INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

The Effectiveness of Animation and Narration in Computer-based Instruction

Tracy Hutcheson

**A thesis presented in partial fulfillment of the requirements for the degree of Master
of Arts**

Department of Psychology

Carleton University

Ottawa, Ontario

January, 1997

© Tracy Hutcheson, 1997

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-22083-4

The undersigned hereby recommend to
the Faculty of Graduate Studies and Research
acceptance of the thesis,

"The Effectiveness of Animation and Narration in Computer-based Instruction"

submitted by

Tracy Hutcheson, B.A. (Hon)

in partial fulfillment of the requirements
for the degree of Master of Arts

Chair, Department of Psychology

Thesis Supervisor

External Examiner

Carleton University

15th January, 1997

Abstract

Subjects were given one of two computer learning tutorials on ANCOVA, the statistical technique. Subjects performed no better on a quiz when experiencing an animation tutorial with voice narration than subjects who experienced a non-animation tutorial with no voice narration. Analyses on the quiz data, following the tutorial, revealed that both groups performed equally poor on fundamental concept questions. Subjects performed poorly on animation-spatial questions, specially designed to test the ANCOVA concepts in the animations. One of two possibilities exist, 1) that animations do not facilitate learning compared to non-animation static graphics or, 2) this study, and much of the research, is deficient in testing animations appropriately. A view is put forward that there is the possibility that animation has not been tested appropriately and that it may yet prove to facilitate learning.

Acknowledgments

I would like to thank both of my thesis advisors, Dr. Richard Dillon and Dr. Chris Herdman. Dr. Dillon's unwavering support and availability far exceed any professor's I have yet encountered. His attention to detail in writing has given me a new appreciation for the incredible hard work involved in works of this nature. Dr. Herdman's calming effect in the midst of student panic, and his intelligent ideas together with great writing skills have taught me a great deal over the past several years.

I would also like to express my gratitude to Dr. Jo Wood for her focusing ideas.

Lastly, but by far the most important, I would like to thank my partner of 12 years, Julie Makrai. Without Julie's terrific insight, patience, and occasional document production expertise the completion of this thesis would never have been possible. I cannot thank her enough for her support, love, and understanding.

Table of Contents

Acceptance Form	ii
Abstract	iii
Acknowledgments	iv
Table of Contents	v
List of Tables	vi
List of Figures	vii
List of Appendices	viii
 Introduction	 1
Background Research	1
Purpose of the Present Research	7
 Method	 10
Subjects	10
Apparatus	10
Materials	10
Tutorials	10
Quiz	13
Procedure	13
 Results	 15
Quiz Question Types	15
Analyses	19
Quiz accuracy	20
Tutorial time and quiz time	24
Page time and number of page accesses	24
 Discussion	 30
The Possibility that Animations do not Help	31
The Possibility that Animations have not been Adequately Evaluated	33
 References	 36
 Appendices	 40

List of Tables

Table 1. Percent Correct on Quiz with Standard Errors of the Mean	21
Table 2. Mean Time in Minutes to Complete the Quiz with Standard Errors of the Mean	24

List of Figures

Figure 1. Animation page.	5
Figure 2. Non-animation page.	12
Figure 3. Question 9: A spatial question	16
Figure 4. Question 20: An animation-spatial question	18
Figure 5. Percentage of people by condition who answered the question correctly for each type.	23
Figure 6. Mean time spent on page type by condition.	26
Figure 7. Number of accesses by page type by condition.	28

List of Appendices

Appendix A. Performance Quiz	40
Appendix B. Animation Narration	57
Appendix C. Mean Number of Page Accesses by Machine Type for Animation Pages	60
Appendix D. ANOVA Summary Table for Percentage Correct on the Quiz for all Accuracy Measures	61
Appendix E. Experimental Instructions	62

The Effectiveness of Animation and Narration in Computer-Based Instruction

There is support for the position that appropriately designed computer graphics facilitate learning in computer-based instruction (CBI) (Alesandrini, 1987). For example, it has been suggested that graphics may serve as a mnemonic for remembering both verbal information and concrete concepts (Paivio & Csapo, 1973; Pressley, 1976). Computer animation as a subset of computer graphics has been encouraged as a good method of presentation in CBI (Alesandrini, 1987; Bork, 1981; Caldwell, 1980). The primary goal of the current study was to determine whether animation facilitated learning as compared with static or non-animation graphics when presented in a computer-based learning tutorial.

Background Research

Computer technology continues to improve to make animation presentation in CBI an attractive method to teach adults in schools and in the workplace. The majority of the animation research with adults, on the other hand, does not show an advantage to learning when using animation graphics as compared to static (non-animation) graphics (e.g., Caraballo, A., 1985; Caraballo, J., 1985; King, 1975; Moore, Nawrocki, & Simutis, 1979; Rieber, 1989; Rieber, 1990; Rieber, Boyce & Assad, 1990; Wong, 1994).¹ Rieber's (1990) review of the literature showed that of all the studies with adults, only one (Baek & Layne, 1988) showed an effect for animation and used sound methodology.

In one typical study, Rieber, Boyce and Assad (1990) presented subjects with one of three CBI tutorials: an animation tutorial, a static graphic tutorial, and a no-graphic tutorial. The tutorials explained and demonstrated Newton's laws of

¹ For a comprehensive review of the literature see Rieber (1990) and Wong (1994).

motion. The motion in the static graphic tutorial was represented by arrows. The no-graphic tutorial was presented with textual information alone. All three tutorials presented the same text information. Immediately following the tutorial lesson, subjects were given a retention test of knowledge of Newtonian principles. Rieber et al. found no differences in performance between subjects in any of the three groups. The absence of a facilitatory effect in the Rieber et al. study, however, may have been due to the lack of direction towards pertinent information during the animations. Subjects may not have known what the important elements were to look for in the animations. Baek and Layne (1988) suggest that the use of animation must draw focus to the important features of the material being presented if learning is to be facilitated.

