3-6-2004

An Experimental Investigation of Productivity Stratification and Knowledge Transfer in an Electronically Mediated Environment

Jachin Sakamoto

AN EXPERIMENTAL INVESTIGATION OF PRODUCTIVITY

STRATIFICATION & KNOWLEDGE TRANSFER IN

AN ELECTRONICALLY MEDIATED ENVIRONMENT

THESIS

Jachin Sakamoto, Captain, USAF

AFIT/GIR/ENV/04M-19

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GIR/ENV/04M-19

AN EXPERIMENTAL INVESTIGATION OF PRODUCTIVITY STRATIFICATION & KNOWLEDGE TRANSFER IN AN ELECTRONICALLY MEDIATED ENVIRONMENT

THESIS

Presented to the Faculty

Department of Systems and Engineering Management

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Information Systems Managment

Jachin Sakamoto, BS

Captain, USAF

March 2004

AFIT/GIR/ENV/04M-19

AN EXPERIMENTAL INVESTIGATION OF PRODUCTIVITY STRATIFICATION &
KNOWLEDGE TRANSFER IN
AN ELECTRONICALLY MEDIATED ENVIRONMENT

Jachin Sakamoto, BS
Captain, USAF

Approved:

//Signed//	06 March 2004
Lt Col David P. Biros, PhD (Chairman)	date
//Signed//	24 February 2004
Capt David D. Bouvin, PhD (Member)	date
//Signed//	01 March 2004
Dr. Joey F. George (Member) Florida State University	date

AFIT/GIR/ENV/04M-19

Abstract

The idea of anywhere and anytime learning is enticing from a military standpoint, given the high deployment rates in the current operational environment. Electronic-based learning is seen as an answer to this requirement. Currently there are many variations in electronic-based instructional media, and little has been done to determine which format or combination of formats is most conducive to facilitating knowledge transfer and learning. The research project explores, through the use of an experiment, three primary constructs of media richness, content flexibility, and forced engagement, in their relation to effectiveness or productivity in facilitating learning in the experimental participants. The instructional subject matter of choice in this experimental research was the art of detecting deceptive communication. Within the confines of this study, little empirical evidence was found to support the idea that any of the specific variations of electronic training medium outperformed the others in facilitating knowledge transfer.

Acknowledgements

Firstly, I would like to thank my wonderful wife for her support and interest during my tour here at the Air Force Institute of Technology. I would also like to extend my sincere gratitude to my thesis committee (Lieutenant Colonel David Biros, Dr. Joey George, and Captain David Bouvin), for their steadfast encouragement and expert guidance. This has truly been a challenging and eye opening experience that has broadened my understanding and given me newfound insight into the research process. I also owe great thanks to Dr. White for his expert guidance on the finer details of statistical analysis. In addition, I want to thank Dr. Mark Adkins, Captain Chris Ford, and Captain Michael Hass for their hard work and valuable inputs into the research process. Again, my sincere thanks to everyone.

Jachin Sakamoto

Table of Contents

Page

Abstract iv

Acknowledgements v

List of Figures ix

List of Tables xi

I. Introduction 1

Overview 1
Problem Statement 4
Narrowing the Study Scope 7
Research Question 7
Investigative Questions 7
Experimental Plan 8

II. Literature Review 9

Introduction 9
Models of Learning 9
Learning Theory of Behaviorism and Cognitivism as a Model 10
Learning Theory of Constructivism and Social Constructivism as a Model 11
Student Learning Styles: Multiple Intelligences 13
Learning Theories and Improving Instruction 14
Impact on Learning Productivity Assessments of Electronic Media 16
Media Richness 17
Content Flexibility 19
Forced Engagement 20
Overlapping Constructs 21

III. Methodology 24

Overview 24
Electronic Medium Selection 24
What the Instrumental Measurements were Designed to Capture 25
Characteristics of Population of Interest 26
Research Methodology 27
Experimental Treatment Group 1 30
Experimental Treatment Group 2 31
Experimental Treatment Group 3 31
Experimental Treatment Group 4 32
Experimental Treatment Group 5 32
The Instruments of Measurement 34

Page

Synopsis of Considerations that Lead to Experimental Design Selection 34
Permission to Conduct Research.. 35
Experiment 36
Planned Analysis Architecture (Operationalizing) 36
Planned Statistical Approach 41

IV. Data Analysis.. 44

Overview 44
Outline of Analysis 44
Introductory Knowledge Pretest Analysis 45
Introductory Knowledge Pretest/Posttest Differential Analysis 52
Analysis in View of Research Hypotheses 53
Analysis Outcomes in Terms of Experimental Hypotheses.. 61
Cues Knowledge Pretest Analysis 62
Cues Knowledge Pretest/Posttest Differential Analysis 69
Analysis Outcomes in Terms of Experimental Hypotheses.. 76
Summary of Statistical Analysis Results 77

V. Conclusion, Recommendations, and Lessons Learned... 78

Summary of Findings... 78
What the Results Mean 79
Limitations to Research Findings.. 80
Limitations to Implications: May be Topic Specific.. 82
Conclusion 84

Appendix A. Definition of Terms.. 85

Appendix B. AFRL Approval Letter 86

Appendix C. Introduction Pretest 87

Appendix D. Introduction Posttest... 89

Appendix D. Introduction Posttest... 89

Appendix E. Cues Lecture Pretest 91

Appendix E. Cues Lecture Pretest 91

Appendix F. Cues Lecture Posttest... 93

Appendix G. Pre- and Post-Judgment Tests 95

Appendix H. Introductory Session Translated Raw Data.. 97

Appendix I. Cues Session Translated Raw Data 102

Page

Bibliography 107

Vita.. 111

List of Figures

Figure Page

1. Three Media and One Topic 5
2. Daft and Lengel (1987:358) Media Richness Model 18
3. Proposed Model 23
4. Agent99 Screenshot 25
5. Histogram of Knowledge Pretest Data 46
6. Pretest ANOVA 47
7. Histogram of Studentized Residuals 48
8. Test for Normality 48
9. Plot of Residuals 49
10. Tests For Equal Variance 49
11. Pretest Cook's D 50
12. Pretest Tukey-Kramer HSD 51
13. Pretest Tukey-Kramer HSD Graphical Illustration 51
14. Histogram of Knowledge Pretest/Posttest Differentials 55
15. ANOVA of Pretest/Posttest Differentials 56
16. Histogram of Studentized Residuals 57
17. Test for Normality 57
18. Plot of Residuals 58
19. Tests For Equal Variance 58
20. Pretest/Posttest Differential Cook's D 59
21. Pretest/Posttest Differential Tukey-Kramer HSD 60
22. Pretest/Posttest Differential Tukey-Kramer HSD Graphical Illustration 60

Figure Page

23. Histogram of Cues Pretest Data 63

24. Pretest ANOVA 64

25. Histogram of Studentized Residuals 65

26. Test for Normality 65

27. Plot of Residuals 66

28. Tests For Equal Variance 66

29. Cues Pretest ANOVA Cook's D Influence Factors 67

30. Pretest Tukey-Kramer HSD 68

31. Pretest Tukey-Kramer HSD Graphical Illustration 68

32. Histogram of Knowledge Pretest/Posttest Differentials 70

33. ANOVA of Cues Pretest/Posttest Differentials 71

34. Histogram of Studentized Residuals 72

35. Test for Normality 72

36. Plot of Residuals 73

37. Tests For Equal Variance 73

38. Cues Pretest/Posttest Differential ANOVA Cook's D Influence Factors 74

39. Cues Pretest/Posttest Differential Tukey-Kramer HSD 75

40. Cues Pretest/Posttest Differential Tukey-Kramer HSD Graphical Illustration 75

41. Pretest/Postest Differential Mean by Treatment Group 81

42. Three Topics and One Medium 82

List of Tables

Table Page

1. Demographic Data 27
2. Experimental Design 28
3. Treatment Sessions 29
4. Instrumental Experiment Timeline 36
5. Linking Hypotheses with Experimental Treatments 41
6. Intro Pretest Means and Standard Deviations 45
7. Hypothesis Treatment Group Pairs 53
8. Intro Posttest Means and Standard Deviations 54
9. Intro Score Differential Means and Standard Deviations 54
10. Hypotheses Outcomes on Introductory Session Analysis 62
11. Cues Pretest Means and Standard Deviations 63
12. Cues Posttest Means and Standard Deviations 70
13. Cues Score Differential Means and Standard Deviations 70
14. Hypotheses Outcomes on Cues Session Analysis 77
15. Summary of Findings 79

AN EXPERIMENTAL INVESTIGATION OF PRODUCTIVITY STRATIFICATION & KNOWLEDGE TRANSFER IN AN ELECTRONICALLY MEDIATED ENVIRONMENT

I. Introduction

Overview

An increasingly fluid and deployed military environment taxed by personnel and fiscal shortfalls has become a modern reality for the United States Armed Forces and has elevated the underlying need for efficient and optimal allocation of time and resources. One segment of fiscal and real-time consideration is the training and education of personnel. The importance of learning and training of military personnel is well-documented in the annals of history, and precedes the rise of the standing army; this need has only continued to become more defined and essential in an increasingly information-centric environment. Hence, the concept of "anywhere" and "anytime" learning—training that adapts to the different time schedules and requirements of military personnel has become alluring.

