Jl. of Educational Multimedia and Hypermedia (2004) 13(1), 23-46

The Effect of Audio and Animation in Multimedia Instruction

CAROL KOROGHLANIAN
University of Wisconsin-Eau Claire
USA
koroghcm@uwec.edu

JAMES D. KLEIN
Arizona State University
USA
james.klein@asu.edu

This study investigated the effects of audio, animation, and spatial ability in a multimedia computer program for high school biology. Participants completed a multimedia program that presented content by way of text or audio with lean text. In addition, several instructional sequences were presented either with static illustrations or animations. The study examined the effects of instructional mode (text vs. audio), illustration mode (static illustration vs. animation), and spatial ability (low vs. high) on practice and posttest achievement, attitude and time. Results indicated that spatial ability was significantly related to practice achievement and attitude. Participants with high spatial ability performed better on the practice items than those with low spatial ability. Participants with low spatial ability responded more positively than those with high spatial ability to attitude items concerning concentration, interest, and amount of invested mental effort. Findings also revealed that participants who received animation spent significantly more time on the program than those who received static illustrations. Implications for the design of multimedia are discussed.

Computer-based instruction (CBI) containing multimedia is increasingly being used as an adjunct to traditional classroom instruction. Advances in CBI and web-based technology make it possible for a designer to include features such as full color, illustrations, audio, animation, and video. However, little research exists to support the notion that adding multimedia features to instruction improves learning and performance. The current study investigated the impact of using two multimedia features—audio and animation—on learning from CBI.

Audio has usually been added to CBI as an afterthought to gain attention and increase motivation with sound effects, musical interludes, and congratulatory phrases. When audio has been added as voiced material, it has often occurred as redundant reading of screen text. Little research has been conducted to provide the instructional designer with guidelines for incorporating audio into CBI to promote learning. When guidelines do exist, they are frequently general and not always based on research (Barron, 1995). Indeed, little research exists to support the notion that adding audio to CBI can improve learning.

Most early research on audio focused on the use of redundant audio (i.e., a voice reading word for word the text that is printed on a page or displayed visually). Hartman (1961) summarized and evaluated early audiovisual studies that looked at audio-print and print only presentation for instruction. He concluded that redundant audio-print instruction was more effective than either audio or print alone. In Hartman's review, most of the studies with positive effects for redundant audio involved grade school children. Hartman himself suggested that the limited reading skills of the children in the studies might have produced the audio-print superiority results. Research findings since Hartman's review are contradictory with some studies indicating that text plus audio is more effective than either alone (Enerson & Tumey, 1984; Hartman, 1961; Lauret, 1998; Menne & Menne, 1972; Nasser & McEwen, 1976). Other studies have found no advantage for audio-print over print alone (Barron & Atkins, 1994; Barron & Kysilka, 1993; Barton & Dwyer, 1987; Furnham, Gunter, & Green, 1990; Koroghlanian & Sullivan, 2000; Nugent, 1982; Rehaag & Szabo, 1995; Shih & Alessi, 1996; Van Mondfrans & Travers, 1964).

These contradictory results can be explained by Paivio's dual coding theory, which proposed that two separate systems are involved in cognition, one for verbal information and another for image formation (Paivio, 1986). In Paivio's view, spoken and written language are both verbal information and are encoded into verbal representations (Clark & Paivio, 1991). In terms of dual coding theory, redundant audio is single channel verbal information and would not be expected to increase learning.

While most audio-text research of the last 40 years has looked at the use of text-redundant audio where the audio is identical to the text and both are presented simultaneously, three studies (Barron & Atkins, 1994; Barron & Kysilka, 1993; Koroghlanian & Sullivan, 2000) have included text density as a treatment variable. These studies found no difference in achievement between full text-no audio and lean text-full audio. All three studies incorporated graphics to explain and illustrate concepts and ideas. Reducing the amount of text on a screen leaves more area available for graphics and labeled illustrations, which are necessary tools for teaching certain types of concepts.

Research examining the integration of illustrations with text and audio is very scarce. Mousavi, Low, and Sweller (1995) and Tindall-Ford, Chandler, and Sweller (1997) looked at static graphics, text and audio and found that shifting a portion of the text to the audio channel improved learning. A related line of research has looked at the integration of illustrations with text. Several studies have examined traditional textbook design where the text refers to a separate and physically distant illustration versus an integrated textbook design with text and illustration in close proximity (Chandler & Sweller, 1991, 1992; Kalyuga, Chandler, & Sweller, 1998; Mayer, 1989; Mayer & Gallini, 1990; Mayer, Steinhoff, Bower, & Mars, 1995; Purnell, Solman, & Sweller, 1991). The majority of these studies have indicated that physically including the text in the illustration and labeling the illustrations improved learning.

The text and illustration proximity result is referred to as the split attention effect. It is postulated that the mental effort required to integrate distant information sources leaves fewer cognitive resources available for learning the information contained within the sources. By physically integrating the information sources, the mental integration effort is reduced, freeing more cognitive resources for learning the information (Chandler & Sweller, 1991).

Similarly, when integration of diverse elements is crucial for learning, as is often the case for scientific or technical illustrations and accompanying text, audio-illustration instruction may decrease the burden on working memory that text-illustration instruction might impose. By distributing the total information between the verbal and visual channels, the burden on working memory for any one channel is lower; this expansion of working memory is an intrinsic tenant of cognitive load theory (Mousavi et al., 1995; Tindall-Ford et al., 1997).

Cognitive load theory is based on the idea that people have a limited working memory (Miller, 1956) and an enormous long-term memory (Chase & Simon, 1973). This cognitive architecture implies that the biggest obstacle to effective learning is working memory's limited capacity. Cognitive load theorists seek techniques to increase working memory by reducing cognitive load, which in turn should result in improved instructional design, learning efficiency, and effectiveness. In terms of audio and text, Baddeley (1992) theorized that working memory consists of two independent processing systems, a visual-spatial sketchpad for visual information and a phonological loop for verbal information, with a central executive coordinating the two systems. If this is the case, then working memory can be expanded by presenting complementary information to both processing systems simultaneously.