Baek and Layne (1988) found that animation facilitated learning when attention was guided to the most pertinent details. They presented different groups of subjects one of three CBI tutorials that contained either non-animation graphics, animation graphics, or text-alone information. Subjects were administered a pre-test one day prior to the tutorial and a post-test immediately following the tutorial session. The pre-test and post-test contained the same questions.

The learning objective was to teach the mathematical rule of average speed which was defined as the distance traveled divided by the time spent. The mathematical rule was defined in the tutorials as a way to show students the important information on which to focus. Subjects then knew what elements on which to focus in the animation examples that followed. Subjects in the animation group did better than subjects in the non-animation group, who in turn, did better than the text-alone group. The results showed a significant but very small learning effect

with a difference of 1 question correct out of 20 in the animation group (12 correct) versus subjects in the non-animation group (10.9 correct). To conclude that this 5 percent difference was significant, number of questions correct on the pre-test was used as a covariate in the analysis of the post-test data. Thus the error term used in the significance test on the post-test scores was reduced by removing the subject differences in the pre-test. The importance of a small, but statistically significant effect is questionable. In practical terms, resource cost (effort, time, and money) associated with creating animations far exceeds those used to create non-animations. The relative merits of using animation as opposed to non-animation graphics is questionable for a 5 percent improvement that brought performance up to 55 percent.

A third study relevant to the present research is Wong (1994). Wong proposed that subjects exposed to well-designed animations would do better than subjects exposed to well-designed, non-animation graphics, but that subjects exposed to poorly designed animations might actually do worse than subjects exposed to well-designed non-animated graphics. Wong developed three computer tutorials that presented information describing concepts of the analysis of covariance (ANCOVA) statistical technique. There were two animation tutorials with one designed using animation guidelines developed by Wong (1994). The second animation tutorial used animations that did not follow the guidelines but reflected common techniques used in commercial animation software. The third tutorial used non-animation graphics. A typical screen for any of the three tutorials included text information at the top and a non-animation graphic on the lower half of the screen. For an animation tutorial, an animation was set in motion by clicking

the mouse on a button with an icon that resembled a movie projector. A typical screen is shown in Figure 1.

5 of 47

To distinguish between reading scores for the two groups, we will plot a separate frequency distribution for each group. Three things are noteworthy regarding the two distributions.

1. There is a difference between means on reading performance.
(Use Animate button.)

Figure 1. Animation page.

Retention was measured by a quiz following the presentation of a tutorial. Retention was worse with poorly designed animations than with well-designed static graphics (Wong, 1994, experiment 1). It is clear from this result that using poorly designed animations is worse than using well designed non-animation graphics. Most important, Wong (1994) found that there was no retention difference between a tutorial with well-designed animations and a tutorial with equally well-designed non-animation graphics. There may be a number of reasons why Wong failed to find retention facilitation with well-designed animations.

First, some questions on the quiz may not have appropriately probed the information covered in the animations. Some quiz questions were not formulated in a form that was consistent with the spatial-relational information in the animations. To understand ANCOVA, subjects must understand the spatial relation of data points and regression lines before and after the ANCOVA technique is applied. In the animation tutorial, spatial-relational information is conveyed in graphical form through animations that show data and regression lines rotating. The questions could better target that kind of information by asking subjects in an analogous graphical form. It is helpful if the cue at the time of testing be a cue encoded at the time of acquiring that knowledge (e.g., Flexser and Tulving, 1978; Tulving and Osler, 1968; Tulving and Thompson 1973; Tulving and Wiseman, 1975; Tulving, 1978). Multiple choice text questions may not accurately assess the complex concepts taught using animations in the Wong (1994) study. Also, not every question used by Wong (1994) targeted content conveyed by the animations. For example, a question that asked which event came first when performing an ANCOVA was purely concerned with the sequence of events in the tutorial and not

with the information directly conveyed in any single animation. This question was asked in a text multiple choice form so it is incongruent with both spatial relation information and animation content. Presumably, as a result, portions of the quiz may have been successfully answered from the textual information alone without the aid of animations.

Secondly, in addition to the shortcomings in the quiz used by Wong, there were problems with the animations themselves. Subjects were not able to simultaneously watch the animation and read the text information that complimented the graphics, although they could control when the animation would occur by clicking a button. Subjects typically read the text information and then watched the animations, or vice versa. Text information did remain available on the top half of the screen during the animation. But this would draw their attention away from the animation and they would miss part of it while attending to the text. Thus, subjects did not have access to important text information when it was most appropriate, during the animations.

Purpose of the Present Research

The current study was designed to extend the Wong (1994) study in two ways.

First, voice narration was added to the animation tutorial developed by Wong (1994) in experiment 1. The narration was used as a guide or cue to information unfolding in the animation and not as a replacement for the text information that was provided with the tutorial. Consistent with the recommendation by Baek and Layne (1988), the voice narration focuses the subject's attention on important information and actions as the animation unfolds.

The narration was a very brief phrase of several words timed to point out important or key aspects of the animation at the most appropriate time.

