cognitive processing (i.e., poor selection of information and ineffective organizational and integration strategies), his or her performance on subsequent tests of transfer suffered. This explanation offers one account for the advantage associated with the agent condition.

Future Directions

One potential explanation for this lackluster outcome of Experiment 3 is the nature of the learning environment itself. Specifically, the worked examples provided in this learning environment consisted of the sequential presentation of problem states in order to highlight problem subgoals. By sequentially presenting problems states, this type of worked example focuses the learners' attention on the process of constructing a solution to a problem, allowing them to examine each component of the example's solution in relative isolation from the one preceding it. That is, instead of appearing on the screen as a completely worked problem as is the case with examples that simultaneously display all of the solution components (i.e., *high* visual search environment), the sequential example appears initially unsolved. Learners then move forward through the example and watch as problem states successively added over a series of pages—similar to an animation, with the final page in the series representing the solution in its entirety.

There are several advantages associated with sequential presentation of problem states. For instance, this feature encourages learners to engage in *anticipative reasoning*—demonstrated to be a successful self-explanation style (Renkl, 1997)—by allowing students to anticipate the next step in an example's solution. Moreover, presenting problem states sequentially, like other forms of dynamic media, can bolster mathematical thinking in general by emphasizing variation over time. As Kaput (1992) posits, "one very important aspect of mathematical thinking is the abstraction of invariance...but, of course, to recognize invariance—to see what stays the same—one must have variation" (p. 526). He also suggests that "in static media, the states of notational objects cannot change as a function of time, whereas in dynamic media the can. Hence, time can become an information-carrying dimension" (p. 525).

Since we felt that this was an important issue to address, we conducted an experiment, Experiment 6, in which we compared the learning gains associated with these two techniques for presenting problem states, that is simultaneous and sequential presentation of problem states. We also examined the possibility for an interaction that might exist between the presence of an agent and the visual complexity of the environment (i.e., simultaneous versus sequential presentation of problem states).

We also suggest that this set of research questions should be reexamined in the context of nonlinear learning environment, one that require the agent to direct learners' attention to items on the screen that are not presented in a linear fashion as was the case in the present experiments (i.e., top to bottom). A study of this nature could potentially provide a better test of an agent's ability to guide and engage learners compared to other visual signals or cues since learners would not be able to simply read in linear fashion from top to bottom of the screen.

Exploring the Impact of Varying an Agent's Degree of Embodiment (Experiment 5 and 6)

To date, research on animated pedagogical agents has yielded favorable results in support of incorporating animated agents into multimedia learning environments. However, little research has been conducted that examines the degree of animation that an agent must possess in order to be effective. In light of this void, Baylor and Ryu (2003) investigated student perceptions of agents who were either static or animated. Seventy-five preservice teachers participated in a computer-based learning environment that presented a case study in which students had to design an instructional plan to teach supply and demand. Participants were assigned to work in one of three versions of the program: (a) fully animated agent condition, which employed gestures, (b) static agent condition using only a static image, or (c) no-image condition which only provided textual instructions. Across each condition, students received identical amounts of guidance, verbal instructions and textual instructions that appeared in a text bubble, which corresponded to the verbal explanations. The animated agent guided learners through the learning environment, provided examples and advisements that promoted the learner's understanding of the assignment. The experimental sessions required students to participate for approximately 90 minutes. Following exposure to the learning environment, learners were administered measures that assessed their perception of the agent – specifically, how engaging, person-like, credible and instructor-like was the agent. Performance measures were also collected which evaluated the learner's accuracy and performance during the learning environment. Results indicated students in the fully animated condition found the agent to be more engaging and more instructor-like than their peers in both the static agent condition and the no-image condition. Further, students exposed to the fully animated condition indicated that the agent was more person-like than students in the static agent condition. Students rated the agent in the fully animated and static conditions more credible than students in the no-image condition. Finally, no statistically significant differences among conditions were found in terms of performance during the learning environment.

Learners in the Baylor and Ryu (2003) study indicated that an agent possessing the most human like characteristics was more engaging, person-like, credible and instructor-like. Therefore, designers of animated agents should develop believable, life-like agents that are fully expressive as opposed to relying on stationary images of agents. Although the fully animated agent appeared superior in terms of student perceptions, it did not produce a greater level of performance relative to the additional two conditions. In order for animated agents to be optimally effective surrogate tutors, they must create social relationships with learners and promote deeper levels of understanding and learning.

Although the value of allowing animated agents to verbally guide learners through learning programs appears salient, the physical attributes and personality of the agents must be considered in order to ensure their optimal impact on learning. According to Johnson et al. (2000), in order for agents to be optimally beneficial in their environment they must be lifelike and believable. Animated agents that possess human like characteristics afford learners more enjoyable and engaging interactions and ultimately a more fulfilling learning experience. Further, agents should display humanistic behaviors because computer-like behaviors present an obvious discrepancy from lifelike characteristics and could interfere with the learner's attention to the content. Lifelike animated agents simulate face-to-face interactions between computers and learners. Upon the learner's establishment of a humanistic connection with a computer-based environment, the residential animated agent can demonstrate learning tasks, guide the learner through tutorials, provide emotive verbal and nonverbal feedback

and direct the learner's attention to the most important aspects of the instruction using gaze, gesture and locomotion.

As substantiated with the previous review of literature, adding an animated pedagogical agent into a learning environment to provide academic lessons in a variety of domains has yielded favorable performance and learning results. Further, research has suggested in order for agents to be maximally effective, they should be life-like, emotive, engaging characters. However, animated agent research has yet to examine the level of humanistic attributes that are necessary for an agent to possess to remain effective in terms of learning performance. Discovering the degree of life-like characteristics needed by agents has practical implications for instructional designers and future research. Specifically, if an animated agent that displays little humanistic traits is equivocal to a fully expressive agent, the programming efforts of designers can be reduced while still offering effective instructional devices. Identifying the effective animated agent will enable future researchers to tease out which physical properties allow the agent to foster learning (i.e., voice, movement, tactics to direct learner attention such as gesture and gaze). Therefore, the current study sought to answer the previous question by manipulating the humanistic properties of three versions of an animated agent to determine which agent aids in the creation of an environment most conducive to learning.

Across two experiments (Experiments 5 and 6), participants received a narrated set of worked-out examples for proportional reasoning word problems spoken by a female native-English speaker. In Experiment 5, participants were assigned to one of two agent-based conditions: (a) fully embodied, where the agent used gaze, gesture and locomotion to highlight each example's problem-solving application and solution, or (b) minimally embodied, where the agent was not programmed to direct learner attention using gesture, gaze, or movement around the workstation. Instead, the minimally embodied agent remained static—with the exception of its mouth, which was synched with the voice—in the top right area of the screen throughout the instructional phase of the learning environment. Moreover, Experiment 5 was similar to Experiment 3 in that the agents were deployed in a simple visual environment (i.e., sequential presentation of problem states). Finally, Experiment 6 examined the main effects and possible interaction associated with three versions of an animated agent (fully embodied, minimally embodied, or voice-only condition) and two types of learning environments (simple or complex). As with the previous experiments (Experiments 1, 2, 3, and 4), both learning process and learning outcome measures were collected. The learning process measures included perceived example understanding, perceived example difficulty, and performance on practice problems. The learning outcome measures included a posttest, which contained both near and far transfer items, and a speaker-rating questionnaire designed to detect the social characteristics attributed to speakers.

Experiment 5

This experiment was designed to address one primary question: Does an agent's degree of embodiment affect learning in a *low* visual search learning environment?

Sample and Design

The participants were 80 undergraduate college students recruited from educational psychology courses at Mississippi State University. They were randomly assigned in equal numbers to one of two conditions, with 40 serving in the fully embodied agent group and 40 serving in the minimally

embodied agent group. The percentage of females was 75% in the fully embodied group and 78% in the minimally embodied group; the percentage of juniors and seniors was 90% in the fully embodied group and 85% in the minimally embodied group; the percentage of students majoring in education or educational psychology was 88% in the fully embodied group and 80% in the minimally embodied group; and the mean GPA was 2.92 for the fully embodied group and 3.08 for the minimally embodied group.