It should be noted that the attractiveness of "anywhere" and "anytime" learning is not unique to the military or the Department of Defense at large. One particular mode of this training that has risen to prevalence over the course of the last decade is electronic learning. Within the confines of this research experiment, electronic learning will take on a broad definition, to be defined as instruction that takes place through some kind of computer interface. Some recent predictions have estimated that gross revenues being generated through this particular medium of learning will surpass $23 billion, on a global scale, by the year 2004 (Pastore, 2001). This trend is true for corporate businesses, public institutions, and the academic community. According to some estimates, the number of institutions of higher learning that will offer some form of electronic learning will exceed 3,300 by 2004 (Pastore, 2001).

It then becomes rather straightforward to conceptualize on the part of the organization that at least one of the factors that has driven the shift toward electronic learning, versus the more traditional form of classroom instruction, is a foundational consideration of an organization's financial bottom line. It has already been openly asserted by numerous academics and practitioners that electronic learning is a cost-effective and flexible way to train and educate today's workforce (Goodgride, 2002; Rosenberg, 2001). As such, it is reasonable to assume that corporate organizations have a desire to utilize the cost/savings benefits of providing training for their members in-house—without the financial burden of costly off-site workshops or highly paid outside experts. These same considerations are not foreign to military leaders and planners—particularly since there has been interest in this topic on the presidential level. Executive Order #13111, issued by the President in 1999, specifically mandates that electronic learning initiatives be used to further the training needs of government employees (13111, 1999).

Thus, electronic learning—which planners and visionaries hope will mitigate constraints on individual time and organizational resources—is rapidly becoming a focal point of migration within the Department of Defense. One of the most highlighted electronic learning projects within the Department of Defense is the Advanced Distributed Learning (ADL) initiative; an initiative which is designed with the goal of modernizing education and training through the smart use of information technology (*Department of Defense Strategic Plan for Advanced Distributed Learning*, 1999). However, while this initiative is one of the most highlighted, it does not and cannot begin to encompass the wide spectrum of electronic learning and training media with which different elements of the military organization deal. In fact, the methodology and the structure of electronic learning media as a whole and within the Department of Defense are strewn with variation. At the seam, where the technology interacts with a person, venues of electronic learning already range from electronic video clips (already employed frequently by commanders on topics such as safety and hot topics such as anthrax vaccine awareness), to voice-enhanced PowerPoint presentations (already employed for topics such as harassment and winter safety), to quasi-interactive environments that include quizzes and tests (such as security training). Diversified with even more variation are the potential topics that may be applied into the electronic learning arena. Topics range from concrete topics (e.g., projectile trajectory calculations) to abstract topics (e.g., leadership methodology).

As stated by Thurston and Earnhadt (2003), a number of recent studies, publications, and conference proceedings concerning the construct of electronic learning have focused on course completion rates; many of these studies have focused on the effects of distractions, feedback, and completion goals. Studies have shown that one form of electronic learning—web-based courses—has a tendency to have a lower course completion rate than the more traditional classroom-based courses; this gap has been shown to be as high as 40% (Carter, 1996; Phipps & Merisotis, 1999; Zielinski, 2000). Research findings have suggested that completion goals can reflect positively on completion—while off-task distractions detract from completion rates (Thurston & Earnhardt, 2003; Thurston & Reyolds, 2002).

Problem Statement

While the direction of these previous venues of study are constructive and relevant—especially within an academic environment of higher learning—the scope of these studies is not complete. To the reader, it should be relatively straightforward to conceive that within the construct of organizational training, especially within the Armed Forces, much of the training is required and not optional. External pressures all but ensure training completion; however, the military has a genuine interest in the idea that actual knowledge transfer and learning take place; training completion is only a means toward that goal. Time spent for training is valuable—so the question becomes, what methods of instruction and training within the electronic learning construct provide the highest rate of return on resources invested? It is rational to suggest that not all forms of electronic learning facilitate knowledge transfer and knowledge retention with the same degree of efficiency. Exploring this question of knowledge transfer, or learning, even in the absence of much internal motivation, is a holistic research construct—a specific segment of which was explored in this experimental study. The following figure illustrates the chain of rational consideration that led to the specific research questions and hypotheses presented later. Specifically, Figure 1 illustrates the conceptualization of the general construct that an individual who proceeds to receive training in an arbitrary topic—in this case Topic A—may find that not all the electronic learning media produce the same learning result, even though roughly identical information is presented. The educational trends, as they were outlined earlier, suggest that the military's demands for the flexible training made available through electronic learning tools will most likely increase in the future. An ever-heightened tempo in United States military operations means that military members will more than ever constantly be on the move, all over the world. It is conceivable that electronic learning and training will be looked at more and more to meet the demands of fluidity and flexibility made real by these ever-increasing operational demands. Given this expectation, it becomes rational to arrive at the conclusion that it is inherently important to achieve an underlying grasp of the comparative effectiveness of different variations in electronic learning media for timely and efficient knowledge transfer.

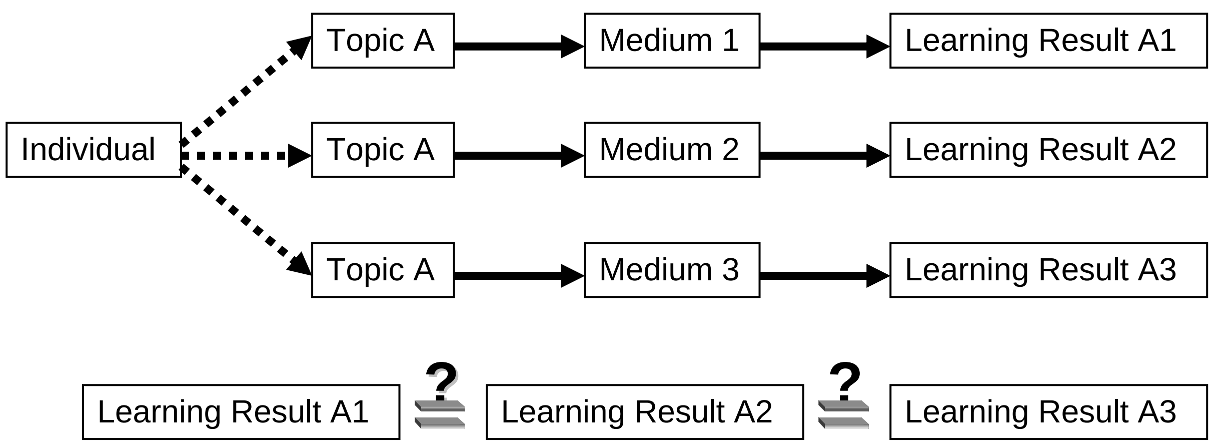

Figure 1. Three Media and One Topic

In light of the military focus in this experimental study, the topics of deception and the detection thereof were selected. Applications within these topics manifests importance and worth in scholarly pursuit. Deception has long been employed as an effective military strategy. Armies have used it for centuries to gain a competitive advantage over their enemies. Operation Desert Storm—otherwise known as the Gulf War—about a decade ago, is decked with examples of the military use of deception (Watson, 1991). One of the most prominent examples of deception within this war was the use of the “Hail Mary” maneuver, used to catch the Iraqi military by surprise, while an amphibious assault was feigned (Watson, 1991). Another ominous example is the case of former Federal Bureau of Investigations agent Robert Hanssen, who spied for the Russians for 20 years (*Ex-FBI Spy Hanssen Sentenced to Life, Apologizes*, 2002).