One way to reduce cognitive load and increase working memory is by physically integrating disparate information such as illustrations and text. In instructional materials, illustrations and text often represent two sources of information, each unintelligible in isolation until mentally integrated into a unified whole. This is particularly true for math, science, and technology content where illustrations are heavily used for instructional purposes (Hegarty & Just, 1989; Mayer, 1994).

While most illustration and text studies have used print media, these research results may be applicable to multimedia CBI and could provide guidance for the design of instruction in that medium. In fact, multimedia CBI allows additional features of illustrations to be exploited for instruction, such as pop-up labels and explanations as well as animations which can be conceptualized as a sequence of illustrations displayed in a timed sequence to provide the illusion of movement or change.

Audio combined with animation is a relatively new research field that evolved from research regarding the effective integration of text with illustrations. Some studies investigating animation and audio narration have found simultaneous audio narration and animation more effective than either alone or nonconcurrent audio narration and animation (i.e., audio narration followed by animation or animation followed by audio narration) (Mayer & Anderson, 1991, 1992; Mayer & Sims, 1994), while another found no difference in achievement (Childress, 1995). Other studies examining various combinations of animation, static illustration, text, and audio have produced mixed findings (Atlas, Cornett, Lane & Napier, 1997; Lee, 1996; Lee, 1997; Moreno & Mayer, 1998; Mayer & Moreno, 1999; Palmiter & Elkerton, 1993; Wilson, 1998).

The efficacy of animation with audio is mixed and may depend upon the function of the animation, type of learning measured and characteristics of the learners. Indeed, Mayer and Sims (1994) found a strong link between spatial ability and animation efficacy with higher spatial ability learners benefiting more from animation with audio narration than lower spatial ability learners.

The role spatial ability plays in learners' interpretation and comprehension of animated and static illustrations is unclear. A few studies have investigated spatial ability in conjunction with animated or static illustrations (Blake, 1977; Hays, 1996; Hegarty & Sims, 1994; Hegarty & Steinhoff, 1997; Mayer & Sims, 1994; Winn, 1982). Some of these studies found animation beneficial to low spatial ability learners (Blake, 1977; Hays, 1996) while the others found animation more beneficial to high spatial ability learners (Hegarty & Sims, 1994; Hegarty & Steinhoff, 1997; Mayer & Sims, 1994).

Mayer and Sims (1994) looked at animation, audio, and spatial ability. They did not examine static illustrations but rather investigated the synchronization of audio with animation and measured problem solving achievement for high and low spatial ability adult participants using scientific explanative content. They found that high spatial ability participants achieved more as measured by transfer/problem solving items than low spatial ability participants. Mayer and Sims explained these findings in terms of dual coding theory; high spatial ability learners can more easily build a visual representation and therefore invest greater cognitive resources towards building referential connections. Low spatial ability learners spend more time and effort building the visual representation leaving fewer cognitive resources available for building referential connections (Mayer & Sims, 1994). Since the Mayer and Sims study did not include a static condition, it is impossible to say whether animation somewhat offset low spatial ability.

The research on audio with illustrations and animations may be broadly summarized into four main points.

1. Audio research has generally indicated no increase in achievement for redundant audio-text versus text only instruction. Some research suggests text can be shifted from the screen to the audio channel without decreasing achievement, but the number of such studies is very small.
2. Audio-illustration research is scarce. What does exist suggests moving text from the screen to the audio channel improves learning if the text information must be integrated with the illustration for comprehension.
3. Audio-animation research is contradictory. The existing research suggests the effectiveness of audio-animation in instruction may depend on the function of the animation, the type of learning measured, and the characteristics of the learners.
4. Research suggests that spatial visualization ability may be related to the interpretation and inferring of motion from diagrams and illustrations of a scientific or technical nature. Low and high spatial ability learners may differentially benefit from animation.

The present study investigated the effects of audio, animation, and spatial ability using a multimedia CBI program concerning a scientific process. The major independent variables were instructional mode (text versus audio), illustration mode (static versus animated) and spatial ability (low versus high). Dependent variables included achievement, attitude, total time-in-program, time-in-instruction, and time-in-practice.

Instructional modes consisted of two versions, text and audio. In the text version, the instruction was presented as screen text, while in the audio version, the instruction was presented as spoken words with limited screen text. The spoken words of the audio version matched the text of the text version.

There were two versions of the illustration mode, static and animated. The static version consisted of a graphic depicting the scientific process with no visual movement to show the process in operation, while the animated version showed the process with visual movement to demonstrate the process in operation.

Spatial ability represented another variable in this study. All participants were classified as low or high spatial ability based on learners' scores on the Paper Folding Test (Ekstrom, French, & Harmon, 1976).

Achievement was measured by practice item scores and a posttest. Attitude survey items measured interest, motivation, and perceived amount of invested mental effort. Time spent on instruction and on practice items, as well as the total time spent in the program, was measured and recorded for each participant.

METHOD

Participants

One hundred and nine students from an urban high school biology course participated in this study. Participants were blocked by spatial ability and randomly assigned to one of four treatments (Text-Static Illustration, Audio-Static Illustration, Text-Animation, Audio-Animation).

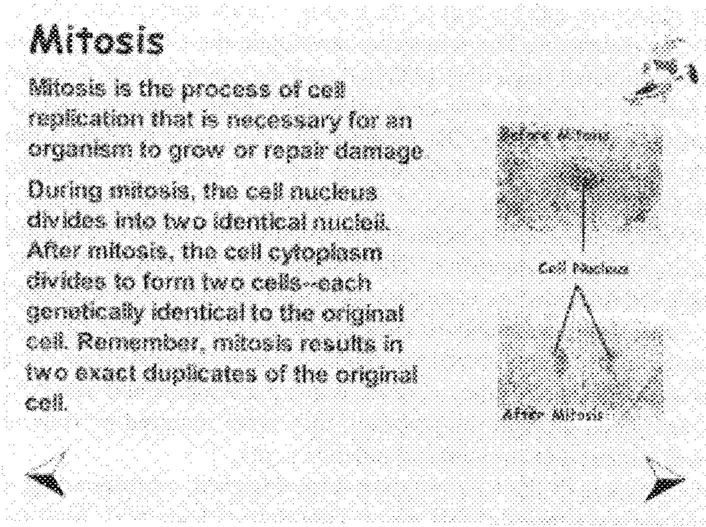
Materials

A computer-based instruction (CBI) program, *The Cell Cycle*, constituted the instructional materials for this study. The CBI program was based upon the objectives and content of the biology course and covered the topics of mitosis and meiosis. The researcher developed the CBI program and incorporated animated sequences obtained with permission from *The Biology Project* website (The University of Arizona, http://www.biology.arizona.edu/). Information sources for the CBI included the classroom teacher, a biology textbook (Wallace, 1997), and the biology web site previously mentioned.