Computer-Based Learning Environment

The learning environment used in Experiments 3 was modified to accommodate the present experiment. Although the voice + agent voice condition from Experiment 3 was simply relabeled the Fully Embodied Agent condition with no other alterations, a new condition—the Minimally Embodied Agent condition—was created out of the voice + agent condition in Experiment 3. The specific modifications are detailed in the next section. As with the learning environment used in Experiment 3, the worked examples provided in this learning environment consisted of the sequential presentation of problem states in order to highlight problem subgoals—which we characterize as a *low* visual search condition for purposes of this experiment. Specifically, the sequential presentations were presented as follows: Initially the examples appeared unsolved. Then the learner proceeded through each example while the problem states were gradually added on the screen until the example was presented in its entirety. This type of worked example focuses the student's attention on the practice of creating a solution to the problem. This practice allows students to study each component of the example's solution in isolation from the one preceding it, because learners can progress through each example, examining each problem state and the transformation required to accomplish the following state. For each example, a control panel was provided thus allowing learners to move throughout each example at their own pace. Throughout each solution step, instructional elaborations were orally provided to highlight the activity in each solution step (i.e., "First, we need to set up a proportional relationship to determine the cost of the travel package without the discount"). The subgoals nested within each example were labeled (i.e., "Initial Amount") in order to distinguish the problem's subgoals from one another.

The learning environment was configurable to run in one of two instructional modes that reflected the two conditions of the present experiment:

Fully Embodied Agent (FE) - The animated agent in this condition consisted of an agent in the form of a parrot – Peedy the Parrot. As the problem states were unfolded, Peedy moved around the workstation – from one subgoal to another – in order to highlight each example's problem-solving application and solution (see Figure 1). The fully embodied Peedy employed gesture and gaze to direct the learner's attention to the appropriate problem state of the example. In addition to directing learner attention via gaze, gesture and locomotion, Peedy also orally provided instructional elaborations that correspond to what is occurring in the appropriate problem state. Using Microsoft Agent software, Peedy was programmed to deliver recorded audio files of a human tutor (a female native-English speaker) providing the example instruction. Further, this software package allows Peedy's mouth to be synchronized with the human tutor's voice.

Minimally Embodied Agent (ME) - Unlike the fully embodied animated agent, capable of both verbal and nonverbal (gaze, gesture and locomotion) modes of communication, the minimally embodied agent was only capable of verbal communication (see Figure 7). That is, the minimally embodied Peedy was not programmed to direct learner attention using gesture, gaze, or movement

around the workstation. The minimally embodied Peedy remained static in the top right area of the screen throughout the instructional phase of the learning environment. However, Peedy's verbal provision of instructional elaborations was indistinguishable from those in the fully embodied condition.

Pencil-Paper Materials

The pencil-paper materials were identical to Experiment 1.

Procedure

The procedure was identical to Experiment 1.

Scoring

The scoring was identical to Experiment 1.

Results and Discussion

The major research question addressed in this experiment concerned whether learners in the fully embodied agent condition reported a significantly different interest in learning and achieved dissimilar levels transfer than learners in the minimally embodied condition. Table 5 shows the mean score (and standard deviation) for each group in Experiment 5 on the perceived example understanding, perceived example difficulty, performance on practice problems, near transfer test, far transfer test, the speaker rating survey, and instructional time. Separate two-tailed t-tests were conducted on these measures, each at alpha = .05. Cohen's d statistic was used as an effect size index where d values of .2, .5, and .8 correspond to small, medium, and large values, respectively (Cohen, 1988).

Does an agent's degree of embodiment affect learning process measures? There were no significant differences between conditions on perceived example understanding, $t(38) = 1.58$, p = ns, perceived example difficulty, $t(38) = 0.26$, p = ns, performance on practice problems, $t(38) = .49$, p = ns, or instructional time, $t(38) = 0.25$, p = ns.

Does an agent's degree of embodiment affect learning outcome measures? There were no significant differences between conditions on near transfer, $t(38) = 0.10$, p = ns, far transfer, $t(38) = 0.43$, p = ns, speaker rating survey, $t(38) = 1.24$, p = ns, or time on posttest, $t(38) = 0.13$, p = ns.

In sum, there were no differences between conditions on both measures collected during the learning process and measures collected as an outcome of learning. Although we cannot conclude with certainty that no difference exists between the two conditions due to the relatively low power of the design, the results suggest the learners (a) found the examples reasonably easy to understand and (b) not very difficult, (c) the practice problems moderately challenging, (d) experienced some success on transfer measures, regardless of what type of agent, fully embodied or minimally embodied, that accompanied the examples.

Experiment 6

The current experiment examined the role of animated pedagogical agents and their ability to mediate learning environments with varying complexity levels. One goal of this study was to manipulate the physical properties of an animated agent and identify the features necessary for the agent to effectively deliver instruction. An additional goal of this experiment was to manipulate the visual search complexity of a multimedia learning environment in order to identify the types of worked examples that encourage optimal student learning. The final goal of the experiment was to explore for a possible interaction between an agent's degree of embodiment and the visual search complexity of the learning environment.

This experiment was designed to address three questions. Specifically, across both low and high visual search conditions: (a) Does an agent's degree of embodiment affect learning, (b) does visual search complexity affect learning, and (c) does an agent's degree of embodiment interact with visual search complexity? Again, without a solid research base from which to base predictions, none were offered the outset of the experiment.

Sample and Design

The participants consisted of 174 undergraduate college students recruited from several education, educational psychology, and psychology courses offered at Mississippi State University. The participants received extra course credit for their participation. The sample was comprised of 43 (24.7%) males, 131 (75.3%) females and included 3 freshmen (1.7%), 14 sophomores (8%), 72 juniors (41.4%), and 85 seniors (48.9%). Of the entire sample, 107 (61.5%) were educational psychology/psychology majors, 46 (26.4%) were teacher education majors (elementary education, secondary education and special education), and 21 (12.1%) reported other as their major. The participants had an average grade point of 2.98 (SD = .50).

The participants were randomly assigned in equal proportions ($n = 29$) to one of six conditions, as defined by the cells of a 2 x 3 factorial design. The first factor was the visual search complexity of the learning environment (low or high); the second factor was the type of animated pedagogical agent present during the tutorial (FE – fully embodied agent, ME – minimally embodied agent, or VO – voice-only).

Computer-Based Learning Environment

The learning environment used in the experiment was configurable to run in one of six instructional modes that reflected the six conditions of the present experiment:

Complexity of Learning Environment

Low Visual Search Learning Environment Conditions - The worked examples provided in these conditions consisted of the sequential presentation of problem states in order to highlight problem subgoals. Specifically, the sequential presentations were presented as follows: Initially the examples appeared unsolved. Then the learner proceeded through each example while the problem states were gradually added on the screen until the example was presented in its entirety. This type of worked example focuses the student's attention on the practice of creating a solution to the problem. This

Experiment 6

The current experiment examined the role of animated pedagogical agents and their ability to mediate learning environments with varying complexity levels. One goal of this study was to manipulate the physical properties of an animated agent and identify the features necessary for the agent to effectively deliver instruction. An additional goal of this experiment was to manipulate the visual search complexity of a multimedia learning environment in order to identify the types of worked examples that encourage optimal student learning. The final goal of the experiment was to explore for a possible interaction between an agent's degree of embodiment and the visual search complexity of the learning environment.

This experiment was designed to address three questions. Specifically, across both low and high visual search conditions: (a) Does an agent's degree of embodiment affect learning, (b) does visual search complexity affect learning, and (c) does an agent's degree of embodiment interact with visual search complexity? Again, without a solid research base from which to base predictions, none were offered the outset of the experiment.

Sample and Design

The participants consisted of 174 undergraduate college students recruited from several education, educational psychology, and psychology courses offered at Mississippi State University. The participants received extra course credit for their participation. The sample was comprised of 43 (24.7%) males, 131 (75.3%) females and included 3 freshmen (1.7%), 14 sophomores (8%), 72 juniors (41.4%), and 85 seniors (48.9%). Of the entire sample, 107 (61.5%) were educational psychology/psychology majors, 46 (26.4%) were teacher education majors (elementary education, secondary education and special education), and 21 (12.1%) reported other as their major. The participants had an average grade point of 2.98 (SD = .50).