However, even despite the prevalent use of deception throughout military history, it has also been shown that the untrained human will only detect, on average, 50 percent of the deception that he or she encounters (Miller & Stiff, 1993; Vrij, Kneller, & Mann, 2000; Zuckerman, DePaulo, & Rosenthal, 1981). Educational resources and electronic teaching forms need to be allocated in a way that will maximize their utility in both cost and effectiveness. The electronic forums need to be compatible with topics of military importance, including the topic of deception. As discussed, deception is a valid military topic because members of the armed services need to be able to detect when someone is attempting to deceive them; this applies to everyone from the front-line soldier to the national level decision maker. Endeavors to use varying knowledge transfer techniques in the arena of deception are particularly challenging, given the difficult and abstract nature of recognizing and understanding deceptive cues. Understanding what teaching forum best lends itself to tacit learning and knowledge retention, given this abstract topic, will help prepare servicemen and servicewomen for combat in an information-centric environment. By learning deception cues, we will be able to detect deceitful actions and be prepared for the intent of the attempted deception.

Narrowing the Study Scope

Many different variations in the format of an electronic teaching medium can be conjectured to have a plausible impact on student learning. However, within the confines of this research project, given limitations on time an resources, three primary variations within the format of an electronic medium were addressed. The constructs chosen were media richness, content flexibility, and forced engagement. The theoretical background to these selections is presented in detail in Chapter 3.

Research Question

Given the constructs addressed in this study, several broad research questions come to mind. Does the level of interactivity in an electronic learning medium significantly influence the deception detection scores of student participants? Does the level of media richness influence the deception detection scores of student participants? Does the level of forced interaction influence the deception detection scores of student participants? These questions, will be considered and then refined into hypotheses through the literature review process, as detailed in Chapter 2.

Investigative Questions

- Does electronic training that uses rich and varied media exceed video lectures have an impact on the learning process for deception detection?

- Does electronic training that uses a non-linear electronic training environment improve upon a linear electronic training environment when teaching the topic of deception detection?
- Does the use of pop-up quizzes in electronic training have an impact on the learning process for deception detection?

Experimental Plan

In order to explore these investigative questions, a solid foundation in relevant theory will discussed in the following chapter. Given the nature of the investigative questions, applicable topics of theory will largely be based on learning theory. Following and in-depth discussion of the theoretical data, coherent hypotheses will be formulized and presented. Additionally, experimental methods/tools for testing the hypotheses will be discussed.

II. Literature Review

Introduction

This section is a substantial review of the relevant published works concerning the applicable construct spheres of electronic learning, teaching theory, deceptive communication, and an electronic learning tool created for a series of related studies on deception detection. Other subtopics such as communications, including computer-mediated communications—and learning motivation are discussed. This section will end with a theoretical model and construct. Within this section, the research problems and questions that were rationally established and comprehensively outlined in the previous chapter are built upon and reviewed using already-established research as a guide.

Models of Learning

Within the academic community there are four major model constructs of learning that are widely recognized. These models are behaviorism, cognitivism, constructivism, and social constructivism (Hung, 2001). Based on these models, a conceptual framework categorizing computer-mediated tools to learning theory construct has been presented. The question raised in this preceding scholarly work is: "Which instructional method would 'most efficiently' enable knowledge to be learned and understood?" (Hung, 2001:282).

Learning Theory of Behaviorism and Cognitivism as a Model

The theory of Behaviorism and Cognitivism has grown in large part out of the stimulus response theory proposed by Burrhus Skinner in 1974 (Hung, 2001; Skinner, 1974). Learning, as understood through this construct, happens when a subject or a "learner" is conditioned to respond to a stimulus (Hung, 2001:281; Skinner, 1974). The actual "inner processes" of the learner were not considered to be key to the learning process: individuals were treated as "black boxes" (Hung, 2001:281; Skinner, 1974). Burrhus Skinner rationalized that since it was not feasible to measure and study a learner's internal workings with any of the available scientific methods, it would be better to focus on the "cause-and-effect relationships" of learning that could be scientifically observed (Hung, 2001:282; Skinner, 1974). "Knowledge is a storehouse of representations, which can be called upon for use in reasoning and which can be translated into language." (Hung, 2001:282) This theory was built upon by several researchers, including Winograd and Flores, in 1986 who asserted that thinking was simply a process that involves the manipulation of representations, and this developed into the idea of cognitivism (Hung, 2001; Winograd & Flores, 1986).

The theory of cognitivism has varied and evolved, giving rise to the premise of Situated Cognition, which asserts that learning or meaning is derived from a learner's interaction with the environment—"meanings are to be taken as relations among situations and verbal or gestural actions" (Hung, 2001:282). This premise refutes the assertion of cognitivism; advocates of Situated Cognition argue that knowledge is not distinct and abstract (Hung, 2001). Instead, under the theory construct, there is an "inextricable link" between contextual environmental constraints such as society, history, and culture, and the acquisition of knowledge (Bredo, 1994; Brown, Collins, & Duguid, 1989; Greeno, 1991; Hung, 2001; Prewat, 1996; Rowe, 1991). Simplified, this means that learning, as understood within the confines of this construct, is "an aspect of person-environment interaction, where activity involves a transaction between person and environment that changes both" (Hung, 2001:282).

Learning Theory of Constructivism and Social Constructivism as a Model

The learning model which has become dominant in many academic circles is that of Constructivism and Social Constructivism (Hung, 2001). Under this paradigm, the assumption is made that learners "interactively work and rework their understanding of an area to construct their own knowledge representations" (Robson, 2000:153). The theory of Constructivism approaches learning from the standpoint of the learner who must engage in "knowledge discovery" which in turn is "the product of an indissociable collaboration between experience and deduction" (Bruner, 1990; Hung, 2001:282; Piaget, 1960/1981:13). The theory builds upon the Neo-Piagetian theory, whose foundation is a model of the cognitive structure of intelligence, to explain how a learner uses his or her intelligence to learn (Robson, 2000). The concept of active learning is emphasized with a focus on the process of assimilation and the accommodation of knowledge (Bruner, 1990; Hung, 2001; Piaget, 1960/1981). Learning, within the confines of this theoretical construct, cannot be separated from the learner (Hung, 2001). The view is that "learning is an active process of constructing rather than acquiring knowledge" (Hung, 2001).

Others researchers such as Vygotsky (1981) and Lave and Wenger (1991) have emphasized and incorporated social and neo-Marxist theories of practice into this learning theory (Hung, 2001). This social emphasis on Constructivism has become known as Social Constructivism. Individual learning, as understood within the confines of this theoretical construct, is conceived to be highly influenced by interactions with people—such as children, parents, and teachers (Hung, 2001). As in Brunner's (1990) own words, children in school, while they are in the process of learning, are actually "participating in a kind of cultural geography that sustains and shapes" what they are doing, "without which there would [] be no learning." This revision has been applied to language learning, as a "means for social co-ordination and adaptation" (Hung, 2001:282; Maturana & Varela, 1987). Language itself can be rationalized to be the tool or mode of knowledge conveyance; learning takes place in the context of language. An individual's language is shaped by social upbringing and culture. As such, Maturana and Varela (1987) argued that in order to understand human learning, it is useful to view it as a process of human communication (i.e. "languageing") (Hung, 2001:282). The intent is to help learners "socially construct knowledge in collaboration within a group using common language, tools, values and beliefs to enable practices, and seek meaning appropriate to the culture of the area of study" (Robson, 2000:153). Hung (2001:282) points out that the process of human communication not only involves the exchange of thoughts and ideas, but also the exchange of "coordinating action and socializing actors as well."

Student Learning Styles: Multiple Intelligences

The theories abstracted above provide a macroscopic construct through which to examine the learning process. However, on a more refined, microscopic level, it may be useful to review theory that addresses the specific learning style variations among individual learners. Gardner, in his theory of multiple intelligences, addressed the learning construct at this level of granularity (Nolen, 2003).

