Effects of Animation and Narration on Recall and Problem-Solving Performance

By

Marcus D. Childress, Ph.D.

Abstract

This study examined the effects of three different formats of instructional multimedia presentations on the recall and problem-solving performance of novice-level adult learners. In this study, participants were randomly assigned to groups viewing computer-based presentations covering two topics (bicycle pumps & brass wind instruments). Participants receiving the words-before-pictures treatments viewed programs containing narration followed by animation; participants receiving the pictures-before-words treatments viewed programs containing animation followed by narration; participants receiving the words-with-pictures treatments viewed programs containing animation and narration simultaneously. Recall and problem-solving tests were administered immediately following the treatments to measure recall and problem-solving performance. There were no significant differences among the three treatment groups. Further investigation revealed that participants taking the recall test first scored higher in recall than participants taking the problem-solving test first. The findings suggest that educators should give consideration to not only presentation format, but also to testing order and procedures for testing.

Keywords
Education Technology, Computer Science, Animation, Narration, Dual-Coding

Introduction

Educators have often deliberated over ways to explain new material to their students in an understandable fashion. Most recently, educators have asked if multimedia instruction can promote understanding through the presentation of sound, graphics, animation, video, and text. Perhaps some of the most dynamic features of multimedia are digital animation and sound. Many software developers now incorporate these features into their instructional products such as multimedia encyclopedias and interactive books that not only allow the user to see information, but also to hear sounds and narration associated with topics of interest. Although educators have embraced multimedia with optimism, little research on how to design multimedia instruction exists. The use of features such as animation and narration has continued to increase due to the widespread availability of computers, authoring systems, and video/graphics tools. This availability has made it possible for educators to design and develop their own instructional presentations using visuals such as animation.

Understanding the relationships between animation and narration and the learning outcomes of recall and problem-solving is one of the keys to designing effective instructional multimedia materials. Theories such as Paivio's (1979) dual-coding theory and frameworks such as Ausubel's (1980) advance organizers and Mayer's (1989) conceptual models of systems are often cited as underlying principles for the use of animation and multichannel instruction. Instructional designers should investigate the attributes of narration and animation and their relationships with outcomes such as recall and problem-solving performance. This study examined the effects of three different formats of instructional multimedia presentations on the recall and problem-solving performance of novice-level learners.

Previous research findings

Previous research has examined the relationships between pictures and words in an instructional setting using films and static illustrations (Baggett, 1984; Baggett & Ehrenfeucht, 1983; Mayer, 1989; Mayer & Gallini, 1990; Mayer, 1999; Mayer & Chandler, 2001). In their study, Baggett and Ehrenfeucht (1983) had students watch an educational movie while being presented with visual and verbal auditory information. They concluded that there is no competition for resources. Encoding in one medium (verbal) does not hinder encoding in another (visual). Furthermore, it was shown that the simultaneous presentation of verbal/visual information promotes recall beyond that of sequential presentation of *verbal*-visual information. However, Baggett found that simultaneous verbal/visual presentation was only 8% better than a sequential *visual*-verbal presentation.

In 1984, Baggett extended the research by having participants watch a 30-minute film, before, after, or in synchrony with narration. Students in the synchronous visuals/narration group and the sequential visuals-narration group scored higher than the narration-visuals group on an immediate recall test. Thus, it was concluded that when recall is desired, information should be presented in a sequential visual-narration or synchronous visual/narration format.

Mayer and Anderson (1991) went one step further by systematically testing a single-code hypothesis, a

separate dual-code hypothesis, and Paivio's dual-code hypothesis. In the first experiment, Mayer and Anderson compared the effects of viewing an animation of a bicycle tire pump with narration (words-with-pictures) with the effects of narration followed by animation (words-before-pictures). A problem-solving test was used to determine performance of the participants. In the study, participants viewed three back-to-back presentations of the same material in a words-before-pictures or simultaneous words-and-pictures format. Results showed that participants in the words-and-pictures group scored significantly higher on the problem-solving test than participants in the words-before-pictures group. When tested on recall of verbal information, there was no significant difference between the words-and-pictures and words-before-pictures groups, which reflects the earlier findings of Baggett (1984) and Baggett & Ehrenfuech (1983).