The participants were randomly assigned in equal proportions ($n = 29$) to one of six conditions, as defined by the cells of a 2 x 3 factorial design. The first factor was the visual search complexity of the learning environment (low or high); the second factor was the type of animated pedagogical agent present during the tutorial (FE – fully embodied agent, ME – minimally embodied agent, or VO – voice-only).

Computer-Based Learning Environment

The learning environment used in the experiment was configurable to run in one of six instructional modes that reflected the six conditions of the present experiment:

Complexity of Learning Environment

Low Visual Search Learning Environment Conditions - The worked examples provided in these conditions consisted of the sequential presentation of problem states in order to highlight problem subgoals. Specifically, the sequential presentations were presented as follows: Initially the examples appeared unsolved. Then the learner proceeded through each example while the problem states were gradually added on the screen until the example was presented in its entirety. This type of worked example focuses the student's attention on the practice of creating a solution to the problem. This

practice allows students to study each component of the example's solution in isolation from the one preceding it, because learners can progress through each example, examining each problem state and the transformation required to accomplish the following state. For each example, a control panel was provided thus allowing learners to move throughout each example at their own pace. Throughout each solution step, instructional elaborations were orally provided to highlight the activity in each solution step (i.e., "First, we need to set up a proportional relationship to determine the cost of the travel package without the discount"). The subgoals nested within each example were labeled (i.e., "Initial Amount") in order to distinguish the problem's subgoals from one another.

High Visual Search Learning Environment Conditions - The worked examples presented in these conditions were identical to those presented in the simple learning environment with one notable exception: the problem states were presented simultaneously. That is, the worked examples simultaneously displayed all of the solution components in their entirety. Identical to the simple learning environment, instructional elaborations were orally provided to emphasize the activity in each solution step. Additionally, the subgoals were labeled in order to distinguish the problem's subgoals from one another.

Type of Animated Agent Present During Instruction

Fully Embodied Agent (FE) Conditions - The animated agent in these conditions consisted of an agent in the form of a parrot – Peedy the Parrot. As the problem states were unfolded, Peedy moved around the workstation – from one subgoal to another – in order to highlight each example's problem-solving application and solution. The fully embodied Peedy employed gesture and gaze to direct the learner's attention to the appropriate problem state of the example. In addition to directing learner attention via gaze, gesture and locomotion, Peedy also orally provided instructional elaborations that correspond to what is occurring in the appropriate problem state. Using Microsoft Agent software, Peedy was programmed to deliver recorded audio files of a human tutor (a female native-English speaker) providing the example instruction. Further, this software package allows Peedy's mouth to be synchronized with the human tutor's voice (see Figure 1 for simple learning environment and Figure 6 for complex learning environment).

Minimally Embodied Agent (ME) Conditions - Unlike the fully embodied animated agent, capable of both verbal and nonverbal (gaze, gesture and locomotion) modes of communication, the minimally embodied agent was only capable of verbal communication. That is, the minimally embodied Peedy was not programmed to direct learner attention using gesture, gaze, or movement around the workstation. The minimally embodied Peedy remained static in the top right area of the screen throughout the instructional phase of the learning environment. However, Peedy's verbal provision of instructional elaborations was indistinguishable from those in the fully embodied condition (see Figure 7 for simple learning environment and Figure 8 for complex learning environment).

Voice-Only (VO) Conditions – These conditions were voice-only conditions in that the instructional lessons were verbally delivered in the *absence* of an animated agent. The oral instructions in this condition were identical to those in the former conditions with one exception: the instructions were presented in the form of a voice-over instead of being presented by an animated agent (see Figure 2 for simple learning environment and Figure 4 for complex learning environment).

Pencil-Paper Materials

The pencil-paper materials were identical to Experiment 1 with one notable exception: a pretest was included in the present experiment. The pretest contained 11 proportional reasoning problems of varying difficulty. It assessed the learner's ability to complete basic mathematical operations as well as solve proportional reasoning word problems prior to treatment exposure. The items on the pretest incorporated four one-step proportion word problems, two multistep problems with one proportional relationship, three multistep problems with two proportional relationships, and two problems not involving proportional reasoning in their solutions (see Appendix F). The items on the pretest were assigned one-point for a correct response and zero-points for an incorrect response. The maximum attainable score on the pretest was 11. An example of a multistep problem on the pretest with two proportional relationships is as follows:

> Sheri, a student architect, wants to establish the difference in height between two buildings, the courthouse and the bank. If Sheri is 6 feet tall and casts a shadow 9 feet long and, at the same time, the shadows of the two buildings are 90 and 120 feet long, what is the difference in height between the two buildings?

Procedure

The procedure was identical to Experiment 1 with one notable expectation: the administration of the pretest. After the participants completed a voluntary consent form, the demographic questionnaire, and studied the eight-page review on solving proportion problems, they took the pretest. After completing the pretest, the participants were exposed to the learning phase of the experiment where they independently studied and worked proportional reasoning word problems in the computer-based learning environment.

Scoring

The scoring was identical to Experiment 1 with one notable expectation: the scoring of the pretest. The items on the pretest were assigned a one (correct) or a zero (incorrect) based on the accuracy of the response. No partial credit was assigned for items on the pretest. Therefore, the maximum score that may be earned on the pretest was 11.

Results and Discussion

Three major research questions were addressed in this experiment: (a) Does an agent's degree of embodiment affect learning, (b) does visual search complexity affect learning, and (c) does an agent's degree of embodiment interact with visual search complexity? Table 6 conveys the descriptive statistics associated with the three conditions in the low visual search environment whereas Table 7 contains the same information for the three conditions associated with the high visual search environment. Each of the tables reported the mean scores and standard deviations of each condition on the perceived example understanding, perceived example difficulty, performance on practice problems, near transfer test, far transfer test, the speaker rating survey, and instructional time. Factorial (2x3) analyses of covariance (ANCOVAs), using the pretest score as the covariate, were conducted to analyze the learning process measures and learning outcome measures, each at alpha = .05. Significant main effects were followed up with Fisher's LSD test, based on a familywise

alpha of .05. Cohen's *f* statistic was used as an effect size index where *f* values of .10, .25, and .40 correspond to small, medium, and large values, respectively (Cohen, 1988).

Does an agent's degree of embodiment affect learning? In terms of learning process measures, there was no statistically significant main effect for type of animated agent present during instruction on practice problem-solving performance, $F(2,167) = 1.07$, MSE = .08, $p > .05$, on perceived worked example difficulty, $F(2,167) = .61$, MSE = 2.01, $p > .05$, or on instructional time, $F(2, 167) = .78$, MSE = 131.48, $p > .05$.

There was, however, a statistically significant main effect for type of animated agent present during instruction on perceived worked example understanding, $F(2,167) = .3.35$, MSE = .38, $p < .05$. Fisher's LSD test indicated that participants in the minimally embodied condition ($M = 1.49$, $SD =$ .63) reported a higher level of understanding of the worked examples than did participants in the no agent condition (M = 1.79, $SD = .74$). Cohen's *f* statistic for these data yields an effect size estimate of .18 for perceived example understanding, which corresponds to a small effect. According to Fisher's LSD test, no differences existed between the other animated agent conditions.

In terms of learning process measures, there was a statistically significant main effect for type of animated agent present during instruction on near transfer test performance, $F(2,167) = 3.67$, MSE = .42, $p < .05$. Fisher's LSD test indicated that participants in the fully embodied condition ($M = 2.21$, $SD = .69$) significantly outperformed their peers in the no agent condition ($M = 1.91$, $SD = .87$) in terms of near transfer test performance. Cohen's *f* statistic for these data yields an effect size estimate of .18, for near transfer test performance, which corresponds to a small effect.

There was also a statistically significant main effect for type of animated agent present during instruction on far transfer performance, $F(2,167) = 3.79$, MSE = .52, $p < .05$. Fisher's LSD test indicated that participants in the fully embodied condition ($M = 1.47$, $SD = .90$) significantly outscored participants in the no agent condition ($M = 1.11$, $SD = .95$) on far transfer test performance. Cohen's *f* statistic for these data yields an effect size estimate of .18, for far transfer test performance, which corresponds to a small effect.