In this theory, learners have eight variations of learning intelligences: verbal, linguistic, musical, mathematical/logical, spatial, bodily-kinesthetic, intrapersonal, and naturalistic. Within the verbal learning variation in intelligence, individuals are thought to learn through their mastery of language. Additionally these learners are thought to have strong auditory skills. Learners with linguistic intelligence are able to “pay special attention to grammar and vocabulary” and tend to be able to memorize best using words. When trying to instruct learners whose strengths lie in these categories, it is important to use language that students can relate to and fully understand. (Nolen, 2003:115)

Those with music intelligence have a strong understanding of rhythm, pitch, and timbre. Music intelligence is thought to be related to mathematical-logical intelligence, because of the comprehension of ratio and regularity, among other music related patterns. Individuals with mathematical-logical intelligence are thought to have an astute “ability to detect patterns, reason deductively, and think logically.” These individuals are often the “model students” because they are able to deftly follow the logical sequencing of instructional material, as it is presented within the traditional classroom. (Nolen, 2003:116)

Individuals with spatial intelligence are able to manipulate and create mental images of problems. When instructing individuals whose strength lies within this Gardner learning variation, it is best to use visual stimuli such as pictures, photographs, films, and overheads. In contrast, individuals with bodily-kinesthetic intelligence are best able to understand the world through the body. The best way to teach individuals with this learning strength is through hands-on activities. (Nolen, 2003)

Within the interpersonal Gardner intelligence learning construct, individuals are thought to have a keen ability "to understand, perceive and discriminate between people's moods, feelings, motives, and intelligences." In order to best instruct individuals with this type of intelligence, it helps to promote team and group work and interaction. Under Gardner's final learning construct of naturalist intelligence, individuals are thought to learn best outdoors, through activities such as field trips and nature hikes. (Nolen, 2003:118)

Learning Theories and Improving Instruction

The pragmatic complement to learning theory, at both the macroscopic and microscopic level, is the study of how to improve instruction. Learning theory is based on the knowledge gain or retention from the perspective of the learner. Instructional approaches balance this, focusing on the vehicle of knowledge conveyance. The learning models of Behaviorism, Cognitivism, Constructivism, and Social Constructivism that were delineated above, when viewed in the context of instruction, can be seen as not only different but virtually at odds with one another. How then can instruction be improved?

One approach to addressing this is through the Carrol (1963) model of school learning (Rathis, 2002). Within this model it is proposed that student learning is dependent on two variables: the quantity of time spent by a student on a task and the quantity of time needed for a student to master that task (Rathis, 2002). Reworded, the quantity of student learning varies directly with time on task and inversely with time needed to learn (Rathis, 2002). In algebraic notation the formula would look as follows:

$$\textit{Learning} = \textit{Time on Task}/\textit{Time needed to learn}$$

Equation 1. Model of School Learning

Since Carroll first proposed this model of school learning, much follow-up research has been conducted and published. As published by Rathis, the following (quoted from his publication) are inferences that instruction has improved (2002:233):

1. If the amount of learning that takes place in a class increases, all things being equal, then one might reasonably infer that instruction has improved.
2. If students increase their time on task within a lesson or a unit of study, all things being equal, then one might reasonably infer that instruction has improved.
3. If the time students need to learn the objectives of the lesson or unit is reduced because of teaching interventions (e.g. scaffolding), all things

being equal, then one might reasonably infer that instruction has improved.

4. If the complexity of the objectives addressed increases across lessons or units, all things being equal, then one might infer that instruction has improved.
5. If the activities assigned to students and the assessments given to students are more closely aligned with a lesson's or unit's objectives, all things being equal, then one might reasonably infer that instruction has improved.

Carroll's theory can then be used as a basis for inferring that instruction has improved, given that one of these five factors holds true. However, in order to have a rational basis to formulate a sound argument for asserting which influence or what differentiation in instruction caused the improvement in instruction, additional theory, especially that concerning the use of electronic media, will be discussed.

Impact on Learning Productivity Assessments of Electronic Media

The relationship between the electronic medium used to present a body of knowledge and the underlying learning theories provides a foundational basis through which to theoretically assess the effectiveness of a particular learning tool. As Robson (2000:154) points out in reference to course design, "decisions concerning context, functionality, display and control of software, there must be a trade-off among various design principles." For example, under the Behaviorist and Cognitive approach, a well-organized sequential layout of information would be paramount, whereas the

Constructivism approach would steer toward a non-linear presentation where students are active and self-regulated. The same principle holds true when considering the impact of Gardner's modes of learning. While a methodical presentation of information in a written format may provide the ideal learning environment for a particular group of individuals predisposed to verbal or linguistic learning, it would be arguably less effective for those whose individual strengths lie in areas defined by the spatial or visual Gardner learning constructs. However, by following the same chain of reasoning, it could be rationalized that an instructional medium tailored to address the underlying ideology of multiple learning theories and individual learning modes would more effectively reach a wider audience. As presented, the constructs addressed by the different learning theories are many and multifaceted and not easily addressed all at once. Hence, the actual learning constructs addressed in this research study have been narrowed to three primary areas of interest: media richness, content flexibility, and forced engagement.

Media Richness

"Media differ markedly in their capacity to convey information" (Lengel & Daft, 2001:226). Hence, it makes rational sense to hypothesize that one key characteristic of a learning medium is the variation of media richness that is used. The richer a presentation medium, the wider the variety of symbolic languages used, such as "graphics, voice inflections, and body gestures" (Lim & Benbasat, 2000:451). All these convey information to the viewer. Each of the symbolic language systems has its own strengths and weaknesses. While a picture may be said to be worth a thousand words, it cannot easily “convey the meaning of conditional events or causes, such as *if, nevertheless, because, or no,* because pictures do not entail any logical connections" (Lim & Benbasat, 2000:451). The media richness hierarchy proposed by Lengel and Daft, as presented in Figure 2 below, illustrates the variation of media richness by media type on a continuum from low to high (Daft, Lengel, & Trevino, 1987). “The reason for richness differences includes the medium’s capacity for immediate feedback, the number or cues and channels utilized, personalization, and language variety” (Daft & Lengel, 1986:560).

.

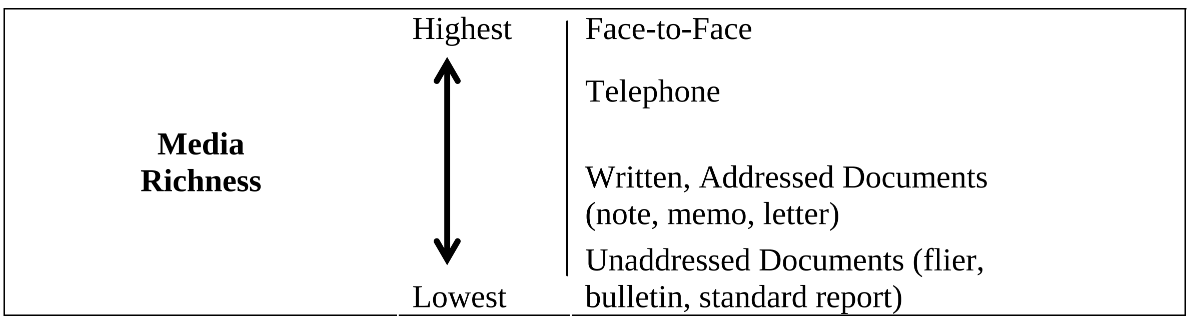

Figure 2. Daft and Lengel (1987:358) Media Richness Model

Furthermore, many human experiences and emotions are not easily encoded; they “resist logic descriptions” and are not easy to “communicate when translated into verbal (or mathematical) descriptions” (Hansen, 1989; Lim & Benbasat, 2000). According to Lim and Benbasat, a multimedia presentation is considered to be media-rich, while pure text is considered to be media-lean. Fundamentally, this suggests that the richer the media of presentation, the better that media is at conveying information. A “rich representation can better convey information, especially nonverbal messages, which may facilitate the understanding of the information” (Lim & Benbasat, 2000:451). This construct, seen in light of the theory of behaviorism—which presents learning in terms of a student being conditioned to respond to stimulus, as already presented at length earlier—can be further rationalized that not only an increase in the richness of the media (which would increase the quality of stimulus), but also an increase in the variation of the media (which would increase the sum total of stimulus), will increase its effectiveness at conveying information to a wider audience. The following hypothesis is proposed, suggesting that media richness will have a significant positive effect on student learning:

H1: Having more media richness within the electronic learning medium will result in significantly higher student learning differentials than media with less media richness.