In addition, presentation of narration with the animation is compatible with Baddeley's working memory model. Baddeley and Hitch (1974) proposed a multicomponent working memory model that consists of an attentional controller component and two slave subsystems: an articulatory loop and a visuo-spatial sketchpad. The articulatory loop is involved in the comprehension and manipulation of speech-based information and the visuo-spatial sketchpad is assumed to set up and manipulate visual images. Baddeley's approach emphasizes the role of working memory in such cognitive tasks as learning, comprehending, and reasoning (Baddeley, 1990). The relevance to the current study is that the visuo-spatial sketchpad is assumed to process and manipulate visual-spatial information in the animations, and the articulatory loop is assumed to process and manipulate the verbal information of the voice narration. The gain with narration is that learners can presumably exploit the compatibility of the two subsystems to synthesize the congruent information more effectively. The current study does not aim to test the model but is consistent with the theoretical position.

As a second change to the Wong study, the quiz in the current study included new questions presented in a visual form as well as questions presented in text form. The quiz is presented in Appendix A. Although Wong (1994) included some questions in a visual form, the questions did not target specific animations and, therefore, may not have probed concepts that the animations directly illustrated. Each question, except the first, in the present study asked information about a specific animation. The new visual questions were expected to better test

the spatial-relational information that animations provided because they were presented in a graphical form that asked information about the relation of the means and regression lines before and after covariate adjustments. The spatial-relational information was considered key in understanding the concepts about ANCOVA in the tutorial.

To summarize, the present study extends Wong's (1994) study and attempts to address the conflicting and weak findings in the literature (e.g., Baek and Layne, 1988; Rieber et al., 1990; Wong, 1994). The narration emphasizes important or relevant information to learners (Baek and Layne, 1988). In the current study, the relevant information was emphasized by voice narration as the animation was taking place. If emphasizing pertinent information to be learned is a key to animation success in learning, then the animation group in the current study should outperform the non-animation group. The animation and voice information should synthesize the information more easily since, in the Baddeley and Hitch (1974) model, they access different information processing subsystems. The questions on the quiz each targeted information that the animations provided and presented them in a spatial form that was congruent with the form of the animation presentation.

Method

Subjects

The study included a total of 40 students recruited from a graduate univariate statistics course (10 subjects) and an undergraduate statistics course (30 subjects) at Carleton University. The graduate students had not taken the portion of the course that dealt with analysis of covariance (ANCOVA) by the time they were tested in this study, and the undergraduate course stopped short of the ANCOVA technique. Volunteer students in the graduate course were motivated to participate in the study because they would cover the material in class following the experiment. The undergraduates were paid 10 dollars for their participation in the study.

Apparatus

Two Apple computers (a Macintosh IIsi and a Power Macintosh 6100/66 AV) were used to present the tutorial material. Clicking with a mouse allowed subjects to move through the tutorial. The Macintoshes recorded time spent on each screen page of the tutorial and time to complete the tutorial.

Materials

Tutorials. Subjects received one of two possible computer tutorial presentations. There was one animation tutorial and one non-animation tutorial developed using the software computer package Director (Macromedia, 1992). Both tutorials included text as well as graphics and covered statistical concepts in analysis of covariance (ANCOVA). The animation tutorial included non-animation graphics as well as animations. The non-animation condition contained only the non-animation pictures and text without animations. Both conditions had a "next"

and a "previous" button icon on each page that allowed them to move forward or backwards through the tutorial when clicked on by the mouse. There was also a "done" button on every page of both tutorials that allowed subjects to quit the tutorial at any time.

The animations were two dimensional dynamic graphics that showed spatial and or temporal changes to depict movement. Each animation was a series of discrete still graphics that were presented to the screen one at a time and quickly replaced by successive graphics until the animation was completed. The rapid presentation of discrete graphics gave the impression of perceived movement. Figures 1 and 2 show that each animation had a corresponding non-animation graphic in the non-animation tutorial. Motion or change in the corresponding non-animation graphics were depicted by arrows.

5 of 47

To distinguish between reading scores for the two groups, we will plot a separate frequency distribution for each group. Three things are noteworthy regarding the two distributions.

1. There is a difference between means on reading performance.

Figure 2. Non-animation page.

The animation tutorial included voice narration that accompanied the animations and guided the learner to important information by cueing important graphical events as they unfolded. The voice narration was the male voice of the experimenter carefully timed to coincide with the important animation events. Reproduction of the voice narrations in text form is shown in Appendix B. Voice narration was recorded using the Director software package. An animation screen consisted of a non-animation picture (e.g., a data graph) in the middle of the screen with text information on the upper portion of the screen. There was a small projector icon at the bottom of the screen that initiated the animation when clicked with the mouse. Subjects were unable to proceed until the animation was viewed at least once. Subjects viewed the animations as many times as they wanted.

Quiz. The quiz was developed to probe statistical information that the animations and corresponding non-animations provided. There was a total of 24 questions subdivided into spatial, textual, and animation-spatial questions.

Procedure

Subjects were randomly assigned to conditions (animation tutorial or non-animation tutorial), and type of computer (Powermac 6100 AV or Macintosh IIsx) as they arrived at the laboratory. They then filled out a consent form. The experimenter described what the subject would see on the screen, and do. Subjects in both conditions were directed to read all of the text information on all screens. Subjects were seated approximately 60 cm from the computer screen and initiated the computer tutorial by clicking the mouse. Subjects proceeded page-by-page through the tutorial by clicking with their mouse on a "next" button, shown in

Figure 1, on the bottom right hand corner of the computer screen. Following the tutorial the subjects moved to a separate room to answer the quiz in paper form.

Results

The dependent measures consisted of accuracy on the quiz, time to complete the tutorial, time to complete the quiz, time spent on each page type, and number of page accesses by subjects. Quiz accuracy was subdivided into four measures.