The social agency survey measured the animated agent on three subscales, (1) superiority, (2) attractiveness, and (3) dynamism. There was no statistically significant main effect for type of animated agent present during instruction on the evaluation of the agent's superiority $F(2,167) = .63$, MSE = 1.12, $p > .05$ or attractiveness $F(2,167) = 1.5$, MSE = 1.76, $p > .05$. However, there was a statistically significant main effect for type of agent on dynamism $F(2,167) = 3.26$, MSE = 1.39, $p <$ .05, in which Fisher's LSD indicated that students in the fully embodied condition ($M = 3.51$, $SD =$ 1.02) rated the agent more dynamic than those students in the no agent condition ($M = 4.04$, $SD =$ 1.17). Cohen's *f* statistic for these data yields an effect size estimate of .20, for dynamism, which corresponds to a small-to-medium effect.

Does visual search complexity affect learning? In terms of the learning process measures, there was no statistically significant main effect for complexity of learning environment on perceived worked example understanding, $F(1,167) = .84$, MSE = .38, $p > .05$, perceived worked example difficulty, $F(1, 167) = 1.07$, MSE = 2.01, $p > .05$, practice problem-solving performance, $F(1,167) = .34$, MSE = .08, $p > .05$, or instructional time, $F(1, 167) = .02$, MSE = 131.48, $p > .05$.

There was, however, a statistically significant main effect for complexity of learning environment on near transfer test performance, $F(1, 167) = 3.94$, MSE = .42, $p < .05$. The participants assigned to the simple learning environment ($M = 2.12$, $SD = .76$) outperformed their peers in the complex learning environment ($M = 1.92$, $SD = .80$) in terms of near transfer test performance. Cohen's f statistic for these data yields an effect size estimate of .14 for near transfer test performance, which corresponds to a small effect.

Moreover, there was a statistically significant main effect for complexity of learning environment on far transfer performance, $F(1, 167) = 4.22$, MSE = .52, $p < .05$. Participants in the simple learning environment ($M = 1.39$, $SD = .93$) outscored participants in the complex learning environment ($M = 1.17$, $SD = .82$) on far transfer test performance. Cohen's f statistic for these data yields an effect size estimate of .14 for far transfer test performance, which corresponds to a small effect.

Finally, there was no statistically significant main effect for complexity of learning environment in the evaluation of the agent's superiority $F(1, 167) = .89$, MSE = 1.12, $p > .05$, attractiveness $F(1, 167) = .03$, MSE = 1.76, $p > .05$ or dynamism $F(1, 167) = .35$, MSE = 1.39, $p > .05$ between students working in the simple learning environment and those working in the complex learning environment.

Does an agent's degree of embodiment interact with visual search complexity?

Although there was no statistically significant interaction of the main effects (type of agent and visual search complexity) on practice problem-solving performance, $F(2,167) = .15$, MSE = .08, $p > .05$, or on instructional time, $F(2, 167) = .07$, MSE = 131.48, $p > .05$, there a statistically significant interaction on perceived worked example understanding, $F(2,167) = 4.37$, MSE = .38, $p < .05$. Cohen's f statistic for these data yields an effect size estimate of .20, for perceived example understanding, which corresponds to a small-to-medium effect. Subsequent analysis demonstrated that there was a simple main effect for type of agent at the high level of the visual search complexity factor, $F(1, 83) = 5.19$, $p < .05$. Cohen's f statistic for these data yields an effect size estimate of .32, which corresponds to a medium effect. Participants in the high visual search environment who received the minimally embodied agent reported higher levels of understanding than their peers in the fully embodied and no agent groups. The remaining simple main effects were not significant.

There was also a statistically significant interaction on perceived worked example difficulty, $F(2,167) = 5.96$, MSE = 2.01, $p < .05$. Cohen's f statistic for these data yields an effect size estimate of .25, which corresponds to a medium effect. Similar to the outcome of the analysis of example understanding, subsequent analysis demonstrated that there was a simple main effect for type of agent at the high level of the visual search complexity factor, $F(1, 83) = 4.50$, $p < .05$. Cohen's f statistic for these data yields an effect size estimate of .32, which corresponds to a medium effect. Specifically, the participants in the high visual search environment who received the minimally embodied agent reported lower levels of perceived example difficulty than their counterparts in the fully embodied and the no agent conditions. The remaining simple main effects were not significant.

On the other hand, there was no statistically significant interaction on near transfer test performance, $F(2,167) = 1.05$, MSE = .42, $p > .05$, or on far transfer performance, $F(2,167) = .35$, MSE = .52, $p > .05$. Moreover, there was no statistically significant interaction on the evaluation of the agent's superiority $F(2, 167) = .15$, MSE = 1.12, $p > .05$, attractiveness.

In conclusion, this experiment provides modest evidence to support the research claim that multimedia learning environments encompassing animated pedagogical agents as virtual learning assistants are superior to multimedia learning environments that include learner assistance in the form of verbal instructions. Even though the effect sizes were smaller than anticipated, the results indicate that multimedia learning environments can be optimized when they are coupled with animated agents. Moreover, the transfer performance of learners can be exploited when they are exposed to a learning environment containing an animated agent. This experiment also highlights the benefit of presenting worked examples that encourage learners to study and process each problem state before viewing subsequent problem states (subgoal-oriented examples in the simple learning environment) rather than examples that concurrently present several problem steps (simultaneous-oriented examples in the complex learning environment). Similar to the effect sizes revealed with the agent factor, the measurable effects of the complexity of the learning environment were also small. However, the results of this experiment clearly specify that learners studying proportional reasoning during a computer-based program can benefit from receiving instructions via subgoal-oriented worked examples.

Conclusions Regarding the Impact of Varying an Agent's Amount of Embodiment

This set of experiments contributes to the growing literature on animated pedagogical agents, worked examples as well as multimedia learning environments. First, the results of Experiment 6 replicated the results of Experiment 4, which suggests incorporating an animated agent into a computer-based learning environment enhances learning more than conditions in which agents are not included (i.e., voice-only). Second, Experiment 6 also empirically investigated the difference between sequentially presented worked examples (i.e., examples with sequentially presented subgoals) and simultaneous presented worked examples (i.e., examples in which the subgoals were simultaneously presented). Since the subgoal-oriented examples (i.e., low visual search environment) proved superior to simultaneous-oriented examples (i.e., high visual search environment), the current study suggests that a *sequential* principle exists, in which sequentially presented subgoals are superior to simultaneously presented subgoals.

Implications

Experiments 5 and 6 investigated whether various types of animated agents, designed to provide instructional elaborations during a computer tutorial involving proportional reasoning were able to increase participants' performances on learning measures. Although no differences were found between the conditions in Experiment 5 (low visual search environment), the findings from Experiment 6 indicated that students receiving instructions from a fully embodied agent outperformed their peers in no agent condition in terms of near and far transfer performance (using the conceptual rubric); however, the measurable effects of this difference were small. The lack of a significant difference between the fully embodied condition and the minimally embodied condition in Experiment 6 is consistent with the findings of Experiment 5. The findings from the Experiment 6 in combination with those of the previous experiment suggest that in a linear computer-based environment, the visual presence of an animated agent is a critical factor in optimizing learning outcomes whereas an agent's mobility is a less important factor.

Results from this Experiment 6 indicate that learning from worked examples is optimized when an example's subgoals are presented in sequential fashion (i.e., low visual search environment).

Determining which type of worked examples benefit student learning and understanding has direct implications for educators and instructional designers. Specifically, worked examples that are provided in textbooks and on mathematics worksheets to serve as expert models for solving mathematics problems should consist of sequentially presented problem states similar to those presented in the simple learning environment. Worked examples should be designed to encourage learners to process and encode each solution step of an example in relative isolation in an effort to increase the chances of recalling strategies when solving subsequent problem-solving tasks in particular domains. Although this study and numerous previous studies suggest the benefit of employing subgoal-oriented examples to model expert problem-solving steps and solutions (Catrambone, 1994, 1996, 1998; Renkl, 1997) scores of textbooks and classroom based instructional activities continue employing conventionally based examples that concurrently present an example's entire set of problem states as well as the final solution (i.e., such as the worked examples included in the complex learning environment).