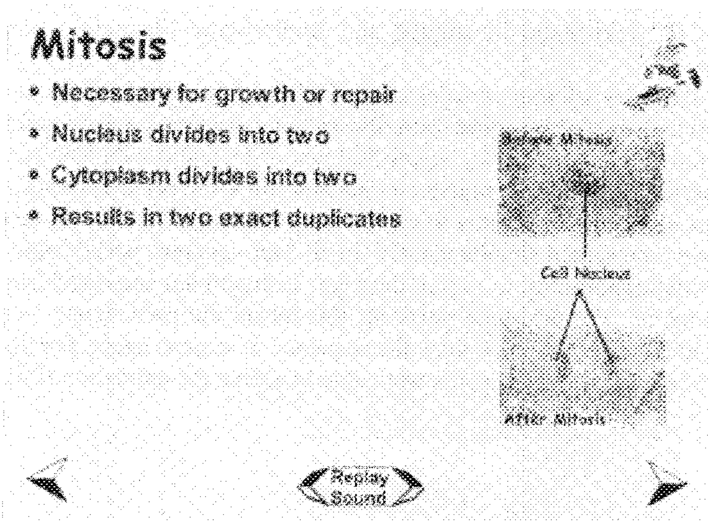
The CBI program consisted of five sections: (a) an introduction, (b) a review of prerequisite knowledge, (c) instruction on mitosis, (d) instruction on meiosis, and (e) a conclusion. The instructional portions of the CBI program included information, examples, activities, practice with feedback, and review. Participants were able to view previous screens within each section of the program and were offered the choice of viewing eight instructional sequences as many times as they choose. While participants had some choices as they navigated through the program, they were required to view all instructional screens and complete all practice items.

Four versions of the CBI program were designed, corresponding to the four treatment conditions of this study: Text-Static Illustration, Audio-Static Illustration, Text-Animation, and Audio-Animation. The CBI program took approximately 55 minutes to complete. Figure 1 shows sample instructional screens for Text-Static Illustration and Audio-Static Illustration program versions.

In summary, all four versions of the CBI program contained the same illustrations delivered by either static illustrations or animations, the same information delivered either by screen full text or by audio with lean text, the same activities, the same practice items and the same reviews. Both audio versions allowed participants to replay the audio and in both, the text and audio were presented simultaneously. In all four versions, participants were offered the opportunity to replay the eight instructional sequences suitable for animation. For the static versions, this meant the participants were allowed, at their discretion, to view the sequence of screens again. For the animated versions, the participants were able to replay the animation if they so chose.



Text-Static Illustration Version Instructional Screen



Corresponding Audio-Static Illustration Version Instructional Screen

Figure 1. Sample instructional screens for text-static illustration and audio-static illustration program versions

Procedures

A spatial ability test was administered to the participants approximately one week prior to the study. Scores from all participants were ranked and a median split was used to classify participants as high or low spatial ability. For this study, 103 participants completed the Paper Folding Test with a median score of 12 out of a total possible score of 20. Participant assignment to each of the four treatments was counterbalanced by spatial ability. On the first day of the study, participants received instructions from the researcher and worked through the CBI program. On the second day of the study, participants completed the CBI program, an attitude survey and a posttest. All events occurred during normally scheduled class time.

The participant information sheet, spatial ability test, attitude survey, and posttest were paper-based. The participant information sheet and spatial ability test took approximately 20 minutes to administer while the treatment, attitude survey and posttest required 55-90 minutes over two days. All events occurred during normally scheduled class time. All participants wore headphones while working through the CBI program.

Criterion Measures

There were three criterion measures employed in this study: an attitude survey, practice item results, and posttest scores. Enroute time data was also collected and examined.

A 10-item Likert-type attitude survey (5 point scale from Strongly Agree to Strongly Disagree) was administered prior to the posttest. The Cronbach Alpha reliability of the attitude survey was .79. The attitude survey included items to measure interest, motivation, and amount of invested mental effort (AIME).

AIME was measured using three items similar to those developed by Salomon (1984). These three items concerned the amount of effort and concentration expended by the participants as well as how well they thought they understood the material. These items were, "The information presented in this program was easy to understand," "I concentrated on learning the material throughout the entire program," and "I tried hard to understand the information presented in the program." The Cronbach Alpha reliability of the AIME measure was .57.

Achievement was measured by a 27-item posttest. The posttest included 15 selected response and 12-constructed response items with each item worth one to three points for a total of 30 possible points. The reliability of the posttest was .82.

The practice items were similar in form and content to the posttest and included 17 selected response and 11-constructed response items with each item worth one to three points for a total of 30 possible points. The reliability of the practice items was .70.

Enroute time data, including total time-in-program, time-in-instruction, and time-in-practice were collected in a data file. For the purposes of this study, total time-in-program was defined as the time elapsed between the participant entering and exiting the CBI program; time–in-practice was defined as the time the participants spent completing the practice items within the CBI program; and time-in-instruction was defined as the difference between total time-in-program and time-in-practice. Since the practice screens were identical in all four treatments, time-in-instruction represented the time participants spent within treatments (Text versus Audio, Static illustration versus Animation).

Design and Data Analysis

This study was a posttest only control group design. It was a 2 (text versus audio) x 2 (static illustration versus animation) x 2 (high versus low spatial ability) factorial design. Instructional mode, illustration mode and spatial ability were between-participants variables.