Quiz Question Types

Each of the questions shown in Appendix A, except question 1, addressed specific information provided by the animation and corresponding non-animation graphics. Question 1 determined whether subjects had a basic grasp of the terms dependent variable, independent variable, and covariate, that were used throughout the tutorials. All 24 questions were subdivided into three types a) text questions, which were straight-forward multiple choice questions presented in text format with no graphics, b) spatial questions and, c) animation-spatial questions which both used graphics as part of the question. The animation-spatial questions may facilitate mental manipulation of a spatial representation similar to an animation to arrive at the correct answers. The questions do not necessarily require mental spatial manipulation, but answers may be facilitated for the animation group. Spatial question answers could be determined by choosing among the provided choices without mental manipulation of the graphical elements necessarily as aid to arriving at the answer. For example, Figure 3 is considered to be a spatial question. It presumably can be answered quite easily by comparing the four graphs and noting that the first one has a steeper slope than the other choices. It is considered spatial because it is presented with graphics. The animation may have facilitated understanding of the issue, but a direct reconstruction and mental manipulation of visuo-spatial information may not facilitate learning further.

9) With ANCOVA, which of the following relations between the dependent variable and the covariate will yield the most reduction in variability of dependent variable scores?

Figure 3. Question 9: A spatial question.

The animation-spatial questions, by comparison, mapped the question and answers to a mental unfolding of the animation itself. The question was more congruent with the way that the animation material was presented. For example, the question graph in Figure 4 corresponds to the beginning of an animation. The answer corresponds to a final state of the animation for the regression lines presented. Researchers have demonstrated the potential importance of visualizing and mentally rotating objects in learning and memory performance (Cooper and Shepard, 1973; Kosslyn, 1976; Shepard and Metzler, 1971). Thus, answering the question using some kind of visual-spatial manipulation of the information such as spatial rotation may be facilitated for the animation group. It is assumed that recalling the animation would make it easier to answer the question although it is not required to answer the question.

20) The diagram below shows the mean differences for the dependent variable before the ANCOVA adjustment for the two groups. Circle the graph that best represents the mean differences on the dependent variable after adjustment.

Figure 4. Question 20: An animation-spatial question.

Analyses

Factorial between groups ANOVAS for the two conditions were performed on all of the quiz accuracy measures, time to complete the tutorial, and time to complete the quiz. Mixed factorials were used to analyze the page time and page access data for the animation and non-animation groups.

The tutorial crashed during the experiment for the first subject from the animation group. The tutorial was restarted and the subject started the session from the point where it crashed. During the data analysis it became apparent that there were artifacts in the data for this subject regarding times to complete the tutorial, time spent on each page, and the number of times each page was accessed. There was no reason to believe that this subject's learning performance scores on the quiz were affected by the computer crash. Therefore, the subject is included for the quiz accuracy analysis and excluded for the measures that were affected by the computer crash. For the quiz accuracy analysis, there was no difference in the outcome with the subject included or excluded even though this subject had a quiz score of 100 percent.

Preliminary factorial analyses for the dependent measures also included machine type (Macintosh Powermac AV6100 vs. Macintosh IIsi) and student type (graduate vs. undergraduate). For these analyses, all F 's showed p values greater than .05 except one. The analysis of number of page accesses with factors condition (animation, non-animation) by machine type showed a main effect for machine, $F(1,35) = 5.5$, $p < .05$ but no interaction with condition. There is no reason to assume this result, shown in Appendix C, is other than spurious. Thus,

it is assumed that the fewer mean page accesses for subjects on the IIsi compared with the Powermac were due to chance alone.

Quiz accuracy. Subjects answered the quiz and were measured on number of correct answers. The quiz which comprised 24 questions was divided into four dependent measures including total quiz accuracy and three subset measures. The first measure consisted of number of questions correct on all 24 questions. The second measure consisted of accuracy on 10 questions that formulated the questions in a spatial form. The third measure was accuracy on 5 questions expressed entirely in text form, and the fourth measure comprised accuracy on nine questions in animation-spatial form. The three subset measures were mutually exclusive and added up to the total number correct.

A between groups analysis of variance for the two groups (animation group vs. non-animation group) was performed for overall accuracy on the quiz. The three subset measures (text questions, spatial questions, and animation-spatial questions) for the quiz were analyzed in the same manner. There were no statistically significant differences for accuracy on the quiz for any of the dependent measures (see Table 1). The animation tutorial did not significantly contribute to quiz accuracy compared with the non-animation tutorial. All F 's are shown in Appendix D. Proportion of variance accounted for (r squared) by knowing which group subjects were in, was .04. Power, the probability to reject the null hypothesis when it is false, with the effect obtained was only .19 at the .05 alpha level. The number of subjects required to detect an effect of this size with power of approximately .8, would be 100 subjects per condition (Cohen, 1988).

As shown in Table 1, subjects achieved an average of 71 percent overall on the quiz and 53 percent on the animation-spatial questions where animation was expected to have greatest effect. Subjects in both groups answered the spatial questions well, with an average of 87 percent.