Future Directions

The results of this set of experiments offer several opportunities for future research, some specifically related to the further examination of animated pedagogical agents, and others to computer-based programs that provide contemporary instructional aides, such as worked examples. First, in order to evaluate which physical properties contribute to an agent's effectiveness in various learning environments, animated pedagogical agents should be incorporated into nonlinear environments that require the agent to direct learners' attention to items on the screen that are not presented in a linear fashion (i.e., not left to right or top to bottom). A study of this nature may provide support that an agent's ability to guide learners by moving from one-step to the next may optimize learning in nonlinear environments whereas the agent's locomotion may not be a crucial factor in directing learners in a linear environment. In other words, in environments that include problems presented in an easy to follow format (i.e., examples that unfold from top to bottom and read from left to right), employing a static on-screen agent as a visual and verbal indicator may independently optimize performance in lieu of movement. Therefore, a fully expressive agent should be included into a nonlinear learning environment and be programmed to direct learner's attention to randomly located problem steps to effectively assess the benefit of designing a fully embodied agent rather than a minimally embodied tutor.

An additional component for future research relative to animated pedagogical agents is investigating the impact of allowing learners to receive instructions from an animated agent that possesses the same characteristics of the learner (i.e., age, gender, and ethnicity). In particular, learning may be optimized if learners are able to obtain academic lessons from an agent that shares their similar physical properties. Findings from the counseling literature may have direct implications for research involving animated pedagogical agents. That is, computer-based learners may be similar to clients who are receiving assistance because the learners are receiving instructional elaborations from an animated agent and the animated agent may be similar to the counselor because both the agent and the counselor facilitate the clarity and understanding of the current issue. Multicultural counseling research suggests that African American clients prefer receiving counseling from African American counselors (Tien & Johnson, 1985) just as Asian American clients desire Asian American counselors (Atkinson, Poston, Furlong, & Mercado, 1989). Likewise, research has provided evidence that gender preference also plays a role in the counseling process. For instance, a study by Fowler, Wagner, Iachini, and Johnson (1992) indicated that female clients would rather participate in counseling sessions in which the counselor was female. Assuming these findings are transferable to

animated agents it would be worthwhile for future research to examine the influence that matching the learner and animated agent physical characteristics have on learning. In addition to investigating the interaction of learner and agent characteristics, it may also be valuable to include the subject matter in order to examine which populations benefit from which type of agents when providing instructions for various domains.

Finally, future research should also examine cognitive load as it relates to learners interacting with agents during computer tutorials. This set of experiments implemented a direct, subjective measure of cognitive load; however, the measure consisted only of a statement inquiring about the difficulty of the examples and practice problems, not the agents. Although the cognitive load measure used has been implemented in numerous empirical studies, future research should employ more direct and objective assessment methods to gauge the degree to which an agent facilitates understanding as well as decreases the difficulty level of problem-solving tasks. A study of this nature would help elucidate which physical properties are necessary for an agent to optimize a multimedia learning environment and maximize learning outcomes. For instance, a minimally embodied agent may promote high levels of understanding and low levels of perceived difficulty while not overtaxing the capacity of the working memory. Brünken, Plass, and Leutner (2003) recently proposed a promising method of assessing cognitive load in multimedia learning environments, the dual-task approach. This approach involves simultaneously engaging in two activities that both require the same amount of mental resources and examining the amount of attentional resources allocated to either the primary or the secondary task to measure cognitive load. Using this method, research could examine whether the cognitive load imposed by the agent interferes with problem solving performance.

REFERENCES

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

Atkinson, R. K., & Derry, S. J. (2000). Computer-based examples designed to encourage optimal example processing: A study examining the impact of sequentially presented, subgoal-oriented worked examples. In B. Fishman & S. F. O'Connor-Divelbiss (Eds.), *Proceedings of the Fourth International Conference of Learning Sciences* (pp. 132-133). Hillsdale, NJ: Erlbaum.

Baylor, A. & Ryu, J. (in press). Does the presence of image and animation enhance pedagogical agent persona? *Journal of Educational Computing Research.*

Brünken, R., Plass, J. L. & Leutner, D. (2003). Direct Measurement of cognitive load in multimedia learning. *Educational Psychologist, 38,* 53-61.

Brünken, R., Steinbacher, S., Plass, J.L. & Leutner, D. (2002). Assessment of cognitive load in multimedia learning using dual-task methodology. *Experimental Psychology, 49,* 109-119.

Cassell, J., Sullivan, J., Prevost, S., & Churchill, E. (2000). *Embodied conversational agents.* Cambridge, NA: MIT Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Ed.). Hillsdale, NJ: Erlbaum.

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.

Graesser, A.C., Moreno, K., Marineau, J., Adcock, A., Olney, A., & Person, N. (2003). AutoTutor improves deep learning of computer literacy: Is it the dialog or the talking head? In U. Hoppe, F. Verdejo, and J. Kay (Eds.), *Proceedings of Artificial Intelligence in Education.* (pp. 47-54). Amsterdam: IOS Press

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41-58). New York: Academic Press.

Jeung, H., Chandler, P., & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. Educational Psychology, 17, 329-433.

Johnson, W. L., Rickel, J. W., & Lester, J. C. (2000). Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education, 41,* 41-78.

Kaput, J. J. (1992). Technology and mathematics education. In D. A. Grouws (Ed.), *Handbook of research on mathematics and learning: A project of the National Council of Teachers in Mathematics* (pp. 515-556). New York: Macmillan.

Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T.., Stone, B. A. & Bhogal, R. (1997). The persona effect: Affective impact of animated pedagogical agents. In the *Proceedings of CHI 97 Human Factors in Computing Systems* (pp. 359-366). New York: Association for Computing Machinery.

Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology, 93*, 377-389.

Mayer, R. E. (1999). *The promise of educational psychology*. Upper Saddle River, NJ: Merrill Prentice Hall.

Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90,* 312-320.

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.

Mousavi, S., Low, R., & Sweller, J.(1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology, 87,* 319-334.

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.

Paas, F., & Van Merrienboer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combining mental-effort and performance measures. *Human Factors, 35,* 737-743.

Reeves, B., & Nass, C. (1996). *The media equation*. New York: Cambridge University Press.

Renkl, A. (1997). Learning from worked-out examples: A study of individual differences. *Cognitive Science, 21,* 1-29.

Zahn, C. J., & Hopper, R. (1985). Measuring language attitudes: The speech evaluation instrument. *Journal of Language and Social Psychology, 4,* 113-122.

TABLE 1

Mean Scores and Standard Deviations by Condition on the Measures of Experiment 1.

	Human Voice		Machine Voice	
	M	*SD*	*M*	*SD*
Perceived Example Understanding	1.32	.45	1.31	.35
Perceived Example Difficulty	2.21	.75	1.99	.58
Performance on Practice Problems	2.67[a]	.63	2.09	.85
Posttest - Near Transfer	2.23[a]	.71	1.62	.77
Posttest - Far Transfer	1.32[a]	.90	.77	.69
Speaker Rating Survey	2.29[a]	.84	3.10	1.30
Instructional Time	39.2	9.2	40.4	17.1

Note: [a]denotes human voice group scored significantly higher than machine voice group at $p < .05$; $n = 25$ for each group; instructional time is reported in minutes.

TABLE 2

Mean Scores and Standard Deviations by Condition on the Measures of Experiment 2.

	Human Voice		Machine Voice	
	M	*SD*	*M*	*SD*
Perceived Example Understanding	1.65	.45	1.51	.43
Perceived Example Difficulty	2.44	.75	2.40	.79
Performance on Practice Problems	2.33[a]	.64	1.80	.86
Posttest - Near Transfer	2.51[a]	.59	1.84	.86
Posttest - Far Transfer	1.74[a]	.70	1.15	.82
Speaker Rating Survey	3.19[a]	1.05	4.23	1.30
Instructional Time	40.7	16.4	42.1	10.3

Note: [a]denotes human voice group scored significantly higher than machine voice group at $p < .05$; $n = 20$ for each group; instructional time is reported in minutes.