A 2 x 2 x 2 analysis of variance (ANOVA) was conducted for practice scores, posttest scores, attitude responses, and enroute time data. A significance level of .05 was set for all statistical tests.

RESULTS

While 109 participants completed the CBI program, attitude survey, and posttest, 103 participants completed the spatial ability measure. Due to network failures and computer problems during the first day of the study, enroute time data for 25 participants was lost or incomplete. Therefore, complete enroute time data were collected for 84 participants.

Practice Achievement

Means and standard deviations for practice achievement are reported in Table 1. The overall mean practice score was 15.45 (SD = 4.41) or 51.5%. A 2 x 2 x 2 ANOVA performed on the practice data indicated that participants with high spatial ability (M = 16.44, SD = 4.09) performed significantly better on the practice items than those with low spatial ability (M = 14.41, SD = 4.54), F (1, 76) = 4.041, p = .048, ES = .46. No other significant results were found for practice achievement.

Table 1

Practice Achievement Means and Standard Deviations

Instructional Mode		Ilustration Mode		Spatial Ability	
Text		Static		Low	
M	15.95	*M*	15.59	*M*	14.41
SD	4.47	*SD*	4.68	*SD*	4.54
N	44	*N*	41	*N*	41
Audio		Animation		High	
M	14.90	*M*	15.33	*M*	16.44
SD	4.33	*SD*	4.19	*SD*	4.09
N	40	*N*	43	*N*	43

Note. 30 points possible with a minimum score of 4 and a maximum score of 24.

Posttest Achievement

Mean scores and standard deviations for posttest achievement are reported in Table 2. The overall mean for the posttest was 17.34 (SD = 5.50) or 57.8%. A 2 x 2 x 2 ANOVA performed on the posttest data revealed no significant results for instructional mode (Text, M = 17.76; Audio, M = 16.92), illustration mode (Static, M = 17.93, Animation, M = 16.64), or spatial ability (Low, M = 16.48; High, M = 18.15).

Table 2

Posttest Achievement Means and Standard Deviations

Instructional Mode		Illustration Mode		Spatial Ability	
Text		Static		Low	
M	17.76	*M*	17.93	*M*	16.48
SD	5.68	*SD*	5.64	*SD*	5.61
N	51	*N*	56	*N*	50
Audio		Animation		High	
M	16.92	*M*	16.64	*M*	18.15
SD	5.38	*SD*	5.30	*SD*	5.31
N	52	*N*	47	*N*	53

Note. 30 points possible with a minimum score of 5 and a maximum score of 28.

Student Attitude

There were 10 items on the attitude survey and participants responded to each on a 5-point Likert scale ranging from 5 (strongly agree) to 1 (strongly disagree). In general, participants found the program interesting ($M = 3.96$, $SD = 0.75$), easy to understand ($M = 4.03$, $SD = 0.79$), and motivational ($M = 3.93$, $SD = 0.92$). Participants indicated they tried hard to understand the information presented in the program ($M = 4.11$, $SD = 0.96$) and believed the pictures/animations made the explanations easier to understand ($M = 4.28$, $SD = 0.78$). A 2 x 2 x 2 ANOVA was performed on each of the attitude items (Stevens, 1996). Results indicated that illustration mode had a significant effect for the item, "The information in this program was easy to understand," $F(1, 95) = 4.029$, $p = .048$, $ES = .38$. Participants in the Animation treatment ($M = 4.19$) responded more positively to this item than those in the Static illustration treatment ($M = 3.89$).

Results also revealed that spatial ability was significantly related to three attitude items. For all three items, participants with low spatial ability responded more positively than their high spatial ability counterparts for the following items: (a) "Learning on a computer keeps me motivated and interested in the material," $F(1, 95) = 5.868$, $p = .017$, $ES = .48$, (b) "I would like to learn more about biology," $F(1, 95) = 6.028$, $p = .016$, $ES = .50$, and (c) "I concentrated on learning the material throughout the entire program," $F(1, 95) = 5.412$, $p = .022$, $ES = .45$.

Three of the attitude items were used to measure the Amount of Invested Mental Effort (AIME). Means and standard deviations for AIME by instructional mode, illustration mode and spatial ability appear in Table 3. The overall mean for AIME was 4.04 ($SD = 0.65$). A 2 x 2 x 2 ANOVA performed on the AIME data revealed a significant difference between Low and High spatial ability participants, $F(1, 95) = 7.065$, $p = .009$, $ES = .52$, with low spatial ability participants ($M = 4.21$, $SD = 0.58$) indicating greater AIME than high spatial ability participants ($M = 3.87$, $SD = 0.67$). No other significant results were found.

Table 3

Amount of Invested Mental Effort Means and Standard Deviations

Instructional Mode		Illustration Mode		Spatial Ability	
Text		Static		Low	
M	3.96	*M*	3.98	*M*	4.21
SD	0.66	*SD*	0.71	*SD*	0.58
N	51	*N*	56	*N*	50
Audio		Animation		High	
M	4.11	*M*	4.11	*M*	3.87
SD	0.63	*SD*	0.56	*SD*	0.67
N	52	*N*	47	*N*	53

Note. The three items comprising the AIME were summed and divided by 3.

Time

Means and standard deviations for total Time-in-Program by instructional mode, illustration mode, and spatial ability are presented in Table 4. The overall total Time-in-Program was 53.30 minutes (SD = 10.87). A 2 x 2 x 2 ANOVA performed on the Total Time-in-Program data indicated that participants in the Animation treatment (M = 56.35, SD = 9.48) spent significantly more Time-in-Program than those in the Static illustration (M = 50.40, SD = 11.41), F (1, 72) = 5.866, p = .018, ES = .55. No other significant results for Time-in-Program were found.

Table 4

Time-in-Program Means and Standard Deviations

Instructional Mode		Illustration Mode		Spatial Ability	
Text		Static		Low	
M	*54.46*	*M*	50.40	*M*	53.32
SD	10.42	*SD*	11.41	*SD*	13.05
N	44	*N*	41	*N*	37
Audio		Animation		High	
M	51.89	*M*	56.35	*M*	53.29
SD	11.38	*SD*	9.48	*SD*	8.73
N	36	*N*	39	*N*	43

Note. Time-in-Program is reported in minutes.