In an attempt to extend their research, Mayer and Anderson (1992) changed their design while investigating what they call the contiguity principle. This instructional design principle implies that the effectiveness of multimedia instruction increases when words and pictures are presented contiguously. In their 1992 study, Mayer and Anderson included control groups and presented participants with concurrent (A+N), successive (AN or NA), animation-only (A), and narration-only (N) presentations. In the control group, participants received no instruction. The concurrent group was presented with three concurrent animation and narration presentations of the same instructional material (A+N, A+N, A+N). The three successive groups were presented with three types of presentations (ANANAN, NANANA, or AAANNN).

The animation-only group received presentations in the form of AAA, and the narration-only group received presentations in the form of NNN. In support of their contiguity principle, results showed that the control group performed worse in retention than all other groups and that there was no significant difference in the performance of those other groups. In measuring problem-solving, the concurrent groups performed significantly better than the successive, animation-only, narration-only, and the control groups, which did not differ significantly from one another.

Hypotheses. Novice learners who are presented with information in a simultaneous words-with-pictures format should perform better in *problem-solving* than novice learners who are presented information in a sequential words-before-pictures or pictures-before-words format (Mayer, 1989, 1993; Mayer & Anderson, 1991, 1992; Mayer & Gallini, 1990). If this is the case, the first research hypothesis should be true:

1. Novice-level participants in the words-with-pictures group will perform significantly better than novice-level participants in the words-before-pictures, and pictures-before-words groups in problem-solving performance.

Researchers such as Mayer and Anderson (1992, 1993) have also concluded that learners *recall* virtually the same amount of information regardless of whether the information is presented in a words-with-pictures, words-before-pictures, or pictures-before-words format. The research of Baggett and Ehrenfeucht (1983) and Baggett (1984), however, leads to the conclusion that learners viewing instruction in a words-with-pictures and pictures-

before-words format scored higher in *recall* than learners viewing instruction in a words-before-pictures format. Based upon the research findings of the Baggett and Ehrenfeucht, the second research hypothesis was generated:

2. Novice-level participants in the words-with-pictures and pictures-before-words groups will perform significantly better than novice-level participants in the words-before-pictures groups in recall performance.

Method

Participants and Design

Participants for this study were a group of 106 college students enrolled in two Virginia universities. The participants were chosen because of their inclusion in a wide range of academic backgrounds, ease of access, and size of classes. The total number of participants consisted of 86 females and 20 males. Participants ranged in age from 19 to 50. The mean age was 22 and the mode was 21. Participants were given a set of questions designed to determine the level of knowledge of the presentation topics (bicycle pumps or brass instruments). Only the participants shown to be "naive" in the subject matter were included in the primary and secondary analyses. Of the 106 participants in the study, 21 participants were excluded, leaving 85 novice-level learners for the analyses.

Materials and Apparatus

Computer-based materials consisted of six presentations showing a bicycle pump mechanism or the mechanics of producing sound with a brass musical instrument. Presentations were developed in

HyperCard 2.1. Animation and narration were presented in the form of *QuickTime* movies embedded into *HyperCard* stacks. The source for video and narration for the brass topic was an animation of the playing of a brass instrument from the Encyclopedia Britannica Education Corporation's videodisc *Instruments of the Symphony Orchestra*. Color video from the source videodisc was digitized and edited through *Premiere* and saved as *QuickTime* movies. Narration for both examples were recorded with *SoundEdit 16* . The source for video for the second topic was a color animation of a bicycle pump constructed in *Director*. Video and narration were then edited through *Premiere* into *QuickTime* movies. All presentations were made on *Macintosh* computers with color monitors. Stereo headphones were provided for participants, enabling them to hear the narration portion of the presentation without outside interference.

Selection Measures

Prior to seeing the bicycle pump or brass instrument presentations, each participant completed a pretest to determine his or her knowledge level (novice vs. non-novice) of bicycle pumps or brass instruments. Participants scoring above novice-level (21 participants) were excluded from the analyses. The criteria used in this study were the same used in Mayer's 1991 study on animation and narration. These novice-level participants were used in the primary analyses of the hypotheses and in the secondary analyses.

Questionnaires and Tests

Paper-and-pencil materials included subject questionnaires and test sheets. The subject

questionnaire for bicycle pump knowledge was a replica of the questionnaire used by Mayer in his 1991 study. The subject questionnaire for brass instrument knowledge was developed from the same framework as the bicycle pump questionnaire and was made to conform as closely as possible to the bicycle pump questionnaire (see Appendix A). The recall and problem-solving tests for the bicycle pump presentations were replicas of those used in Mayer's 1991 study. Brass instrument problem-solving and recall tests were developed with the same specifications to closely match the format and intent of Mayer's bicycle pump tests (see appendices B & C).