TABLE 3

Mean Scores and Standard Deviations by Condition on the Measures of Experiment 3.

	Condition					
	Voice-Only		Voice + Highlighting		Voice + Agent	
	M	*SD*	*M*	*SD*	*M*	*SD*
Perceived Ex. Understanding	1.60	.57	1.59	.67	1.53	.52
Perceived Ex. Difficulty	2.63	.89	2.25	.74	2.38	.82
Performance on Practice Prob.	2.51	.59	2.56	.69	2.50	.54
Instructional Time	35.56	13.09	33.52	9.40	35.48	9.18
Posttest - Near Transfer	2.08	.86	2.01	.86	2.29	.74
Posttest - Far Transfer	1.43	.81	1.40	.93	1.42	.63
Time on Posttest	32.36	8.41	30.36	8.82	28.60	6.84
Speaker Rating - Superiority	2.47	1.38	2.02	.99	1.95	1.08
Speaker Rating - Attractiveness	2.67	1.11	2.50	1.16	2.62	1.57
Speaker Rating - Dynamism	4.10	1.39	3.77	1.15	3.23	1.24

Note: $n = 25$ for each group; instructional time is reported in minutes.

TABLE 4

Mean Scores and Standard Deviations by Condition on the Measures of Experiment 4.

	Condition					
	Voice-Only		Voice + Highlighting		Voice + Agent	
	M	*SD*	*M*	*SD*	*M*	*SD*
Perceived Ex. Understanding	2.64	.84	2.44	.97	2.66	.82
Perceived Ex. Difficulty	1.97	.87	1.80	.76	1.61	.72
Performance on Practice Prob.	1.38	1.14	2.11[a]	1.08	2.02	.89
Instructional Time	34.62	10.18	35.00	11.56	37.69	9.03
Posttest - Near Transfer	1.75	.98	1.79	1.00	1.95	.98
Posttest - Far Transfer	.73	.87	1.02	.97	1.42[a]	.87
Time on Posttest	23.96	7.97	26.58	8.35	32.27[a, b]	8.28
Speaker Rating - Superiority	2.36	1.26	2.38	1.17	2.30	1.31
Speaker Rating - Attractiveness	3.25	1.58	3.06	1.58	2.32[a]	1.01
Speaker Rating - Dynamism	3.81	1.34	3.88	1.07	3.79	1.38

Note: [a]differs statistically from voice-only at $p < .05$; [b]differs statistically from voice + highlighting at $p < .05$; $n = 26$ for each group; instructional time is reported in minutes.

TABLE 5

Mean Scores and Standard Deviations by Condition on the Measures of Experiment 5.

	Fully Embodied		Minimally Embodied	
	M	*SD*	*M*	*SD*
Perceived Example Understanding	2.64	.88	2.34	.82
Perceived Example Difficulty	1.71	.62	1.75	.68
Performance on Practice Problems	1.61	1.05	1.73	.99
Instructional Time	35.63	11.69	35.00	10.80
Posttest - Near Transfer	1.81	.83	1.79	.79
Posttest - Far Transfer	1.16	.95	1.08	.89
Speaker Rating Survey	3.04	1.12	2.76	.86
Time on Posttest	27.58	11.12	27.28	9.71

Note: $n = 40$ for each group; instructional time is reported in minutes.

TABLE 6

Unadjusted (for Covariate) Mean Scores and Standard Deviations by Condition in the Low Visual Search Environment on the Measures of Experiment 6.

	Simple Conditions					
	Fully Embodied Agent		Minimally Embodied Agent		No Agent	
	M	*SD*	*M*	*SD*	*M*	*SD*
Perceived Ex. Understanding	1.40	.48	1.68	.70	1.72	.65
Perceived Ex. Difficulty	2.86	1.28	3.5	1.79	2.78	1.22
Performance on Practice Problems	1.94	.86	1.84	.95	1.81	.95
Instructional Time	35.17	12.78	32.41	14.34	33.48	9.87
Near Transfer Posttest Items	2.35	.63	1.89	.70	2.12	.89
Far Transfer Posttest Items	1.63	.98	1.27	.78	1.30	1.01
Social Agency Survey Rating	2.72	.87	2.99	1.01	3.29	1.01

Note: $n = 29$ for each group; instructional time is reported in minutes.

TABLE 7

Unadjusted (for Covariate) Mean Scores and Standard Deviations by Condition in the High Visual Search Environment on the Measures of Experiment 6.

	Complex Conditions					
	Fully Embodied Agent		Minimally Embodied Agent		No Agent	
	M	*SD*	*M*	*SD*	*M*	*SD*
Perceived Ex. Understanding	1.91	.80	1.39	.52	1.78	.84
Perceived Ex. Difficulty	3.71	1.57	2.63	1.25	3.51	1.71
Performance on Practice Problems	1.63	1.14	1.91	1.03	1.78	1.03
Instructional Time	34.97	9.57	32.93	7.97	32.62	13.31
Near Transfer Posttest Items	2.05	.73	1.88	.84	1.82	.85
Far Transfer Posttest Items	1.31	.80	1.13	.79	1.05	.89
Social Agency Survey Rating	2.76	.89	2.93	1.05	3.16	.92

Note: $n = 29$ for each group; instructional time is reported in minutes.

FIGURE CAPTIONS

Figure 1. Voce + Agent (fully embodied) condition (*low* visual search environment).

Figure 2. Voice-only condition (*low* visual search environment).

Figure 3. Voice + Highlighting condition (*low* visual search environment).

Figure 4. Voice-only condition (*high* visual search environment).

Figure 5. Voice + Highlighting condition (*high* visual search environment).

Figure 6. Voce + Agent (fully embodied) condition (*high* visual search environment).

Figure 7. Minimally embodied agent (*low* visual search environment).

Figure 8. Minimally embodied agent in a (*high* visual search environment).