Means and standard deviations for Time-in-Instruction by instructional mode, illustration mode and spatial ability are presented in Table 5. The overall Time-in-Instruction was 41.66 minutes (SD = 8.45). A 2 x 2 x 2 ANOVA performed on the Time-in-Instruction data revealed that participants in the Animation illustration mode (M = 44.30, SD = 7.17) spent significantly more Time-in-Instruction than those in the Static illustration mode (M = 39.16, SD = 8.89), F (1, 72) = 8.100, p = .006, ES = .61. There were no other significant results for Time-in-Instruction.

Table 5

Time-in-Instruction Means and Standard Deviations

Instructional Mode		Illustration Mode		Spatial Ability	
Text		Static		Low	
M	*42.51*	*M*	39.16	*M*	42.17
SD	7.67	*SD*	8.89	*SD*	10.25
N	44	*N*	41	*N*	37
Audio		Animation		High	
M	40.63	*M*	44.30	*M*	41.23
SD	9.32	*SD*	7.17	*SD*	6.62
N	36	*N*	39	*N*	43

Note. Time-in-Instruction is reported in minutes.

The overall Time-in-Practice was 11.15 minutes ($SD = 2.84$). A 2 x 2 x 2 ANOVA performed on the Time-in-Practice data indicated that there were no significant results for instructional mode, illustration mode or spatial ability.

DISCUSSION

The purpose of this study was to investigate the effects of audio, animation, and spatial ability in a multimedia program for biology. Participants completed a CBI that presented instructional material by way of text or audio with lean text. In addition, eight instructional sequences were presented either with static illustrations or animations. All participants were classified as either low or high spatial ability and assigned to one of the four treatments (Text-Static Illustration, Audio-Static Illustration, Text-Animation, or Audio-Animation). The study examined the effects of instructional mode (text vs. audio), illustration mode (static illustration vs. animation) and spatial ability (low vs. high) on practice and posttest achievement, attitude, and time.

This study found no main effect results for posttest achievement. The posttest results for Illustration mode are consistent with the half of all previous animation research findings indicating no benefit for animation over static illustration (Caraballo, 1985; Caraballo-Rios, 1985; ChanLin, 1998; Lai, 1998; Park, 1998; Peters & Daiker, 1982; Reed, 1985; Rieber, 1989; Rieber & Hannafin, 1988; Robison, 1996; Schnotz, Boeckheler, & Grzondziel, 1999; Spangler, 1994; Spotts & Dwyer, 1996; Towers, 1994; Wilson, 1998; Wright, Milroy, & Lickorish, 1999). Additionally, the finding for Instructional mode indicating no loss of achievement when moving some text from the screen to audio is also consistent with previous research (Barron & Atkins, 1994; Barron & Kysilka, 1993; Koroghlanian & Sullivan, 2000). However, the posttest achievement results for spatial ability are not consistent with previous research. The present study does not support previous findings that animation benefits low spatial ability learners in terms of achievement (Blake, 1977; Hays, 1996) nor does it support the finding that high spatial ability learners benefit more from animation than low spatial ability learners (Mayer & Sims, 1994). It should be noted that the Hays (1996) study examined gains scores while the Mayer and Sims (1994) study examined achievement in terms of transfer and problem solving. The current study did not examine gains scores nor was transfer and problem solving the sole type of achievement measured.

Closer examination of the posttest data reveals some interesting patterns. High Spatial Ability participants achieved more than Low Spatial Ability participants in the Static Illustration mode for both Text and Audio. This result was expected as high spatial ability learners probably build visual representations more easily than low spatial ability learners and thus have more cognitive resources available for integrating the verbal information with static illustrations (Hays, 1996; Hegarty & Sims, 1994; Hegarty & Steinhoff, 1997; Mayer & Sims, 1994).

Looking at the posttest achievement for Animation mode, High and Low Spatial Ability participants achieved comparably but High Spatial Ability participants achieved less in the Animation mode than High Spatial Ability participants in the Static mode. This result contradicts the explanation offered previously for high spatial ability learners and static illustrations. If high spatial ability learners build visual representations more easily than low spatial ability learners, it seems logical that providing an animation instead of static illustrations would allow easier and faster building of visual representations; thus leaving more cognitive resources available for integrating verbal information with the animations. However, as Hays (1996) and Schnotz et al. (1999) observed, participants with higher cognitive abilities must use those abilities for learning advantages to occur.

Alternatively, high spatial ability participants in this study may have found the combination of animation with text or audio distracting and failed to focus on the pertinent visual and verbal information. Some recent research studies have suggested animations can increase rather than decrease cognitive load in some circumstances for some learners (Kalyuga et al., 1998; Lowe, 1999; Wright et al., 1999).

In summary, in terms of posttest achievement, the inclusion of animation or audio was neither beneficial nor detrimental for Low Spatial Ability participants. However, the inclusion of animation and possibly the inclusion of audio appeared to be detrimental for High Spatial Ability participants.

Results for practice achievement indicated that High Spatial Ability participants achieved more than Low Spatial Ability participants. Again, this result was expected as high spatial ability learners probably build visual representations more easily than low spatial ability learners, freeing cognitive resources for integrating verbal and visual information (Hays, 1996; Hegarty & Sims, 1994; Hegarty & Steinhoff, 1997; Mayer & Sims, 1994).

Closer examination of the practice data reveals a pattern similar to that found in the posttest achievement data. High spatial ability participants achieved less in the Animation than in the Static mode and within the Animation mode, participants in the Audio treatment achieved less than those in the Text treatment. Again, the inclusion of animation or audio was neither beneficial nor detrimental for Low Spatial Ability participants but High Spatial Ability participants may have found animation or audio detrimental to achievement. While the high spatial ability learners have more cognitive resources available, they may fail to put forth the effort necessary to integrate the verbal and visual information as well as they might (Hays, 1996; Schnotz et al., 1999). This lack of effort might be related to their perception of the nature and role of animation and audio in media (Salomon, 1984).