Table 1

Percent Correct on the Quiz with Standard Errors of the Mean

Condition	Question Type			
	Overall	Anim- spatial	Spatial	Text
	Quiz			
Animation	74	55	89	78
S.E.	.03	.06	.03	.04
Non-animation	69	50	85	70
S.E.	.02	.04	.03	.04
Both Conditions	71	53	87	74
S.E.	.02	.04	.02	.03

Figure 5 shows the percent correct for each question by conditions. The vertical axis shows the 24 questions, and the horizontal axis shows the percent of subjects who answered each question correctly. Accuracy for subjects in the animation group is shown in hatched bars and for the non-animation group in black bars. The questions are grouped into text, spatial, and animation-spatial categories. The graph shows that the questions did not discriminate well between the two groups. Subjects performed worse on the animation spatial questions relative to other question types. For example, only 25 percent of people for both groups

correctly answered Question 24, an animation-spatial question. With four alternative multiple choice questions, 25 percent represents chance level of responding if subjects were just guessing. In contrast to poor performance on the animation-spatial questions, with one exception, subjects in both conditions answered the textual questions well. Subjects from both groups answered text question 16 poorly. The graph shows that subjects did not perform well on the fundamental ANCOVA concepts questions (questions 16, and 18 through 24). For these questions subjects needed to know about the effects of the removal of the mean differences in the independent variable due to the covariate and the removal of error on the dependent variable attributable to the covariate. In addition, for question 16, they needed to know how these adjustments would affect the obtained F value.

Figure 5. Percentage of people by condition who answered the question correctly for each question type.

Tutorial time and quiz time. Time to complete the tutorial for each condition (animation vs. non-animation) was assessed by ANOVA which revealed that the non-animation tutorial group completed the tutorial significantly faster (21 min.) than the animation group (25 min.), $F(1, 38) = 4.75$, $p < .05$, $MSe = 33.67$. There was no difference in how long it took to complete the quiz. Mean times to complete the quiz are shown in Table 2 for the two groups.

Table 2

Mean Time in Minutes to Complete the Quiz with Standard Errors of the Mean

Condition	Quiz Time	
	Mean	S.E.
Animation	16	1.05
Non-animation	15	.84

Page time and number of page accesses. Analyses were performed on average time spent on pages and average number of accesses to pages. For these analyses “animation pages” are defined as pages with animations from the animation tutorial and the corresponding static graphics pages from the non-animation tutorial. Similarly, the static graphics pages from the animation tutorial and the same static graphics pages from the non-animation tutorial are defined as “non-animation pages.”

Wong (1994) found that subjects spent more time on animation pages relative to non-animation pages. To determine whether there were similar differences in the present study a two (animation page vs. non-animation page) x

two (animation tutorial vs. non-animation tutorial) mixed design analysis with repeated measures on the first factor was performed on time spent to complete the page. There was a significant main effect of page type, $F(1, 37) = 70.61$, $p < .001$, and a significant interaction between page type and condition, $F(1, 37) = 78.16$, $p < .001$. This interaction is shown in Figure 6.

Figure 6. Mean time spent on page type by condition

Follow up multiple comparisons using the Sheffé procedure showed that subjects in the animation group spent more time on average for animation pages (38 sec) than they did on non-animation pages (26 sec), $F(1,37) = 147.97$, $p < .01$. The animation tutorial group also spent more time on average per animation page than the non-animation tutorial group on corresponding pages (27 sec), $F(1,37) = 22.38$, $p < .01$. There was no statistically significant difference for time spent on the two page types for the non-animation group, and also no significant difference for time spent on non-animation pages between the two conditions.

A similar two (animation vs. non-animation) x two (animation tutorial vs. non-animation tutorial) mixed design analysis with repeated measures on the first factor was performed on number of times the pages were accessed by subjects. There was a significant main effect of page type, $F(1, 37) = 7.46$, $p = .01$, and a significant interaction between page type and condition, $F(1, 37) = 5.78$, $p < .05$. The interaction is shown in Figure 7.

Figure 7. Number of accesses by page type by condition.

Follow up multiple comparisons showed that subjects in the animation tutorial group accessed the animation pages significantly less often (1.03 mean number of page accesses) than they accessed the non-animation pages within their tutorial (1.22 mean number of page accesses), $F(1,37) = 11.65$, $p < .01$. The animation group also visited animation pages on average less often (1.03 mean number of page accesses) than did their non-animation group counterparts on corresponding pages (1.25 mean number of animation page accesses), $F(1,37) = 5.24$, $p < .05$. The non-animation group showed no difference in number of accesses to the two page types. The number of non-animation page accesses did not differ between the two groups. Both groups accessed the non-animation pages equally often.

In summary, analyses of accuracy on the quiz revealed no learning differences between the animation and non-animation groups overall and for the spatial, text, and animation-spatial components of the quiz. The time to complete the tutorial was greater for the animation tutorial group than the non-animation tutorial group. The animation group spent more time on animation pages than the non-animation group on corresponding pages but accessed those pages less often than the non-animation tutorial group. There was no difference between the two groups for mean time on number of accesses for non-animation pages.

Discussion

The primary goal of this study was to assess whether adult learning was better with animation compared with non-animation graphics. The study replicated Wong (1994) in that students learned how an abstract statistical procedure (ANCOVA) works using either an animation tutorial or a non-animation tutorial. To evaluate the effectiveness of animations as thoroughly as possible, the study made two substantial modifications to the materials used by Wong. As suggested by Baek and Layne (1988), emphasis was placed on key issues and important points in the animations. Voice narrations called subjects attention to important concepts and actions as they occurred. The quiz was also modified to explicitly probe concepts which the animations illustrated, and questions were presented in a spatial form that were congruent with the spatial information presented by the animations. With these modifications, it was felt that, if animations help learning, this study would demonstrate it.