Procedure

Participants were randomly assigned to one of three treatment groups: words-and-pictures, words-before-pictures, and pictures-before-words. The words-with-pictures groups received a simultaneous words-with-pictures presentation, followed by either a words-before-pictures or pictures-before-words presentation covering a topic other than the first presentation; the words-before-pictures groups received sequential words-before-pictures presentation, followed by either a pictures-before-words or words-with-pictures presentation covering a topic other than the first topic; the pictures-before-words groups received a sequential pictures-before-words presentation, followed by either a words-before-pictures or words-with-pictures presentation covering a topic other than the first topic. Participants were randomly assigned to treatments and were tested individually. A list of random numbers was generated and each of the participants were randomly assigned to a presentation/testing group according to the generated

number. Participants were randomly assigned to one of the following six possible presentation combinations and test formats, as seen in Table 1.

Table 1

Presentation Combinations and Test Order

Topic	Format	Test Order	Topic	Format	Test Order
Brass	Words-with-Pictures	Problem-Solving/Recall or Recall/Problem-Solving	Pump	Pictures-before-Words	Recall/Problem-Solving or Problem-Solving/Recall
	Words-before-Pictures			Words-with-Pictures	
	Pictures-before-Words			Words-before-Pictures	
Pump	Words-with-Pictures		Brass	Pictures-before-Words	
	Words-before-Pictures			Words-with-Pictures	
	Pictures-before-Words			Words-before-Pictures	

Presentation

The participant was first prompted to "Click here to begin." Upon the clicking of the mouse, the participant viewed a color, digitized video of either 1) how a brass instrument produces a sound or 2) how a bicycle pump works (narration presented before video in the words-before-pictures groups, video presented before narration in the pictures-before-words groups, and narration presented simultaneously with video in the words-with-pictures group). The words-with-pictures group viewed a presentation lasting 40 seconds. Due to the sequential nature of the presentations, the pictures-before-words and words-before pictures presentations lasted 80 seconds.

The narration for the brass instrument presentation:

> When a player buzzes into the mouthpiece, each air burst from the vibrating lips causes the molecules in the tube to crowd together and push against each other. The pushing and shoving creates a wave-like ripple that moves rapidly out toward the bell, where it leaves the tube and continues through open air until your ear hears it as sound.

The narration for the bicycle pump presentation:

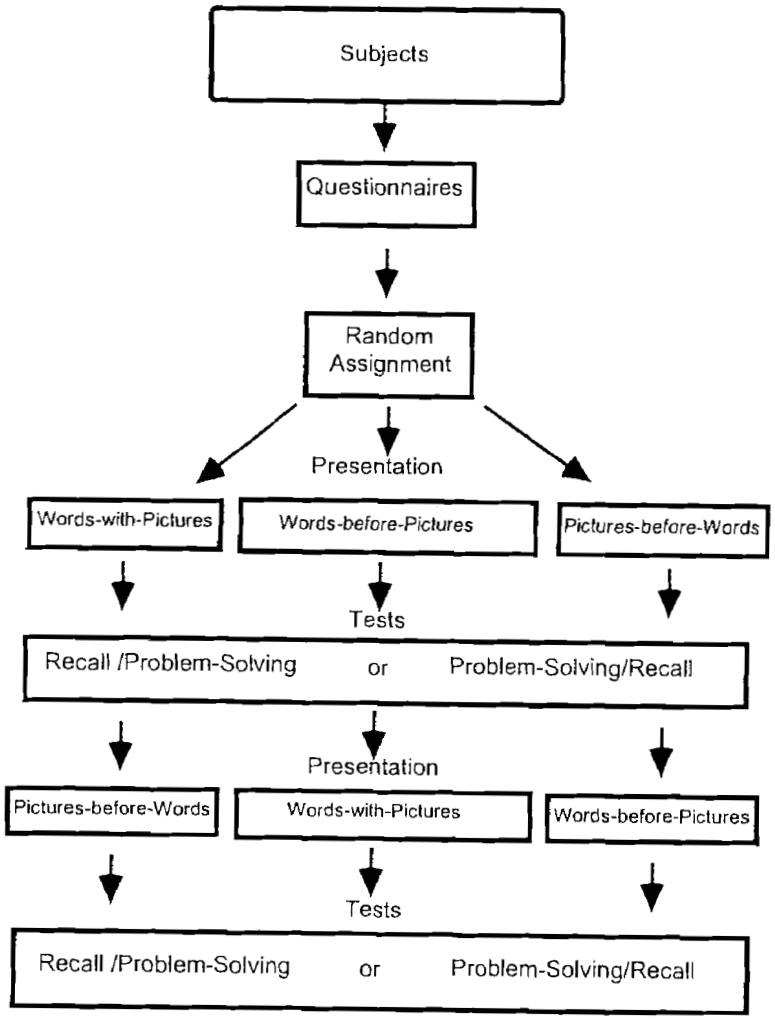
> When the handle is pulled up, the piston moves up, the inlet valve opens, the outlet valve closes, and air enters the lower part of the cylinder. When the handle is pushed down, the piston moves down, the inlet valve closes, the outlet valve opens, and air moves out through the hose. When the handle is pulled up, the piston moves up, the inlet valve opens, the outlet valve closes, and air enters the lower part of the cylinder. When the handle is pushed down, the piston moves down, the inlet valve closes, the outlet valve opens, and air moves out through the hose.