Instructions

Initial Amount

4 Students = **55 Students**

$1,377.97 **X**

4X = $1,377.97 * 55

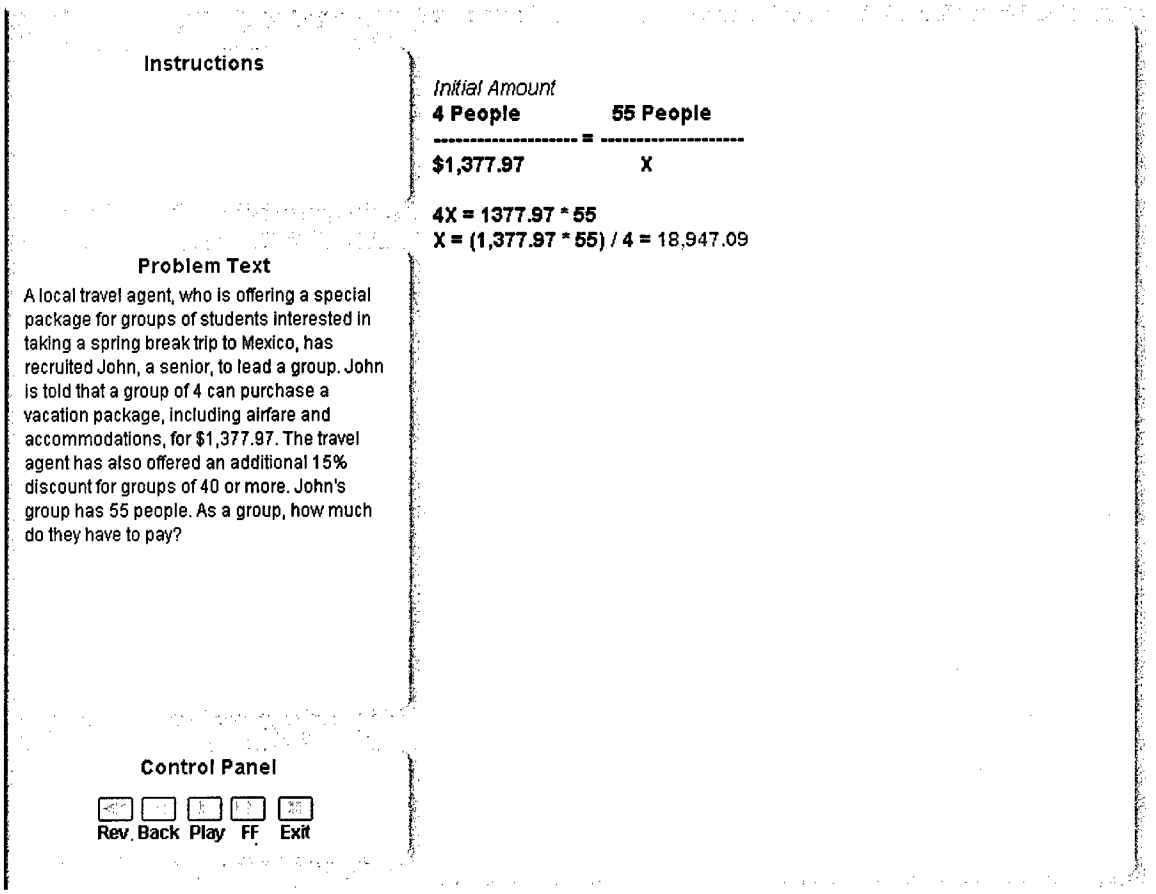
X = ($1,377.97 * 55) / 4 = $18,947.09

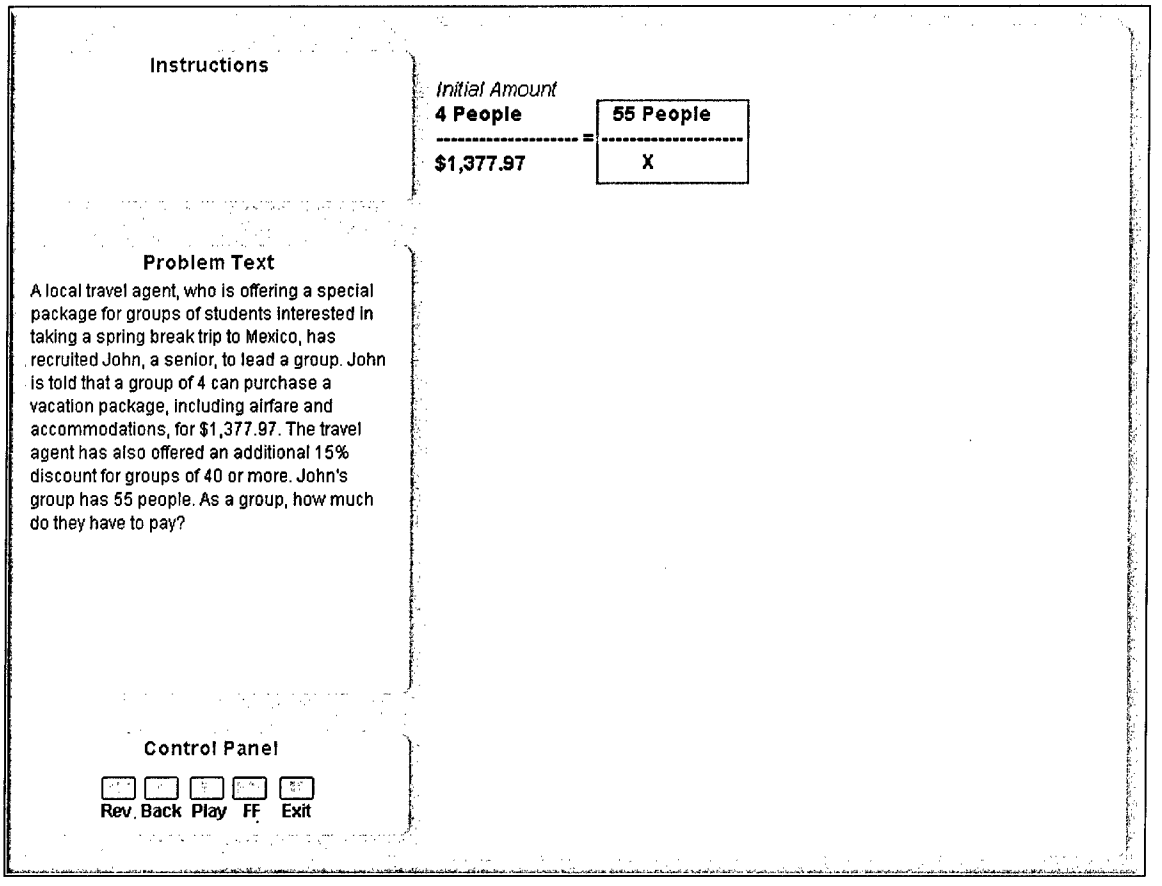
Problem Text

A local travel agent, who is offering a special package for groups of students interested in taking a spring break trip to Mexico, has recruited John, a senior, to lead a group. John is told that a group of 4 can purchase a vacation package, including airfare and accommodations, for $1,377.97. The travel agent has also offered an additional 15% discount for groups of 40 or more. John's group has 55 people. As a group, how much do they have to pay?

Control Panel

Rev. Back Play FF Exit


Instructions
Problem Text
A local travel agent, who is offering a special package for groups of students interested in taking a spring break trip to Mexico, has recruited John, a senior, to lead a group. John is told that a group of 4 can purchase a vacation package, including airfare and accommodations, for $1,377.97. The travel agent has also offered an additional 15% discount for groups of 40 or more. John's group has 55 people. As a group, how much do they have to pay?
Control Panel
Rev
Back
Play
FF
Exit
Initial Amount
4 People
55 People
$1,377.97
X
4X = 1377.97 * 55
X = (1,377.97 * 55) / 4 = 18,947.09


Instructions
Initial Amount
4 People
$1,377.97
=
55 People
X
Problem Text
A local travel agent, who is offering a special package for groups of students interested in taking a spring break trip to Mexico, has recruited John, a senior, to lead a group. John is told that a group of 4 can purchase a vacation package, including airfare and accommodations, for $1,377.97. The travel agent has also offered an additional 15% discount for groups of 40 or more. John's group has 55 people. As a group, how much do they have to pay?
Control Panel
Rev
Back
Play
FF
Exit

Instructions

Problem Text

A local travel agent, who is offering a special package for groups of students interested in taking a spring break trip to Mexico, has recruited John, a senior, to lead a group. John is told that a group of 4 can purchase a vacation package, including airfare and accommodations, for $1,377.97. The travel agent has also offered an additional 15% discount for groups of 40 or more. John's group has 55 people. As a group, how much do they have to pay?