Contrary to predictions, the study failed to find any facilitation to learning from the use of animations in computer-based instruction as compared to well designed non-animation graphics with adults. Retention was approximately 71 percent for both groups. Seven important questions of the 24 addressed the combined main processes of ANCOVA where differences in the covariate are removed from the independent variable by sliding group means along the regression lines, and error in the dependent variable attributable to the covariate is removed by flattening the regression lines. Questions 18 through 24, shown in Appendix A, addressed these conceptual processes directly. Performance on these and the two remaining questions of the spatial-animation questions was poor, relative to overall

retention, at 53 percent. It was not surprising that performance was poor relative to the overall retention on these very important, but difficult concepts, for subjects who received the non-animation tutorials. It is telling that subjects who received the animation tutorials, which were specifically designed to address these concepts, did no better.

What can we conclude about the effectiveness of animations for learning?

The current study resulted in another failure to find an advantage to using animations. There are two possibilities, 1) animations do not facilitate learning or, 2) all the studies that show no improvements when animations are used are deficient in some way.

The Possibility that Animations do not Help

Some people are reluctant to accept the possibility that animations do not facilitate learning (Wong, 1994). In spite of much experimental evidence, some cannot, or do not want to, accept that well-designed animations, which are interesting and impressive to watch, do not help. The only study that showed facilitated conceptual learning using animation with adults found a small effect (Baek and Layne, 1988). The current study incorporated Baek and Layne's suggestion of pointing out relevant animation material to learners by the use of voice narration. The narration did not help. As well, subjects in the Baek and Layne study achieved approximately 55 percent overall accuracy on a post-test, 5 percent higher than a control group with non-animation graphics. But even for this single positive result, the practical implications of animations with such a small effect are questionable considering the resources that must be spent on animations for such a small return. There is also the practical considerations of detecting a

significant effect with the current obtained effect size. Approximately 100 subjects per group would be required to detect an effect with power of .8. If it is this difficult to develop animations that help so little, why put resources into making them?

Even if animations do not improve learning, they might be of value if they provide equivalent learning in less time than non-animations. Unfortunately, subjects who took the animation tutorial spent more time on the tutorial than subjects who took the non-animation tutorial. In fact, the extra time it takes to complete the animation tutorial is entirely attributable to the additional time to play the animations, approximately 5 minutes.

Wong (1994) suggested that retention was not better with the animations because subjects used a poor learning strategy. The present study addressed and rejects this suggestion. In the Wong study, subjects in the animation tutorial group spent less time on non-animation pages than they spent on animation pages, and also less time than subjects in the non-animation tutorial group spent on non-animation pages. Wong speculated that animations might facilitate learning, but that essential non-animation material was ignored by subjects with the availability of compelling animations. Instructions in the present study, shown in Appendix E, emphasized the necessity to attend to all pages. Subjects in the animation tutorial group and in the non-animation tutorial group spent the same amount of time on the non-animation pages. Thus, Wong's (1994) failure to show an affect of the animations cannot be attributed to neglect of essential non-animation materials by subjects. Perhaps there is enough evidence to accept that well-designed static graphics are as effective in learning as animations.

The Possibility that Animations have not been Adequately Evaluated

It is always risky to accept the null hypothesis that something, in this case animation, does not have an effect. It is possible that the studies conducted so far had shortcomings that result in the failure to demonstrate real advantages of animations. A number of issues should be resolved before learning advantages of animation are dismissed.

First, in the present study, animations did not adequately convey the fundamental concepts. Mean retention on seven fundamental concept questions was poor as shown in Figure 5, with the poorest at 25 percent of subjects in both groups answering question 24 correctly. Subjects did the worst on questions 22 and 24, those questions that were more difficult because they applied the fundamental concepts of ANCOVA without being directly illustrated in the tutorial. Subjects did just as poorly on question 16, which is also a fundamental ANCOVA concept question, but in a text form.

Wong (1994) did not report retention scores separately for individual questions. Based on her analyses, she assumed that subjects did just as well with animations and equally well designed non-animation graphics. The picture that emerges from the current study is that subjects are doing equally bad on important fundamental concept questions. It is premature to conclude that animations are ineffective when particular animations on these essential concepts were ineffective. It is possible that improved animations might help.

The animations could possibly be improved in a number of ways. Wong (1994) designed the tutorials for an audience with a fairly low level of statistical sophistication. For example, the ANCOVA animations are not related to previously

learned statistical concepts and formulas that might help with understanding. Residualizing, residuals and partialing are not mentioned. People who understand partialing out the effect of nuisance variables probably will not integrate what is animated with previous knowledge about the nature of residuals unless the relation is made explicit.

Second, subjects viewing the animations used in the present study were passive, rather than active, learners. It is possible that animations which engage subjects in problem solving might be more effective than passive demonstrations (J. Gregory Trafton, personal communication, October, 1996). Subjects could be asked to predict what would happen before it happens in the animation. Immediately after an animation, subjects could be required to show that they understand what was taught, perhaps by manipulating regression lines. A greater variety of data sets could be presented to help subjects generalize the concept. Animations may not be as effective as free play or exploratory simulations.

The strongest recommendation to show successful facilitated learning with animations is to convey the fundamental concepts through clear and concise use of an appropriate number of animations and background information. Presumably, the best way to achieve this goal would be to add animations or elaborate on the existing animations to convey the fundamental concepts in tandem with some kind of problem solving technique. There may not be a single key aspect for animations to facilitate learning, but a combination of key elements such as problem solving, pointing out important unfolding events, and sufficient learning of the fundamental concepts. The quiz might be improved. Perhaps the questions on fundamental concepts are inadequate, and the problem is not that subjects in the animation

condition don't understand the concepts, but that the questions do not allow them to demonstrate this knowledge. Until these possibilities are ruled out, it is premature to conclude that animations provide no help in learning.