Testing

After finishing the first presentation, participants were tested on problem-solving and recall. Participants completed the tests in a problem-solving/recall format (four test questions for problem-solving, followed by ten test questions for recall) or in a recall/problem-solving format (ten test questions for recall, followed by the four test questions for problem-solving). Participants were allowed 2.5 minutes to answer each of the four problem-solving questions, for a total of ten minutes on the problem-solving test. Participants were allowed five minutes to answer the recall section. Upon completion of the first presentation and tests, participants moved on to the next section of the experiment and watched the second remaining presentation (either brass instrument content or bicycle pump content). Participants then completed the recall and problem-solving tests in the same manner as the first presentation. A schematic summary is shown in figure 1, illustrating the design of this investigation.

Figure 1

Schematic Design of Research Procedure



Scoring

Recall. For the recall portion of the test, participants were given one point for each correctly identified action recalled from the presentations, for a total of ten points. When scoring the recall responses, judges referred to the following checklists containing each of the ten actions involved in the brass instrument and bicycle pump presentations (see Table 2).

Table 2

Checklist of Actions

Brass Instrument	Bicycle Pump
1. player buzzes into mouthpiece	1. handle is pulled up
2. air burst from vibrating lips	2. piston moves up
3. causes molecules in tube	3. inlet valve opens
4. crowd together and push against each other	4. outlet valve closes
5. pushing and shoving	5. air enters cylinder
6. creates wave-like ripple	6. handle is pushed down
7. moves rapidly out toward bell	7. piston moves down
8. leaves tube	8. inlet valve closes
9. continues through open air	9. outlet valve opens
10. ear hears it as sound	10. air moves out through hose

Problem-Solving. When scoring the problem-solving responses, the judges referred to a list of possible appropriate responses. Participants were given one point for each correct answer on the problem-solving tests, a maximum of four points for questions one, two, and three; a maximum of two points for question four.

The identity of the treatment group to which the participants were assigned was unknown to the two judges. Inter-judge agreement measured for the problem-solving analyses was .95 (forced agreement

on 5%). Inter-judge agreement measured for the recall analyses was .96 (forced agreement on 4%).

Design

A 3 X 2 X 2 mixed model design was used to analyze the test data. Main effects and interactions were considered for presentation type (words-with-pictures, words-before-pictures, pictures-before-words), question type (problem-solving, recall), and topic (bicycle pump, brass instrument). The level of significance of all of the analyses was set at .05. In the secondary analysis, a one-way Analysis of Variance was used to test whether there was a significant interaction between presentation type and test order.

Results

Primary Analysis

Hypothesis one: Novice-level participants in the words-with-pictures group will perform significantly better than novice-level participants in the words-before-pictures and pictures-before-words groups in problem-solving performance.

Mean scores from the problem-solving tests for each group were subjected to an analysis of variance. The analysis was adjusted for unequal cell size. The resulting *F* statistic was not significant for the three experimental groups. None of the experimental groups scored significantly higher on problem-solving than any of the other groups.

The first hypothesis was rejected at the .05 level. Table 3 shows the mean and standard deviation of responses on the problem-solving test as a function of type of presentation which was used with each treatment group.

Table 3

Mean Problem-Solving Scores as a Function of Presentation Type

Presentation Type	*M*	*SD*	*n*
Words-with-Pictures	5.137	1.929	51
Words-before-Pictures	5.127	1.954	55
Pictures-before-Words	5.016	1.879	63

Hypothesis two: Novice-level participants in the words-with-pictures and pictures-before-words groups will perform significantly better than novice-level participants in the words-before-pictures groups in recall performance.