Participants in general, responded favorably to the CBI program. There was one item for which a difference by Illustration mode was significant, "The information in this program was easy to understand," with participants in the Animation treatment responding more positively than those in the Static Illustration treatment. This may represent a perception on the part of the participants that animation is an easier to understand medium than static illustrations, somewhat similar to Salomon's (1984) findings for television and print. Other studies examining attitudinal differences between audio and text (Barron & Atkins, 1994; Barron & Kysilka, 1993, Koroghlanian & Sullivan, 2000; Rehaag & Szabo, 1995) or between animation and static illustrations (Kim, 1998; Szabo & Pookay, 1996) have generally found no attitudinal differences.

In all three items for which significant differences were found by spatial ability, the Low Spatial Ability participants responded more positively than High Spatial Ability participants. The items concerned concentration and interest (Learning on a computer keeps me motivated and interested in the material, I would like to learn more about biology, I concentrated on learning the material throughout the program) and the responses may be a reflection of High Spatial Ability participants perceiving themselves to be efficacious and the material to be easy. In contrast, Low Spatial Ability participants may have realized that the material was difficult and the support provided may have increased their interest and concentration (Salomon, 1984).

Not surprisingly, AIME results differed by spatial ability with Low Spatial Ability participants reporting greater AIME than High Spatial Ability participants. Again, low spatial ability learners have to try harder to visualize processes presented through static illustrations or animations, so naturally they perceive a greater amount of mental effort than high spatial ability learners. The largest differences in AIME between High and Low Spatial Ability participants occurred for Audio-Static Illustration and Audio-Animation treatments. This provides further evidence for the idea that providing support to learners who do not need support (in this case, high spatial ability learners) may reduce demands on the learner to perform mental processes for themselves that are essential to learning and achievement (Schnotz, et al., 1999).

Since the practice screens were identical in all treatments, it was not surprising there were no differences in terms of Time-in-Practice. However, this finding is somewhat contradictory to findings of Rieber, who in a number of studies found animation as opposed to static illustrations or no illustrations can shorten the amount of time necessary to answer posttest items (Rieber, 1989, 1990, 1991; Rieber, Boyce, & Assad, 1990). Rieber postulated this time difference was probably a result of more efficient storage and retrieval of information to and from long term memory followed by efficient reconstruction of the information in short term memory. Although this study looked at Time-in-Practice rather than posttest, a similar result might have been expected for practice but was not found.

While significant differences for total Time-in-Program and Time-in-Instruction by illustration mode were found, with the Animation mode taking five to six minutes longer than Static Illustration mode, this difference was expected and supports the finding of Baek and Layne (1988). The eight Animated instructional sequences took 9 minutes and 20 seconds to play assuming no sequence was replayed by the participant. The corresponding eight Static Illustration instructional sequences should have taken less time for the participants to complete as they need not have waited for an animation to finish before proceeding.

More interesting findings are buried in the time data. Participants spent approximately the same amount of Time-in-Program for the Text as the Audio mode (54.56 minutes for Text, 51.89 minutes for Audio). Based on previous studies (Barron & Kysilka, 1993; Koroghlanian & Sullivan, 2000) and common sense, one would have expected the Audio to take longer than the Text and thus have a lower learning efficiency. It should be noted that 41 minutes was the shortest possible Time-in-Program for the Audio-Static Illustration mode while 30 minutes was the shortest possible Time-in-Program for the Text-Static Illustration mode. The corresponding Animation versions took an additional five minutes. These shortest possible times suggest Text participants would have been expected to finish 10 to 15 minutes sooner than the Audio participants, but this was not the case. Perhaps participants in the Text mode spent those 10 to 15 minutes rereading the text, examining the static illustration/animation or trying to integrate the text information with the static illustration/animation. Investigating the reasons for this time equalization between Text and Audio mode might prove interesting.

The results of this study supports previous research that suggests moving some text from the screen to audio neither hinders nor improves learning. This finding has important implications for multimedia development. If screen "real estate" is needed for something other than instructional text, which is especially true for simulations and concepts difficult to explain with words alone, then text can be moved from the screen to audio with no loss in achievement. This is an important and useful instructional technique for instructional designers to consider, especially when designing materials with scientific or technical content.

The implications of this study are less straightforward in terms of animation. Animation did not improve learning for this content and age group. Animation did take more instructional time than static illustrations with no corresponding improvement in achievement or difference in attitude. Whether to include animation or not in multimedia CBI programs is still a matter of instinct, not research, and the final decision may be dictated by pragmatic concerns such as budget or time.

While this study was conducted with computer-based instructional materials; the results have wider implications for multimedia instruction in general. Web-based instruction, for example, increasingly incorporates multimedia attributes such as audio and animation. The incorporation of these attributes should be based on instructional design principles and research to ensure effective and efficient instruction.

Several avenues of future research are suggested by the findings of this study. Although spatial ability appeared to be related to some aspects of this study, it is possible that other learner characteristics might have a more direct relationship. Perhaps examination of reading ability, self-efficacy, and general intelligence, as well as spatial ability, would provide insight into the processes involved in learning from illustrations and animations. It may also be more fruitful in the future to use ability measures as a covariate rather than a blocking variable or to include in the data analysis only those participants scoring in the top and bottom quarter of any learner characteristic. Including a large number of individuals, who score in the middle and are arbitrarily assigned to high or low ability groupings when in fact their abilities are very similar, may obscure the details necessary to see significant differences. If animation, for example, is postulated to principally benefit learners with low abilities, this benefit may be obscured by the bulk of learners of intermediate ability lumped into the low ability category by a median split.

Another area that warrants further investigation is the physical combination of audio and animation. Some researchers might argue that the present study did not minimize the split attention effect and thereby did not optimize the instruction or research conditions. Future research could examine superimposing text on the illustrations and animations as well as utilizing audio only with illustrations or animations followed by text at the end of the sequence. These sequencing and layout situations would tend to minimize the split attention effect and might clarify research results and subsequent instructional design decision making.