Content Flexibility

Over the last few decades, educators have shifted "away from teaching structured sets of facts in predetermined order" toward "participatory learning environments" that promote "students' individual development" (Kraidy, 2002:95). This same philosophy has been applied to computers with the expectation of creating a customizable learning process. According to Kraidy (2002:96), many studies have been done that classify the information delivery system of these technologies into "linear and non-linear media." Examples of linear instructional methods include books and lectures, whose order are predetermined by the author or instructor; these are considered part of the traditional teaching methodology. The use of computers has brought about non-linear teaching technologies such as the use of hypertext, which can provide "multidimensional access to information," which can then be focused on giving the individual a personalized and customized learning experience (Kraidy, 2002:96).

One of the basic rationales behind a non-linear approach, versus a linear approach, is to allow students to develop their own learning paths by having the freedom of navigation through the material of interest (Chen, 2002). This construct seen in light of the theory of constructivism, (discussed earlier with its implications on student interaction with the environment, along with the distinct variations learning styles), makes it feasible to rationalize that if a non-linear learning capability of the instructional software allows participants to tailor the material presented in a way that meets their own cognitive needs, then their own particular learning styles are more likely to be addressed, and their learning should increase. This forms the basis of the following hypothesis that an increase in the content flexibility within the teaching medium will lend itself to higher student learning differentials:

H2: An electronic medium with more content flexibility will result in significantly higher student learning differentials than media with less content flexibility.

Forced Engagement

The well-known Socratic teaching method is based on a teacher asking his or her pupils questions, so that they can learn how to think clearly. Asking students questions forces them to engage in the material at hand: it forces student engagement. As extracted from the works of Bloom (1976) and Hecht (1978) by Kumar (1991:50) in his meta-analysis of instruction-engagement, "if an instructional method [has] an influence on student achievement," then "student engagement in a learning task must mediate the relationship." An appropriate way to define student engagement is the time a student "actively participates in learning" (Kumar, 1991:50). Participation includes "experimenting, attending, participating in discussion, questioning, answering, taking notes, etc." (Kumar, 1991:50). Thus, one way in which to actively engage a student with the material is through the active use of questions—which are implemented readily enough into an electronic teaching medium. The expectation, then, is that the increased interactivity of the participant with the medium will lead to an enhanced learning experience. This would fall in line with an extension of the learning theory of social constructivism, which, as presented earlier, defines learning as an interaction between the individual and other people: in this case, the creator of the electronic medium interacts with the student through the medium by a set of predetermined questions. This leads to the third construct hypothesis formulated for this study, that the presence of forced engagement will cause a participant to interact or engage the material being presented will lead to higher learning:

H3: The presence of forced engagement of the student with the material will result in significantly higher student learning differentials than media with no forced engagement.

Overlapping Constructs

Furthermore, given the reasoning presented above, it becomes feasible to conjecture that when the influences addressed by the forgoing three hypotheses are simultaneously introduced into a learning medium, the effect might be additive or synergistic. In other words, increasing media richness and introducing content flexibility should also lead to an increase in student learning. For example, one might deduce that the introduction of both media richness and content flexibility will also have a positive impact on improving instruction, and ultimately student learning. The same principle would he expected to hold true for media richness and forced engagement, and between forced engagement and content flexibility. The overlaps between all three influences will be explored, and thus yield the following four additional hypotheses:

H4: The sum total of media richness and content flexibility will cumulatively result in significantly higher student learning differentials than that in media with lower levels of media richness and less content flexibility.

H5: The sum total of media richness and forced engagement will cumulatively result in significantly higher student learning differentials than that in media with lower levels of media richness and no forced engagement.

H6: The sum total of content flexibility and forced engagement will cumulatively result in significantly higher student learning differentials than that in media with less content flexibility and no forced engagement.

H7: The sum total of media richness, content flexibility, and forced engagement will cumulatively result in significantly higher student learning differentials than that in media with lower levels of media richness, less content flexibility, and no forced engagement.

The sum total of all seven hypotheses is visually illustrated in a construct model, as presented in Figure 3. As shown, each of the three constructs: media richness, content flexibility, and forced engagement, and their overlapping effects, are expected to have a positive impact on student learning.

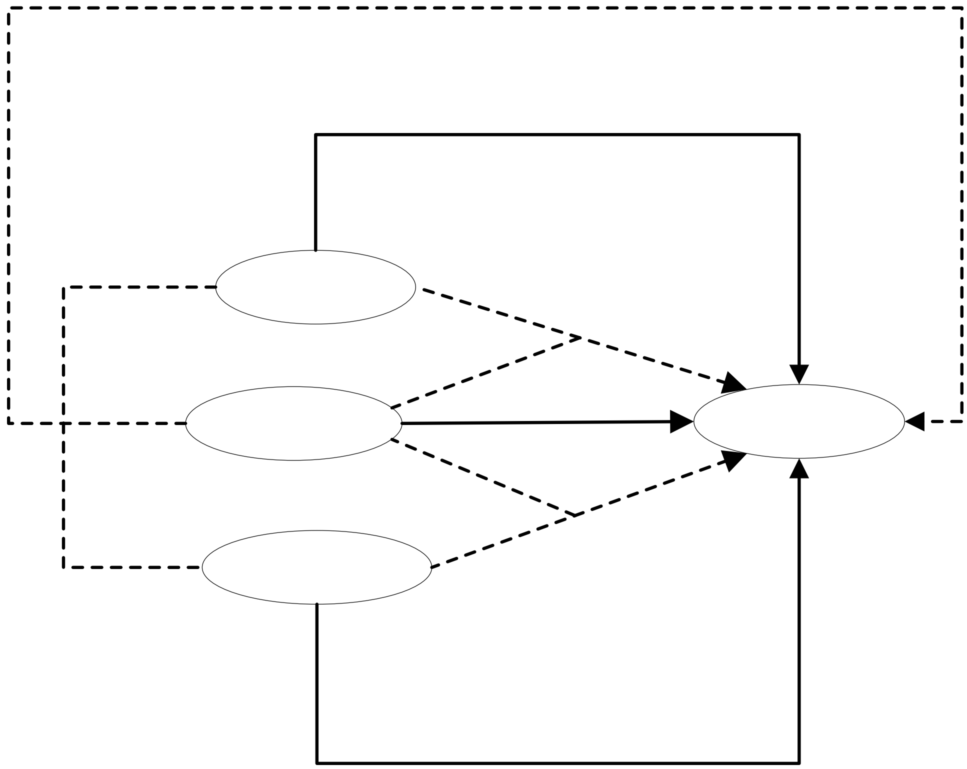

Figure 3: Proposed Model

III. Methodology

Overview

Thus far, a comprehensive summary of pertinent scholarly research regarding the concepts of learning, electronic forums, and deception has been presented; in-depth review of this available data, mixed with a rational conceptualization of present-day trends, laid the foundation which provided the basis for the research problem which was raised and the hypotheses that have been presented. This chapter describes in a methodical and purposeful manner the modes and mechanisms through which the research question was explored, and the controls that were put in place to elevate accuracy and care for the underlying constructs of experimental validity.

Electronic Medium Selection

It is in the military's best interest to understand which training and education medium formats are the most effective. The tool that will be used in this research project to study variations in electronic media is the Agent99 trainer. A visual snapshot of what the Agent 99 medium looks like is presented below in Figure 4 below. The top left quadrant is where the video lecture is presented. The top right quadrant is an outline of the lecture. The bottom left-hand quadrant is used to provide full text to the lecture. The bottom right quadrant is used to provide a PowerPoint slide relevant to that section of the lecture. Agent99 was developed at the University of Arizona and is structured in a way that facilitates the exploration of the various factors that were discussed in Chapter Two. For this research experiment, five variations of medium presentation were created using the Agent99 trainer. These variations are discussed in greater detail in the following sections.