Means scores from the recall tests for each group were subjected to an analysis of variance. The analysis was adjusted for unequal cell size. The resulting *F* statistic was not significant for the three experimental groups. None of the experimental groups scored significantly higher on recall than any of the other groups. The second hypothesis was rejected at the .05 level. Table 4 shows the mean and standard deviation of responses on the recall test as a function of type of presentation which was used with each treatment group.

Table 4

Mean Recall Scores as a Function of Presentation Type

Presentation Type	*M*	*SD*	*n*
Words-with-Pictures	4.333	2.330	51
Words-before-Pictures	4.036	1.688	55
Pictures-before-Words	3.889	1.867	63

Secondary Analysis

With the use of two testing formats: 1) problem-solving followed by recall (PS/R) and 2) recall followed by problem-solving (R/PS), it seemed logical to investigate the relationship, if any, between test scores and test order. When analyzing differences in problem-solving scores of novices who had taken the recall tests first and the problem-solving tests second versus the problem-solving scores of novices who had taken the problem-solving tests first and the recall tests second, no significant differences were found. Novices who took the tests in a *problem-solving*/recall format did not score significantly higher than novices who took the tests in a *recall*/problem-solving format.

There was no significant difference in novice problem-solving/recall test takers and recall/problem-solving test takers in problem-solving performance. Table 5 shows the mean and standard deviation of responses on the problem-solving test as a function of test order.

Table 5

Mean Problem-Solving Scores as a Function of Test Order

Test Order	*M*	*SD*	*n*
Problem-Solving / Recall	5.023	2.133	88
Recall / Problem-Solving	5.160	1.639	81

Significant differences were found when analyzing differences between recall scores of novices

who had taken the recall tests first and the problem-solving tests second, versus the recall scores of novices who had taken the problem-solving tests first and the recall tests second. Novices who took the tests in a recall/problem-solving format scored significantly higher than novices who took the tests in a problem-solving/recall format. Table 6 shows the mean and standard deviation of responses on the recall test as a function of test order.

Table 6

Mean Recall Scores as a Function of Test Order

Test Order	*M*	*SD*	*n*
Problem-Solving / Recall	3.534	1.838	88
Recall / Problem-Solving	4.654	1.938	81

Discussion

In this study, presentations were developed in contiguous words-with-pictures, isolated words-before-pictures, and isolated pictures-before-words formats. There were two research hypotheses applied to this study. The first hypothesis was rejected. Participants in the words-with-pictures group did not perform significantly better than participants in the words-before-pictures, and pictures-before-words groups in problem-solving. Primary results from the study showed that there were no significant differences in problem-solving or recall performance among any of the treatment groups (words-with-pictures, pictures-before-words, words-before-pictures). These results disagreed with the findings of Mayer (1991) in which contiguous words-with-pictures presentations produced higher problem-solving performance. However, these results seem to confirm Rieber's (1990) claim that "The few serious attempts to study the instructional attributes of animation have reported inconsistent results" (p.84).

The second hypothesis was also rejected. Participants in the words-with-pictures and pictures-before-words groups did not perform significantly better than the participants in the words-before-pictures groups in recall. The results of this study are consistent with the findings of Mayer (1991). When recall is the desired outcome, the presentation order may not be as important as the content of the presentation in which verbal information is presented. The placement of verbal information with, before, or after animation does not seem to matter, as long as the verbal information is indeed presented.

When examining the effect of test order on recall and problem-solving, participants who took the recall test first and the problem-solving test second scored

significantly higher in recall than participants who took the problem-solving test first and the recall test second. These results agree with previous studies which show that recall may be hindered by the length of delay between presented material and recall of the material. With this in mind, participants who are asked to recall material first should and do score higher on recall than participants who are asked to complete a problem-solving test prior to a recall test. Differences of recall scores between problem-solving/recall and recall/problem-solving test takers may also be explained by interference. The problem-solving test may act as interference for those taking the recall test second, resulting in lower recall scores for the problem-solving/recall test takers.

Upon closer examination, it became apparent that alternative approaches to the design of this study could have been taken. Participants could have viewed the same presentation style for both topics, with the test order counterbalanced. Also, each presentation style for the first topic could have been matched to each other presentation style for the second topic. Therefore, order of presentation style per topic was not tested.