One puzzling and fascinating result of the present study concerns the activities and mental processes of the participants. Participants spent the same amount of time in the Text as in the Audio versions although participants in the Text version would have been expected to finish sooner than those in the Audio version. Interposing questions during the instruction or formally observing participants as they work through the CBI program might provide information of use and interest to the researcher and instructional designer.

Further research into text density and structure would be valuable to instructional designers designing both traditional CBI and web-based instruction. Reducing instructional screen text while providing the majority of instruction through an audio track, is an extremely useful technique in situations with highly complex processes or simulations where there is a need to maximize screen space for nontext purposes. Research into the amount of text required when text is combined with audio and the manner in which that text should be structured and presented warrants further investigation.

Audio and animation are powerful tools for the instructional designer. Deciding when and how to use these tools is an important field of inquiry that deserves more attention and effort.

References

Atlas, R., Cornett, L., Lane, D.M., & Napier, H.A. (1997). The use of animation in software training: Pitfalls and benefits. In M. A. Quinones & A. Ehrenstein (Eds.), *Training for a rapidly changing workplace: Applications of psychological research* (pp. 281-302). Washington, DC: American Psychological Association.

Baddeley, A. (1992). Working memory. *Science, 255*, 556-559.

Baek, Y.K., & Layne, B.H. (1988). Color, graphics, and animation in a computer-assisted learning tutorial lesson. *Journal of Computer Based Instruction, 15*(4), 131-135.

Barron, A. (1995). Digital audio in multimedia. *Educational Media International, 32*(4), 190-193.

Barron, A.E., & Atkins, D. (1994). Audio instruction in multimedia education: Is textual redundancy important? *Journal of Educational Multimedia and Hypermedia, 3*(3-4), 295-306.

Barron, A.E., & Kysilka, M.L. (1993). The effectiveness of digital audio in computer-based training. *Journal of Research on Computing in Education, 25*(3), 277-289.

Barton, E.A., & Dwyer, F.M. (1987). The effect of audio redundancy on students' ability to profit from printed-verbal/visualized instruction. *International Journal of Instructional Media, 14*(2), 93-98.

Blake, T. (1977). Motion in instructional media: Some subject-display mode interactions. *Perceptual and Motor Skills, 44,* 975-985.

Caraballo, J.N. (1985). The effect of various visual display modes in computer-based instruction and language background upon achievement of selected educational objectives (Doctoral dissertation, The Pennsylvania State University, 1985). *Dissertation Abstracts International, 46,* 1494.

Caraballo-Rios, A.L. (1985). An experimental study to investigate the effects of computer animation on the understanding and retention of selected levels of learning outcomes (Doctoral dissertation, The Pennsylvania State University, 1985). *Dissertation Abstracts International, 46,* 1494.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*(4), 293-332.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology, 62*(2), 233-246.

ChanLin, Lih-Juan. (1998). Animation to teach students of different knowledge levels. *Journal of Instructional Psychology, 25*(3), 166-175.

Chase, W.G., & Simon, H.A. (1973). Perception in chess. *Cognitive Psychology, 4,* 55-81.

Childress, M.D. (1995). Effects of three multimedia instructional presentation formats containing animation and narration on recall and problem-solving performance (Doctoral dissertation, Virginia Polytechnic Institute and State University, 1995). *Dissertation Abstracts International, 56,* 3430.

Clark, J.M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*(3), 149-210.

Ekstrom, R.B., French, J.W., & Harmon, H.H. (1976). *Kit of factor referenced cognitive tests.* Princeton, NJ: Educational Testing Service.

Enerson, M., & Tumey, D. (1984). An experiment with speech on an electronic text service. *Videodisc and Optical Disc, 4*(4), 306-317.

Furnham, A., Gunter, B., & Green, A. (1990). Remembering science: The recall of factual information as a function of the presentation mode. *Applied Cognitive Psychology, 4,* 203-12.

Hartman, F.R. (1961). Single and multiple channel communication: A review of research and a proposed model. *AV Communication Review, 9,* 235-262.

Hays, T.A. (1996). Spatial abilities and the effects of computer animation on short-term and long-term comprehension. *Journal of Educational Computing Research, 14*(2), 139-155.

Hegarty, M., & Just, M.A. (1989). Understanding machines from text and diagrams. In H. Mandl & J. Levin (Eds.), *Knowledge acquisition from text and pictures* (pp. 171-194). Amsterdam, The Netherlands: North-Holland.

Hegarty, M., & Sims, V.K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory and Cognition, 22*(4), 411-430.

Hegarty, M., & Steinhoff, K. (1997). Individual differences in use of diagrams as external memory in mechanical reasoning. *Learning and Individual Differences, 9*(1), 19-42.

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors, 40*(1), 1-17.

Kim, H.W. (1998). Effects of animated graphics of plate tectonics on students' performance and attitudes in multimedia computer instruction (Doctoral dissertation, The University of Texas at Austin, 1998). *Dissertation Abstracts International, 59*, 3395.

Koroghlanian, C.M., & Sullivan, H.J. (2000). Audio and text density in computer-based instruction. *Journal of Educational Computing Research, 22*(2), 217-230.

Lai, Shu-Ling. (1998). The effects of visual display on analogies using computer-based learning. *International Journal of Instructional Media, 25*(2), 151-157.

Lauret, D.T. (1998). The auditory display in interactive courseware: Moving human factors into computer education (Doctoral dissertation, Nova Southeastern University, 1998). *Dissertation Abstracts International, 59*, 2459.

Lee, H.L. (1997). The use of animation as a tool for concept learning (Doctoral dissertation, The Ohio State University, 1997). *Dissertation Abstracts International, 58*, 1669.

Lee, S. (1996). The effects of computer animation and cognitive style on the understanding and retention of scientific explanation (Doctoral dissertation, Virginia Polytechnic Institute and State University, 1996). *Dissertation Abstracts International, 57*, 4248.

Lowe, R.K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Educ*ation, *14*(2), 225-244.

Mayer, R.E. (1989). Systematic thinking fostered by illustrations in scientific text. *Journal of Educational Psychology, 81*(2), 240-246.