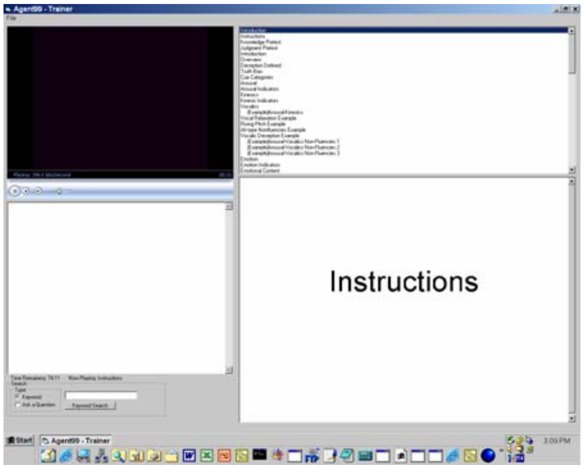

Figure 4. Agent99 Screenshot

What the Instrumental Measurements were Designed to Capture

As presented in the preceding chapters, the aim of this particular experiment was to analyze the comparative effectiveness of five varying forums of electronic computer-based training. Since this experiment, as will be explained in the following section, was a collaborative effort with multiple exploratory goals, the topic of training use was predefined to be in the area of deception detection. Within this area of learning, it was decided that the subtopics of focus would be introductory material and cues training. Introductory material training covered an overview of the deception detection topic from a holistic and general viewpoint. Cues training honed in on the specifics of telltale signs, whose presence during the course of communication may expose the underlying presence of deception. The knowledge transfer that took place between the training medium and the experimental participant was measured using two instruments of testing: a knowledge-based test and an application-based test. The test instruments themselves can be viewed in appendices C, D, E, and F, respectively. It should be noted that the information presented across each of the five varying electronic media was identical, and that the testing instruments used were also identical on an intra-session basis. It should be noted that treatments 4 and 5, as described below, were given additional examples of content already presented, but no new information was given. Additionally, the reader should note that the pretest and the posttest differed by design in order to mitigate the memory effect. The only notable difference between the varying training treatments, with the exception of added examples in treatments 4 and 5, was the interface through which the information was presented.

Characteristics of Population of Interest

The target population selected for sampling in this study was comprised of communications students in a military communications training program. The group primarily consisted of officers, who had obtained a minimum educational level of a bachelor's degree and were comfortable in a computer-mediated environment. The specific demographic characteristics of the sampling population as well as their proportionate makeup are summarized in Table 1 below. The characteristics of this sampling population well suit the requirements of this research study; however, they may have implications in making this study generalizable across other cross-sections of the United States population.

Table 1. Demographic Data

Variables	Military Participants	Reference Percentage
Gender		
Male	159	83.7
Female	31	16.3
Rank		
Civilian	2	1.1
2Lt	177	93.2
1Lt	4	2.1
Capt	5	2.6
Maj	2	1.1
LtCol	0	0
Age		
Average (years)	27.6	N/A
Years in Communications Career Field		
Average (years)	2.5	N/A
Education		
Bachelor's Degree	174	91.6
Master's Degree	14	7.4
Doctoral Degree	0	0
Duty Day on Computer		
Average (%)	51.6	N/A
Off Duty Hours on Computer		
Average (hours per week)	14.7	N/A
Online Courses Taken		
Average	6.4	N/A

Research Methodology

The experimental design which was finally selected and constructed to explore the research question was collaboratively constructed with other academics who were to working on the Air Force Office of Scientific Research (AFOSR) research project on the detection of deceptive communication. The other academics working on the AFOSR research project where primarily from the University of Arizona and Florida State University. Special thanks to Dr. Judee Burgoon, Dr. Joey George, Dr. Mark Adkins, Lt.

Col. Biros, John Kruze, Karl Wiers, and Christopher Steinmeyer. These individuals were responsible in large part for the creation of the lecture, the measuring instruments, and programming involved in the Agent99 production. The final research design selected, as agreed upon by the research teams, was a quasi-experiment and is graphically depicted in Table 2 below (Dooley, 2001). A quasi-experimental approach was taken because a fully random assignment of subject groups was not viable, given some scheduling restrictions that were applicable to the sample population; the reasoning is explained in greater detail in the following section. Each of the X_{nn} variables represents a different training construct; these are summarized in Table 3.

Table 2. Experimental Design

Subject Groups:	
1	$O_1\ X_{11}\ O_1\ O_1\ X_{12}\ O_1$
2	$O_2\ X_{21}\ O_2\ O_2\ X_{22}\ O_2$
3	$O_3\ X_{31}\ O_3\ O_3\ X_{32}\ O_3$
4	$O_4\ X_{41}\ O_4\ O_4\ X_{42}\ O_4$
5	$O_5\ X_{51}\ O_5\ O_5\ X_{52}\ O_5$

Table 3. Treatment Sessions

	X_{11}	X_{21}	X_{31}	X_{41}	X_{51}
Session 1: Intro. Material	Video Lecture	Linear Agent99	Non-Linear Agent99 +Ask a Question	Non-Linear Agent99 +Ask a Question +More Content	Non-Linear Agent99 +Ask a Question +More Content +Quizzes
	X_{12}	X_{22}	X_{32}	X_{42}	X_{52}
Session 2: Cues	Video Lecture	Linear Agent99	Non-Linear Agent99 +Ask a Question	Non-Linear Agent99 +Ask a Question +More Content	Non-Linear Agent99 +Ask a Question +More Content +Quizzes

Each of the subject groups, including the control group, begins at time zero with an initial observation, depicted by "O". The initial observation was used both to collect demographic information from each participant and to gather baseline experimental data—a baseline from which each experimental participant was able to either improve or deteriorate. This observation included pretests in both knowledge and application of the material presented during the treatment session for each experimental group. (Each of the tests alluded to in this paragraph is illustrated in greater detail in later sections of this chapter). Following the pretest, each participant group was exposed to their distinct training regimen—depicted by "X". The training regimen was immediately followed by a posttest assessment in both knowledge and application. This process was repeated for an additional training regimen, for each experimental group, barring the control group. Analyzed experimental results are discussed in detail in Chapter Four; summarized results are available in Appendices H and I. The five varying treatment measures are summarized in the following paragraphs.

Experimental Treatment Group 1

This treatment group was trained with both the introductory material and cues areas of emphasis via a pre-constructed lecture video. Several factors were considered in the selection of a taped lecture, versus a real-time and in-person lecture. Firstly, the video lecture lent itself well to the electronic medium topic of the research problem, as the video was in a format that was viewable on personal computers. Secondly, an internal validity consideration was made, in response to the anticipated condition of the experimental environment that the same lecture had to be presented multiple times to accommodate the schedule of the experiment participants—ensuring that the lecture that each participant viewed would be identical, thereby mitigating the potential for response or instrumentation threats. Such errors, if uncontrolled for, have the potential to cause skewing of the data. The video lecture itself was prepared at a media laboratory on Wright Paterson Air Force Base in July 2003. It should also be noted that the video lecture presentation was non-interactive.

Experimental Treatment Group 2

This treatment group was trained via a linear version of Agent99. The concept of linear training is that the order of the training material is pre-determined and unchangeable by the participant. The linear version of Agent99 added features to the basic video presented to treatment group 1: in addition to the basic video, a content outline—which was accompanied by PowerPoint slides of the material—was presented. Additionally, participants were able to see a full transcript of the material presented, as it was being spoken on the video. Hence, the sum media richness of the presentation was increased. However, as in treatment group 1, the order and timing of the material presented could not be altered by the participant.

Experimental Treatment Group 3

This treatment group was trained via a non-linear version of Agent99 with an interactive "ask-a-question" routine. The concept of non-linear training is that the order in which the training material is presented can be altered by the experimental participant. The rationale behind this feature is that it provided the trainee with additional flexibility of tailoring the educational program to suit his or her particular learning style. The "ask-a-question" routine was an additional feature that allowed the trainee to prompt the computer for additional information on select topics. This could be done at any time during the lecture presentation.

Experimental Treatment Group 4

This treatment group was trained via a non-linear version of Agent99 with an interactive "ask-a-question" routine and some extra content, in the form of extra examples of the topic areas presented. The additional content included additional examples of the same material—aimed at reinforcing the content already presented.

Experimental Treatment Group 5

This treatment group was trained via a non-linear version of Agent99 with an interactive "ask- a-question routine", additional content, and a quiz routine. The quizzes (12 for the introductory session, and 14 for the cues session) were dispersed throughout the lecture and had to be answered by the participant before continuing with the lecture. The intent was to force participants to be engaged with the material.

While the topics of interest for the purpose of training were predetermined, they adapted themselves well to the research and exploration of the research area of interest. The intent of the particular design construct was to validate or reject the research hypotheses presented in Chapter Two. The design construct was established in such a way that the pre-learning and post-learning performance levels of all participant groups could be analyzed for significant findings. Additionally, as an important predetermining factor to internal validity, the pre-learning scores across the groups could be compared to establish a reasonable assertion of cross-group equivalence. The design considerations taken with respect to experimental internal validity are presented further in the following section.