References

- Alesandrini, K. (1987). Computer graphics in learning and instruction. In H. Houghton & D. Willows (Eds.) *The Psychology of Illustration: Vol. 2. Instructional Issues*. New York: Springer-Verlag.
- Baddeley, A., D. (1990). The role of memory in cognition: Working memory. *Human Memory: Theory and Practice*. (pp. 67-95) London: Allyn and Bacon.
- Baddeley, A., D., & Hitch, G. (1974). Working memory. In G.A. Bower (Ed.), *Recent Advances in Learning and Motivation*, Vol. 8. New York: Academic Press.
- Baek, Y., & Layne, B. (1988). Colour, graphics, and animation in a computer-assisted learning tutorial lesson. *Journal of Computer-Based Instruction*, 15, (4), 131-135.
- Bork, A. (1981). *Learning with Computers*. Billerica, MA: Digital Press.
- Caldwell, R. (1980). Guidelines for developing basic skills instructional materials for use with microcomputer technology. *Educational Technology*, 20 ,(10), 7-12.
- Caraballo, A. (1985). An experimental study to investigate the effects of computer animation on the understanding and retention of selected levels of learning outcomes. Unpublished Doctoral Dissertation, Pennsylvania State University.

Caraballo, J. (1985). The effect of various visual display modes of selected educational objectives. Unpublished Doctoral Dissertation, Pennsylvania State University.

Cohen, J. (1988). *Statistical Power Analyses for the Behavioral Sciences: Second Edition*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Cooper, L. A., & Shepard, R. N., (1973). Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visual Information Processing*. New York: Academic Press.

Flexser, A. J., & Tulving, E. (1978). Retrieval independence in recognition and recall. *Psychological Review*, 85, 153-171.

King, W. A. (1975). A comparison of three combinations of text and graphics for concept learning. (Report No. NPRDC-TR-76-16). San Diego, CA: Navy Personnel Research and Development Center. (ERIC Document Reproduction Service No. ED 112 936).

Kosslyn, S. M. (1976). Can imagery be distinguished from other forms of internal representation? Evidence from studies of information retrieval time. *Memory and Cognition*, 4, 291-297.

Macromind (1992). Macromind Director. San Francisco, California.

Moore, M. V., Nawrocki, L., & Simutis, Z. (1979). The instructional effectiveness of three levels of graphics displays for computer-assisted instruction. (Report No. ARI-TP-359). Arlington, VA: Army Research Institute for the Behavioral and Social Sciences. (ERIC Document Service No. ED 178 057).

Paivio, A., & Csapo, K. (1973). Picture superiority in free-recall: Imagery or dual-coding? *Cognitive Psychology*, 5, 176-206.

Pressley, M. (1976). Mental imagery helps eight-year-olds remember what they read. *Journal of Educational Psychology*, 68, 355-359.

Rieber, L. (1989). A review of animation research in computer-based instruction. Proceedings of the 1989 Annual Convention of the Association for Educational Communications and Technology, 11, 369-390.

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

Rieber, Lloyd, P., Boyce, Mary, J., & Assad, Chahriar (1990). The effects of computer animation on adult learning and retrieval tasks. *Journal of Computer-Based Instruction*, 17, No. 2, 46-52.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, 701-103.

Tulving, E. (1978). Relation between encoding specificity and levels of processing. In L.S. Cermak and F. I. M. Craik (Eds.) *Levels of Processing and Human Memory*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Tulving, E., & Osler, S. (1968). Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology*, 77, 593-601.

Tulving, E., & Thompson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, 352-373.

Tulving, E., & Wiseman, S. (1975). Relation between recognition and recognition failure of recallable words. *Bulletin of the Psychonomic Society*, 6, 78-82.

Wong, A (1994). The use of animation in computer assisted instruction. Unpublished Doctoral Dissertation, Carleton University.

Appendix A

Performance Quiz

The text questions as presented to subjects in the quiz are: 1,7,15,16,17. The questions that were presented in a spatial form were: 2,3,4,6,9,10,11,12,13,14.

The questions that were presented in animation-spatial form were:
5,8,18,20,21,22,23,24.

1. Match the following items used in the tutorial.

c reading performance scores
a teaching method
b IQ

a. independent variable
b. covariate
c. dependent variable

2) Circle the graph from the 4 choices below that best represent the mean lines for group 1 and group 2.

Shading is the answer to question 3

3) In the frequency graph above shade the area of overlap between the two groups.

4) Assuming that there is a positive linear relation between reading performance and the covariate, place the points that are circled into the appropriate cell(s) in the grid below.

5) Choose the graph that best represents the regression line(s) that the ANCOVA process generates to perform the adjustment on the covariate for the data below (The two groups have equal means on the covariate).

- 6) The diagram below shows what a set of scores for one group looks like before the covariate adjustment is made. Select the diagram which shows what the scores will look like after the adjustment process.

7) Circle the correct answer.

After adjustment of the covariate:

- a) the covariate could still account for differences on the D.V.
- ☒ b) the regression line will have a slope equal to zero
- c) there will be more variability in the D.V. scores when compared to the original dependent variable
- d) both b and c.

8) The point circled is the data for Mary Smith before adjusting for the covariate. On the grid below, carefully draw the data point for Mary Smith after adjusting for I.Q..

9) With ANCOVA, which of the following relations between the dependent variable and the covariate will yield the most reduction in variability of dependent variable scores?

10) Which of the covariates listed in the question above will yield the least adjustment?

a. b. ☒ c. d.

For each of the following cases, indicate whether an adjustment of the means on the covariate is needed?