Mayer, R.E. (1994). Visual aids to knowledge construction: Building mental representations from pictures and words. In W. Schnotz & R. W. Kulhavy (Eds.), *Comprehension of graphics* (pp. 125-138). Amsterdam, The Netherlands: Elsevier Science.

Mayer, R.E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology, 83*(4), 484-490.

Mayer, R.E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology, 84*(4), 444-452.

Mayer, R.E., & Gallini, J.K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82*(4), 715-726.

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*(2), 312-320.

Mayer, R.E., & Sims, V.K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology, 86*(3), 389-401.

Mayer, R.E., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. *Educational Technology Research and Development, 43*(1), 31-43.

Menne, J.M., & Menne, J.W. (1972). The relative efficiency of bimodal presentation as an aid to learning. *AV Communication Review, 20*(2), 170-180.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychology Review, 63*(2), 81-97.

Moreno, R., & Mayer, R.E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology, 91*(2), 358-368.

Mousavi, S.Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology, 87*(2), 319-334.

Nasser, D.L., & McEwen, W.J. (1976). The impact of alternative media channels: Recall and involvement with messages. *AV Communication Review, 24*(3), 263-272.

Nugent, G.C. (1982). Pictures, audio, and print: Symbolic representation and effect on learning. *Educational Communication Technology Journal, 30*(3), 163-174.

Paivio, A. (1986). *Mental representations: A dual coding approach.* Oxford, England: Oxford University Press.

Palmiter, S., & Elkerton, J. (1993). Animated demonstrations for learning procedural computer-based tasks. *Human-Computer-Interaction, 8*(3), 193-216.

Park, O. (1998). Visual displays and contextual presentations in computer-based instruction. *Educational Technology, Research and Development, 46*(3), 37-50.

Peters, H.J., & Daiker, K.C. (1982). Graphics and animation as instructional tools: A case study. *Pipeline, 7,* 11-13, 57.

Purnell, K.N., Solman, R.T., & Sweller, J. (1991). The effects of technical illustrations on cognitive load. *Instructional Science, 20*(5-6), 443-462.

Rehaag, D.M., & Szabo, M. (1995, February). *An experiment on effects of redundant audio in computer based instruction on achievement, attitude, and learning time in 10th grade math.* Paper presented at the annual meeting of the Association for Educational Communications and Technology, Anaheim, CA. (ERIC Document Reproduction Service No. ED 380 123)

Reed, S.K. (1985). Effects of computer graphics on improving estimates to algebra word problems. *Journal of Educational Psychology, 77*(3), 285-298.

Rieber, L.P. (1989). The effects of computer animated elaboration strategies and practice on factual and application learning in an elementary science lesson. *Journal of Educational Computing Research, 5*(4), 431-44.

Rieber, L.P. (1990). Animation in computer-based instruction. *Educational Technology, Research and Development, 38*(1), 77-86.

Rieber, L.P. (1991). Animation, incidental learning, and continuing motivation. *Journal of Educational Psychology, 83*(3), 318-328.

Rieber, L.P., Boyce, M.J., & Assad, C. (1990). The effects of computer animation on adult learning and retrieval tasks. *Journal of Computer Based Instruction, 17*(2), 46-52.

Rieber, L.P., & Hannafin, M.J. (1988). Effects of textual and animated orienting activities and practice on learning from computer-based instruction. *Computers in the Schools, 5*(1/2), 77-89.

Robison, S.A. (1996). The effects of computer animated instruction in college trigonometry on student achievement, conceptual understanding, and student attitudes (Doctoral dissertation, The University of Nebraska—Lincoln, 1996). *Dissertation Abstracts International, 57*, 3435.

Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology, 76*(4), 647-658.

Schnotz, W., Boeckheler, J., & Grzondziel, H. (1999). Individual and cooperative learning with interactive animated pictures. *European Journal of Psychology of Education, 14*(2), 245-265.

Shih, Y.F., & Alessi, S.M. (1996). Effects of text versus voice on learning in multimedia courseware. *Journal of Educational Multimedia and Hypermedia, 5*(2), 203-218.

Spangler, R.D. (1994). The effects of computer-based animated and static graphics on learning to visualize three-dimensional objects (Doctoral dissertation, University of Kentucky, 1994). *Dissertation Abstracts International, 55*, 856.

Spotts, J., & Dwyer, F. (1996). The effect of computer-generated animation on student achievement of different types of educational objectives. *International Journal of Instructional Media, 23*(4), 365-375.

Stevens, J. (1996). *Applied Multivariate Statistics for the Social Sciences*. Mahwah, NJ: Lawrence Erlbaum.

Szabo, M., & Poohkay, B. (1996). An experimental study of animation, mathematics achievement, and attitude toward computer-assisted instruction. *Journal of Research on Computing in Education, 28*(3), 390-402.

The University of Arizona (1999). *The biology project.* [Online]. Available: http://www.biology.arizona.edu/default.html/

Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied, 3*(4), 257-287.

Towers, R.L. (1994). The effects of animated graphics and static graphics on student learning in a computer-based instructional format (Doctoral dissertation, University of Kentucky, 1994). *Dissertation Abstracts International, 55,* 1184.

Van Mondfrans, A.P., & Travers, R.M.W. (1964). Learning of redundant material presented through two sensory modalities. *Perceptual and Motor Skills, 19,* 743-751.

Wallace, R.A. (1997). *Biology, the world of life.* Menlo Park, CA: Addison Wesley Longman.

Wilson, F.S. (1998). The effect of time and level of visual enhancement in facilitating student achievement of different educational objectives (Doctoral dissertation, The Pennsylvania State University, 1998). *Dissertation Abstracts International, 59,* 2943.

Winn, W. (1982). The role of diagrammatic representation in learning sequences, identification and classification as a function of verbal and spatial ability. *Journal of Research in Science Teaching, 19*(1), 79-89.

Wright, P., Milroy, R., & Lickorish, A. (1999). Static and animated graphics in learning from interactive texts. *European Journal of Psychology of Education, 14*(2), 203-224.