The original intent was for the establishment of a true controlled experiment, with a complete randomization of the sampling population into the various treatment and control groups; true experimental designs are known to yield the highest internal validity (Dooley, 2001). Environmental experimental conditions were such that limitations were forced into the research process. In a true experiment, every member of each distinct treatment and control group is randomly assigned from the same sampling pool of subjects (Dooley, 2001). Given the rigid classroom environment of the study, all the sampling groups where predetermined as classes ranging from 16 to 20 students. Experimental resources and time were provided to us by the school under the stipulation that the students' regular class schedules could not be changed. However, original class pooling was not predetermined by any measurable performance factors; this assurance was obtained by questioning the faculty. Therefore, although a formal group equalization did not occur, the groups could still be seen as be equally composed, conceptually. This assumption was tested during the analysis of the data. Despite this limitation, a degree of randomness was inserted into the construction of the quasi-experiment by the arbitrary assignment of the groups to each distinct set of training treatments or to the control group. Thus, given this element of randomness, it can be rationally asserted that the experiment conducted had greater internal validity than that of a baseline quasi-experiment. Additionally, the insertion of a pretest for each experimental group allowed for the testing of pretreatment equivalence across groups; these calculations are presented in Chapter Four.

The Instruments of Measurement

As alluded to in the previous section, the assessment used within this experiment consisted of a knowledge pretest and a knowledge posttest. The test instruments can be viewed in Appendixes C through G. When reviewing the test instruments the reader should note that each test was of identical format (multiple choice with four options), with the identical number of questions (10). It should also be noted that the tests across all treatment groups were identical and that no participant was given the same test more than once. The implications this had on the research experiment are delved into further in the following section where considerations to the selection of the experimental design are discussed.

Synopsis of Considerations that Lead to Experimental Design Selection

Several considerations were made in the design of the experiment to maximize internal validity and exert control on experimental shortcomings. One of the prime experimental goals was to validate a relationship of causation between training construct and the impact on knowledge transfer and tacit understanding. As such, a longitudinal experimental design was selected. This aspect of the experimental design mitigated the threat of reverse causation, problematic in pure correlational research designs (Dooley, 2001). Additionally, time threats such as sample mortality and maturation were mitigated by the short time interval between treatment and effect measurement; the intervals between treatment and observation were all under one hour (Dooley, 2001). Given the limited availability in the number of the deception scenarios that each group could be exposed to, an unavoidable (due to time constraints) threat of instrumentation was an experimental vulnerability. This problem arose because the measurement tool for measuring participants' ability to detect deception was particularly abstract. Although the difficulty level of each scenario had been previously scored by experts in the field, such an assessment is qualitative and not a guarantee that the pretests and posttests are of exactly the same difficulty. A less difficult posttest would be forecasted to yield a higher score, even in the absence of real learning. However, since the hypotheses addressed in this study focus on the differential or delta between pretest and posttest treatment observations, the impact of this should be mitigated: the bias that an unequal test would introduce would be equivalent across all treatment groups and therefore be transparent when looking only at the differential. The potential internal validity threat of reactivity was also considered. Given the experimental design, the potential existed for experiment participants to react to the content of the pretest in such a manner as to influence the posttest results. However, as explored earlier, since the research problem being considered is primarily accounted for by cross-group comparisons, the impact of this effect should be minimized, because each group was subjected to identical pretests and posttests.

Permission to Conduct Research

This research was aimed at accomplishing a comparative analysis of training techniques for teaching the abstract art of deception detection to a sampling population of military communications students. All aspects of this research were conducted in accordance with human experimentation requirements as outlined in Air Force Instruction 40-402 (SECAF, 1 Sep 2000). An exemption was requested in order to use

Department of Defense personnel as volunteers for electronic learning research. This exemption was granted by the Air Force Research Laboratory Human Effectiveness Directorate (AFRL/HEH). The research was conducted under AFRL/HEH Case Log approval number F-WR-2003-0080-E and AFIT case log number 2003-047, granted on 02 September, 2003. See Appendix B. AFRL Approval Letter.

Experiment

The experiment itself was conducted entirely based on the design previously presented in Table 2 and Table 3. The timeline of the experiment itself, as pertaining to each observation and corresponding treatment, is summarized in Table 4 below. The between-session timeline is also summarized in Table 4 below.

Table 4. Instrumental Experiment Timeline

Session 1: Introductory Overview	**Time (min)**	**Total Session Time (min)**
Introductory Knowledge Pretest	15	15
Judgment Accuracy Pretest	15	30
Introductory Training	60	90
Introductory Knowledge Posttest	15	105
Judgment Accuracy Posttest	15	120
Session 2: Cues Training	**Time (min)**	**Total Session Time (min)**
Cues Knowledge Pretest	15	15
Judgment Accuracy Pretest	15	30
Cues Training	60	90
Cues Knowledge Posttest	15	105
Judgment Accuracy Posttest	15	120

Planned Analysis Architecture (Operationalizing)

Using the foundational understanding of the differentiation between the treatment groups, as described at length previously, it is possible to draw a construct map between the treatments and the independent variables addressed by the hypotheses in Chapter

Two. Conceptually, each hypothesis will be linked to two treatments: one will form the baseline, and the second will form the basis for measuring the deviation from the baseline, attributable to the hypothesis variable in question.

Considering the first hypothesis—that more media richness within the electronic learning medium will result in significantly higher student learning differentials than media with less media richness—the primary construct of interest is media richness. This can rationally be linked to a differential between experimental treatment one and treatment two. As expounded on earlier, treatment one is based purely on a video presentation of the topic of interest, while treatment two involves the addition of PowerPoint slides and a full transcript of the lecture presentations. These, according to Daft and Lengel's (1986) media richness hierarchy presented in Chapter Two, lead to a sum total increase in media richness. Considering that the variable of time was controlled and learning was measured during the experiment via pretest and posttests (as presented in Appendices C through F), then Carroll's (1963) model of school learning applies (a detailed description of this model was presented in chapter 2), and it can be rationally inferred that any increase in learning can be attributed to the fact that "instruction has improved." If instruction then does indeed improve, the only variation of change between these two treatments would be the introduction of media; hence, a statistically significant increase in learning between experimental treatments one and two would be in support of the first hypothesis.

Considering the second hypothesis—that an electronic medium with more content flexibility will result in significantly higher student learning differentials than media with less content flexibility—the primary variable of interest is the introduction of content flexibility. This can rationally be linked to a differential between experimental treatment two and treatment three. As presented at length earlier, experimental treatments two and three are identical, except for the introduction of content flexibility. Participants in treatment three were allowed to jump around within the presentation, both forwards and backwards, through the use of hyperlinks within the table of contents, and through the use of a search routine that allowed them to jump to a topic of interest. Again, Carroll's (1963) model of school learning applies, and it can be inferred that an increase in participant learning would be attributable to improved instruction; in this case, the improved instruction would come through the form of content flexibility within the learning medium. A statistically significant increase in learning between experimental treatments two and three would be in support of the second hypothesis.

Regarding hypothesis three—that the presence of pop-up quizzes that force the engagement of the student with the material will result in significantly higher student learning differentials—the primary variable of interest is the forced engagement of the student with the material. This can rationally be linked to a differential between experimental treatment four and treatment five. As thoroughly presented earlier, the only difference between treatment four and five was the introduction of pop-quizzes. Experimental participants were forced to pause and answer the quizzes before being allowed to continue with the presentation. Again, Carroll's (1963) model of school learning applies, and it can be inferred that an increase in participant learning would be attributable to improved instruction; in this case, the improved instruction would come through the form of forced engagement within the learning medium. A statistically significant increase in learning between experimental treatments four and five would be in support of the third hypothesis.

Considering the fourth hypothesis—that the sum total of hypotheses one and two will cumulatively result in significantly higher student learning differentials—the two primary variables of interest are media richness and content flexibility. This can rationally be linked to a differential between experimental treatment one and treatment three. Treatment three introduces additional media richness and content flexibility to the material presented in treatment one. Again, Carroll's (1963) model of school learning applies, and it can be inferred that an increase in participant learning would be attributable to improved instruction; in this case, the improved instruction would come through the combination of media richness and content flexibility within the learning medium. A statistically significant increase in learning between experimental treatments one and three would be in support of the fourth hypothesis.