As with any experimental study, there are other factors which may have influenced the results of this study. Because of computer speed, at times the narration portion of the presentations skipped or "clipped-out" for a fraction of a second. Although this clipping of the narration seemed to not interfere with the understanding of the narration, it could be seen as a distraction to the participants encountering this. Also, this study used only written answers for recall and problem-solving questions. Other methods of data collection could have been used, such as drawings, sketches, and verbal interviews.

Although care was taken to pattern the brass questionnaires and tests as closely after Mayer's (1991) as possible, it appears that the questionnaires and tests may not have been equivalent to each other. For example, the bicycle pump questionnaire asks general household repair questions, while the brass instrument questionnaire asks questions specific only to brass instrument playing experience. Therefore, the novice-level bicycle pump participants may have been of a different level than the novice-level brass instrument participants.

Whether the recall tests actually measured recall must also be addressed. The recall tests asked the participant to "Please write down an explanation of how a bicycle tire pump (or brass instrument) works. Pretend that you are writing to someone who does not know much about pumps (or brass instrument)." While recall deals with remembering the verbal narration of the presentation, it appears that the recall test may encourage the participant to reconstruct the presentation for the purpose of explaining, rather than recall. Further research should investigate the types of tests used to evaluate recall and problem-solving performance. On several of the tests, participants sketched diagrams and illustrations of their answers. How can these drawings be included in the scoring? How do these drawings help the participants in recall or problem-solving?

Conclusion

The results of this study have implications for developers of instructional materials using animation and narration. Instructional designers need to be more aware of the different attributes used in the development of a multimedia instructional presentation. They should take into consideration not

only the presentation format (words-and-pictures, pictures-before-words, and words-before-pictures), but also the order of the testing, and the procedure used for testing. Designing instructional multimedia presentations involves much more than incorporating production features such as presentation format.

References

Ausubel, D. (1980). Schemata, cognitive structure, and advance organizers: A reply to Anderson, Shapiro, and Anderson. *American Educational Research Journal, 17* (3), 400-404.

Baggett, P. (1984). Role of temporal overlap of visual and auditory material in forming dual media associations. *Journal of Educational Psychology*, 76, 408-417.

Baggett, P., & Ehrenfeucht, A. (1983). Encoding and retaining information in the visuals and verbals of an educational movie. *Educational Communication and Technology Journal, 31*, 23-32.

Encyclopedia Britannica Educational Corporation. *Instruments of the symphony orchestra*. Chicago, Ill.

Mayer, R.E. (1989). Models for understanding. *Review of Educational Research, 59* (1), 43-64.

Mayer, R.E. (1993). Illustrations that instruct. In R. Glaser (Ed.), *Advances in instructional psychology,* 4, (pp. 253-284). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Mayer, R. E. (1999). Multimedia aids to problem-solving transfer. *International Journal of Educational Research*, 31(7), 611-623.

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology, 83*, 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*, 444-452.

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93*(2), 390–397.

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

Paivio, A. (1979). *Imagery and verbal processes*. Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Rieber, L.R. (1989). The effects of computer animation on intentional and incidental learning and motivation. Paper presented at the annual meeting of the Association for the Development of Computer-Based Instructional Systems, Washington, DC.

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

Marcus D. Childress is an associate professor of Instructional Design and Technology at Emporia State University.

Appendix A
Brass Questionnaire
BQ ID Number _____
Gender (M/F)_____Age_____

Put a check next to the things which apply to you:
_____I own a brass wind instrument (such as a Trumpet or Trombone)
_____I have played a Trumpet or French Horn.
_____I have played a Trombone or Tuba.
_____I have played another type of brass wind instrument.
_____I have blown into the mouthpiece of a brass wind instrument.
_____I have seen the inside of a brass wind instrument.

Put a check mark indicating your knowledge of how to play a brass wind instrument:

_____Very Much _____Much _____Average
_____Little _____Very Little

Appendix B

Brass Problem-Solving Test

BPS ID Number_____

1. What could be done to make the brass instrument more effective, that is, to produce more sound?

2. What could be done to make a brass instrument more reliable, that is, to make sure it would not fail?

3. Suppose you buzz into the mouthpiece of the instrument and no sound comes out. What could have gone wrong?

4. Why does sound enter a brass instrument?

   Why does sound exit a brass instrument?

Appendix C

Brass Recall Test

BR ID Number _____

Please write down an explanation of how a brass instrument makes a sound. Pretend that you are writing to someone who does not know much about musical instruments.

PLEASE KEEP WORKING UNTIL YOU ARE TOLD TO STOP