Control Panel

Rev. Back Play FF Exit

Initial Amount

4 Students **55 Students**

$1,377.97 = **$18,947.09**

4X = $1,377.97 * 55

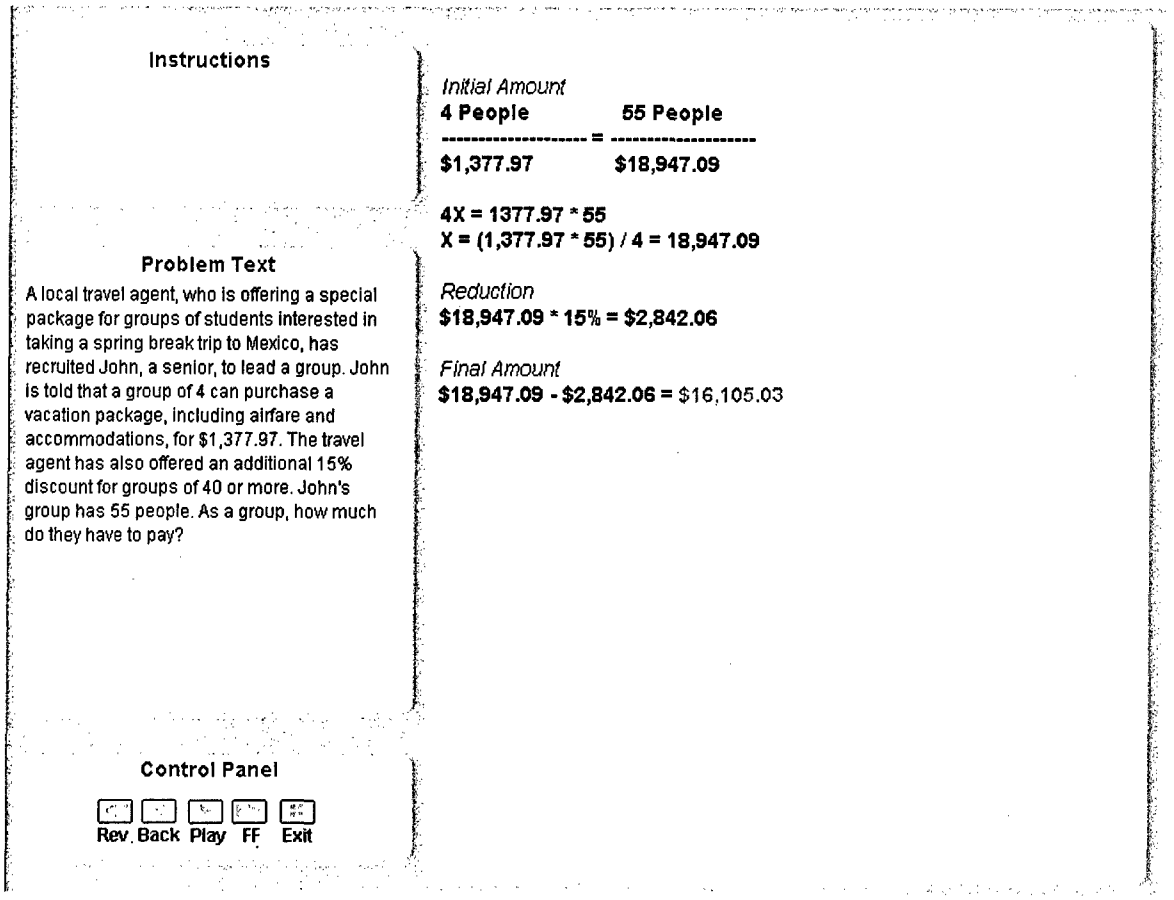
X = ($1,377.97 * 55) / 4 = $18,947.09

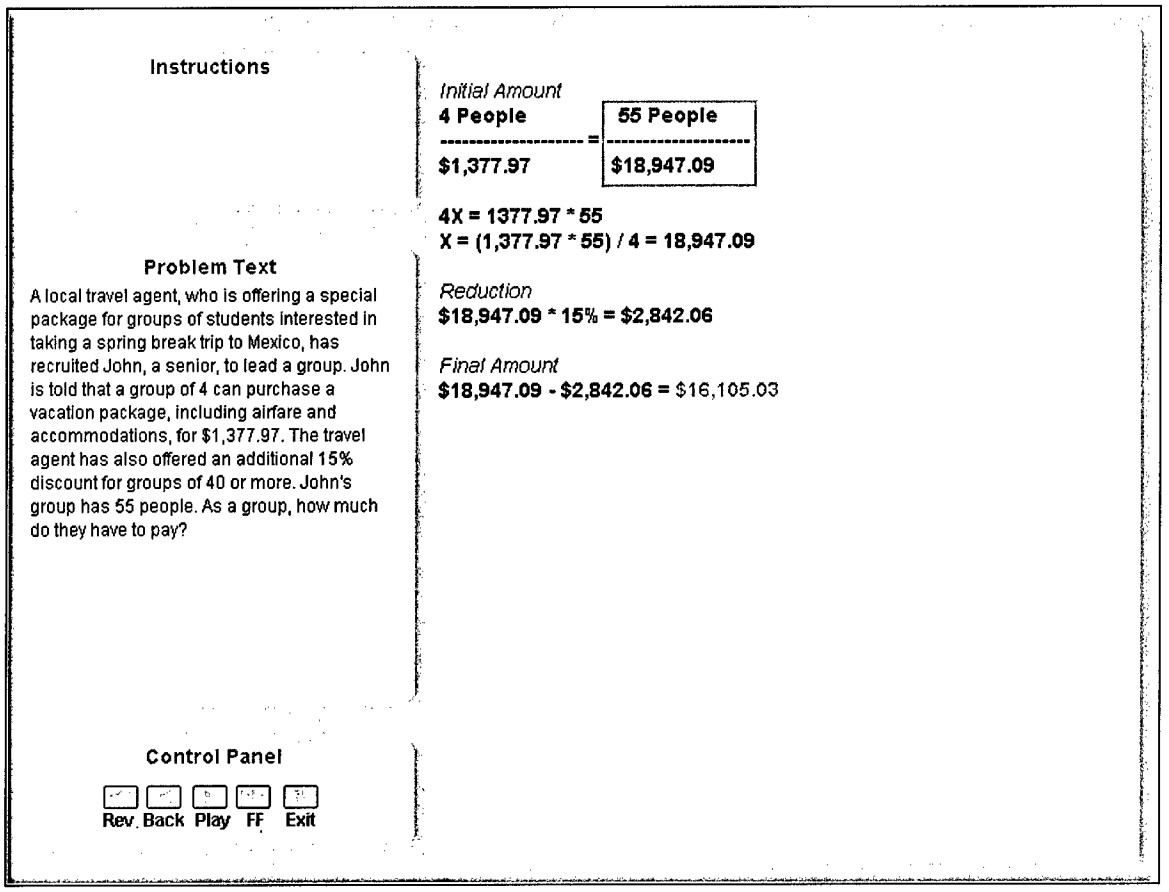
Discount

$18,947.09 * 15% = $2,842.06

Final Amount

$18,947.09 - $2,842.06 = $16,105.03


Instructions
Problem Text
A local travel agent, who is offering a special package for groups of students interested in taking a spring break trip to Mexico, has recruited John, a senior, to lead a group. John is told that a group of 4 can purchase a vacation package, including airfare and accommodations, for $1,377.97. The travel agent has also offered an additional 15% discount for groups of 40 or more. John's group has 55 people. As a group, how much do they have to pay?
Control Panel
Rev
Back
Play
FF
Exit
Initial Amount
4 People
55 People
$1,377.97
$18,947.09
4X = 1377.97 * 55
X = (1,377.97 * 55) / 4 = 18,947.09
Reduction
$18,947.09 * 15% = $2,842.06
Final Amount
$18,947.09 - $2,842.06 = $16,105.03


Instructions
Problem Text
A local travel agent, who is offering a special package for groups of students interested in taking a spring break trip to Mexico, has recruited John, a senior, to lead a group. John is told that a group of 4 can purchase a vacation package, including airfare and accommodations, for $1,377.97. The travel agent has also offered an additional 15% discount for groups of 40 or more. John's group has 55 people. As a group, how much do they have to pay?
Control Panel
Rev
Back
Play
FF
Exit
Initial Amount
4 People
55 People
$1,377.97
$18,947.09
4X = 1377.97 * 55
X = (1,377.97 * 55) / 4 = 18,947.09
Reduction
$18,947.09 * 15% = $2,842.06
Final Amount
$18,947.09 - $2,842.06 = $16,105.03

Instructions

Problem Text

A local travel agent, who is offering a special package for groups of students interested in taking a spring break trip to Mexico, has recruited John, a senior, to lead a group. John is told that a group of 4 can purchase a vacation package, including airfare and accommodations, for $1,377.97. The travel agent has also offered an additional 15% discount for groups of 40 or more. John's group has 55 people. As a group, how much do they have to pay?

Control Panel

Rev Back Play FF Exit

Initial Amount

4 Students		**55 Students**
$1,377.97	=	**X**

4X = $1,377.97 * 55
X = ($1,377.97 * 55) / 4 = $18,947.09

Instructions

Problem Text

A local travel agent, who is offering a special package for groups of students interested in taking a spring break trip to Mexico, has recruited John, a senior, to lead a group. John is told that a group of 4 can purchase a vacation package, including airfare and accommodations, for $1,377.97. The travel agent has also offered an additional 15% discount for groups of 40 or more. John's group has 55 people. As a group, how much do they have to pay?

Control Panel

Rev. Back Play FF Exit

Initial Amount

4 Students **55 Students**

$1,377.97 = **$18,947.09**

4X = $1,377.97 * 55

X = ($1,377.97 * 55) / 4 = $18,947.09

Discount

$18,947.09 * 15% = $2,842.06

Final Amount

$18,947.09 - $2,842.06 = $16,105.03