11)

- ☐ will require adjustment of covariate means
☒ will not require adjustment of covariate means

12)

- ☒ will require adjustment of covariate means
☐ will not require adjustment of covariate means

13)

- ☒ will require adjustment of covariate means
☐ will not require adjustment of covariate means

14)

- ☒ will require adjustment of covariate means
☐ will not require adjustment of covariate means

15) How does the result of an ANCOVA differ from the result of an ordinary ANOVA when the two groups are equal on the covariate mean?

- a. F-value increases, error decreases.
- b. F-value decreases, error decreases.
- c. F-value increases, error increases.
- d. F-value decreases, error increases.
- e. F-value increases or decreases, error decreases.
- f. F-value increases or decreases, error increases or decreases.
- g. F-value increases or decreases, error increases.

16) How does the result of an ANCOVA differ from the result of an ordinary ANOVA when the two groups are not equal on the covariate mean?

- a. F-value increases, error decreases.
- b. F-value decreases, error decreases.
- c. F-value increases, error increases.
- d. F-value decreases, error increases.
- e. F-value increases or decreases, error decreases.
- f. F-value increases or decreases, error increases or decreases.
- g. F-value increases or decreases, error increases.

- 17) A good covariate is one which
- a. has a mean equal 0.
 - b. has little variability.
 - c. has high correlation with the independent variable.
 - ☒ d. has high correlation with the dependent variable.

18) The diagram below shows the mean differences for the dependent variable before the ANCOVA adjustment for the two groups. Circle the graph that best represents the mean differences on the dependent variable after adjustment.

19)

20)

21)

22)

23)

24)

Appendix B

Animation Narration

narration	animation description	tutorial page
Group 1 mean, group two mean. The two groups have different means on the dependent variable, reading performance.	illustrates two groups with different means on the dependent variable	5
The two groups overlap on the dependent variable, reading performance	illustrates the overlap of the two sets of scores	6
Reading performance, only, is shown on the left, it is replotted as reading performance against IQ on the right.	a set of scores is replotted from a frequency distribution to a scatterplot with IQ. The scores float one by one from the first graph to the second.	12
Low score on the covariate, low score on the dependent variable. Hi score on the covariate, high score on the dependent variable	lines animated to show that subjects with lower IQ tend to have lower reading performance. Similarly with high IQ.	13
The separate regression lines generated for each group are parallel to the overall regression line.	group regression lines are generated out of an overall regression line.	18
We adjust the reading performance scores so the slope is zero. Now there is no relation between IQ and adjusted reading performance scores.	illustrate by rotating regression line so that it is flat.	19
Not adjusted, adjusted. IQ related to performance, IQ removed.	toggle regression line between adjusted and pre-adjustment positions.	20
Data for one subject, before adjustment, after adjustment, same distance from the regression line. Another subject, before adjustment, after adjustment, same distance from the regression line.	equality of the distance of a point from the regression line, before and after the adjustment, is illustrated.	21
Group 2, same adjustment.	same description as page 19	22

narration	animation description	tutorial page
After removing the effect of IQ for both groups separately, the two groups still have different means on the dependent variable	the two group regression lines are rotated so they are flat.	24
Look at the mean square error in ANOVA, it is smaller in ANCOVA. As a result, the F value in the ANOVA, is smaller than the observed F in ANCOVA.	ANOVA and ANCOVA summary tables are shown with animation illustrating the error and F values.	28
We slide the data sets along the regression lines creating a common mean IQ.	sliding regression lines showing the removal of the differences in mean IQ	32
Watch the mean reading performance differences decrease as we remove the mean differences in IQ.	sliding regression lines showing the changes in the dependent variable as the differences due to the mean IQ are removed.	33
We now reduce within group error by removing the effect of IQ.	flatten regression lines for both groups by rotating regression lines	34
Common mean on IQ, reduce error.	sliding regression lines followed by flattening of lines	35
In the previous example, a higher IQ score has a corresponding higher reading performance score. In the new example, a lower IQ score has a corresponding higher reading performance score.	shows corresponding positions on both axes for a data point. Shows this for previous and the new example.	36
As before we first find a common mean on the covariate, IQ.	See page 32 description	37
Watch reading performance means increase as we find a common mean IQ.	See page 33 description.	38
Remove within group error variance due to IQ.	See page 34 description	39
Common mean on IQ, reduce error	See page 35 description	40

narration	animation description	tutorial page
There is no relation between reading performance and height	plots scores from a frequency graph onto a scatterplot of reading performance and height demonstrating no relation.	42

Appendix C

Mean Number of Page Accesses by Machine Type for Animation Pages

Condition	machine	
	Macintosh IIsi	PowerMac
Animation	1.09	1.13
Non-animation	1.11	1.46
across conditions	1.1	1.3

Machine by condition ANOVA summary table

Source	SS	DF	MS	F	Sig of F
Within cells	4.37	35	.12		
Condition	.54	1	.54	4.34	.045
Machine	.69	1	.69	5.5	.035
Condition	.42	1	.42	3.35	.076
by machine					

Appendix D

ANOVA Summary Table for Percentage Correct on the Quiz for all Accuracy

Measures

All F tests performed with (1,38) df.

Dependent Variable	Source				F	Sig of F
	SS Between	SS Within	MS Between	MS Within		
Overall	.02	.63	.02	.02	1.5	.23
Animation	.03	2.02	.03	.05	.47	.5
Spatial	.01	.74	.01	.02	.63	.43
Text	.06	1.15	.06	.03	2.11	.15

Appendix E

Experimental Instructions

- 1) proceed through the tutorial a screen at a time until finished by clicking the mouse on the “next” button.
- 2) pay attention to all forms of information on each screen. Please make sure and read all of the text that is provided for every screen.
- 3) you must view each animation at least once before moving on to the next screen.