Considering the fifth hypothesis, that the sum total of hypotheses one and three will cumulatively result in significantly higher student learning differentials—the two primary variables of interest are media richness and forced engagement. As the reader may already have ascertained, this hypothesis is unfortunately not testable in the given experimental construct. Due to the joint effort of the experiment and the limitations of the experimental setting, the number of experimental groups was limited to five; hence, no treatment existed where the only change variables were media richness and forced engagement. In order to test this hypothesis, a linear presentation medium with the addition of pop-up quizzes would have to be introduced, to compare against the baseline of treatment one. This hypothesis was only mentioned in order to present a complete experimental approach; however, since it is not testable given the experimental design, it will be omitted from the remainder of this text.

Regarding the sixth hypothesis—that the sum total of hypotheses two and three will cumulatively result in significantly higher student learning differentials—the two primary variables of interest are content flexibility and forced engagement. This can rationally be linked to a differential between experimental treatment two and treatment five. Treatment five differs from treatment two because of the introduction of content flexibility and pop-up quizzes. Again, Carroll's (1963) model of school learning applies, and it can be inferred that an increase in participant learning would be attributable to improved instruction; in this case, the improved instruction would come through the combination of content flexibility and forced engagement within the learning medium. A statistically significant increase in learning between experimental treatments two and five would be in support of the sixth hypothesis.

Considering the seventh hypothesis—that the sum total of hypotheses one, two and three will cumulatively result in significantly higher student learning differentials—the three primary variables of interest are media richness, content flexibility, and forced engagement. This can rationally be linked to a differential between experimental treatment one and treatment five. Treatment five differs from treatment one because of the introduction of increased media richness, content flexibility, and pop-up quizzes. As before, Carroll's (1963) model of school learning applies, and it can be inferred that an increase in participant learning would be attributable to improved instruction; in this, case the improved instruction would come through the combination of media richness, content flexibility, and forced engagement within the learning medium. An increase in learning between experimental treatments one and five would be in support of the seventh hypothesis.

The construct links between the experimental treatments and the research hypotheses are summarized in Table 5 below. These will form the basis of the planned statistical analysis, described in the following section, that will be used to draw rational inferences from the experimental data.

Table 5. Linking Hypotheses with Experimental Treatments

Hypothesis	Treatment Group Pairs	
H1	1	2
H2	2	3
H3	4	5
H4	1	3
H6	2	5
H7	1	5

Planned Statistical Approach

Since conceptually the data analysis involves calculating the effect of some independent variables on a dependent variable, it can be rationally concluded that a sound statistical approach would be to begin with an analysis of variance (ANOVA). The ANOVA "partitions the total variation of a sample into components" and computes an F-test which can be used to make a judgment on the model's effectiveness. The ANOVA will form the basis for the Tukey-Kramer Honestly Significant Difference (HSD) test to do a simultaneous test across all treatment group means. This is appropriate since, given the link between the experimental treatments and the research hypotheses presented above, the "family of interest is the set of all pair wise comparisons of factor level means" (Neter, Kunter, Nachtsheim, & Wasserman, 1996). A significant difference revealed by the Tukey-Kramer HSD would be of note if it fell into the framework summarized in Table 5. The null hypothesis of the Tukey-Kramer HSD test is that the pairs held in comparison are equal or $\mu_i - \mu_{i'} = 0$. Hence, a positive outcome from this test would lead to a failure to reject the null hypothesis, and the conclusion that the means are not different. In this experiment, such an outcome would lead to the conclusion that the variable of interest had no effect on participant learning. For the bases of the analysis within this research project, an alpha level of 0.05 will be considered adequate. This process will be expounded on as the analysis process is illustrated in Chapter 4.

Since the ANOVA is a parametric test, several assumptions must hold true about the data—and therefore must be tested. These assumptions include continuity of the dependent variable, normality of the regression residuals, and constant variance (Neter et al., 1996). Independence of the stochastic data points will be assumed because of the pseudo-random design of the experiment. Additionally, when using an ANOVA it is prudent to test outliers that might exert an inordinate influence on the model—hence causing a skew in the results. The assumption of continuity will be checked via a histogram of the independent variable, in this case test score. The assumption of normality will be tested for via a histogram of the studentized residuals and a Shapiro-Wilk Goodness-of-Fit test, comparing the residuals to a normal quantile plot. Since the null hypothesis of the Shapiro-Wilk test is that normality holds true, a p-value greater than 0.05 is desired, which would lead to a conclusion of failing to reject the null hypothesis. The assumption of constant variance will initially be tested via a scatter plot of the studentized residuals; the desired outcome would be that no pattern is evident within the scatter plot. The outcomes of this test will be confirmed via the O'Brian test, Brown-Forsythe test, Levene's test, and Bartlett's test for constant variance. The null hypothesis in these tests is that the variance is constant; hence an F-stat value of greater than 0.05 is desired. Finally, in order to test for the inordinate influence of outlying data points, a Cook's Distance test will be conducted. As long as the Cook's Distance value for each data point is calculated to have a weight less than 10-20%, it is considered to be within the normal range, and deemed not to exert an undue influence (Neter et al., 1996).

The statistical process outlined above will be conducted initially on the pretest scores across all treatment groups for both the introductory session and the cues session of the experiment. The expected outcome of the initial test, given the pseudo-random assignment of participants to treatment groups, is that all treatment groups will participate equally well or poorly on the pretest. Hence it is expected that there will be a failure to reject the Tukey-Kramer null hypothesis across all pair-wise comparisons. This will establish group equivalence at the start of the longitudinal experiment. Secondly, the same statistical process outlined above will be conducted on the delta or differential between the pretest and posttest scores for both the introductory session and the cues session of the experiment. These tests will form the basis for either supporting or failing to support the research hypothesis. The analysis process and results are presented in detail in the following chapter.

IV. Data Analysis

Overview

The raw data collected, as previously described, was located in over 370 separate data files. Each session of a student with the training material created a separate data file. These data files where concatenated two primary Microsoft Excel data spreadsheets via a software-driven compiler: one for the Introductory Session, and one for the Cues Session. The data was subsequently cleaned up and translated prior to analysis, firstly by the University of Arizona research team and then by myself. Fundamentally, the analysis followed the investigative plan laid out in the preceding chapters using an Analysis of Variance (ANOVA) statistical method with a Tukey or Tukey-Kramer HSD test. An ANOVA is valid for experimental and observational data, and for both single-factor and multifactor studies; hence, it suits the needs of this study (Neter et al., 1996). All the applicable assumptions were also tested. Unless otherwise stated, all statistical analysis was done through the JMP (Release: 5.0.1.2) Statistical Discovery Software.

Outline of Analysis

The analysis delineated below is laid out in the following order: Firstly, an analysis of the introductory session knowledge pretest data is followed by the analysis of the introductory session knowledge pretest/posttest differential. Secondly, cues session knowledge pretest data is followed by the analysis of the cues session knowledge pretest/posttest differential. Thirdly, a summary of the analysis results from the different analysis phases is presented. Fourthly, an interpretive analysis of how the statistical numbers relate to the hypotheses drawn in Chapter 2 is offered.

Introductory Knowledge Pretest Analysis

The pretest scores were analyzed to establish baseline equivalence across all five treatment groups. All groups, as formerly discussed, were assigned in a pseudorandom manner and were thus expected to be equivalent subject groups. Upon initial review of the data for completeness, it was discovered that one of the entries (entry 5005, see Appendix H) was incomplete, and therefore removed from the data set. The remaining 189 pretest knowledge answers were translated into a composite percentile score for each entry by comparing the student answers to an answer key, and giving a score of one for a correct response and a score of zero for an incorrect response. The total point score accumulated was divided by 10, representing the highest possible score, yielding an aggregate percentile score. A summary of the post-translation data set can be viewed in Appendix H. Additionally, to get an initial understanding of the data, the means of and standard deviations of each treatment group's pretest scores were calculated. These can be viewed in Table 6 below. The means, as presented, appear to be very close together in magnitude, hovering right around 0.51. This adheres to the expectation that all the treatment groups start at an identical state of knowledge.

Table 6. Intro Pretest Means and Standard Deviations

Treatment	Mean	Standard Deviation
Group 1	0.5452	0.1234
Group 2	0.4977	0.1389
Group 3	0.5333	0.1241
Group 4	0.5114	0.1498
Group 5	0.5225	0.1209

However, this initial feel for the data was followed by a more rigorous statistical approach. Following data translation, the pretest scores were visually analyzed for the initial ANOVA statistical assumption of continuity via a histogram . The shape of the histogram suggests that the data is continuous enough to proceed with the ANOVA approach to analysis. The histogram is shown in Figure 5 below.

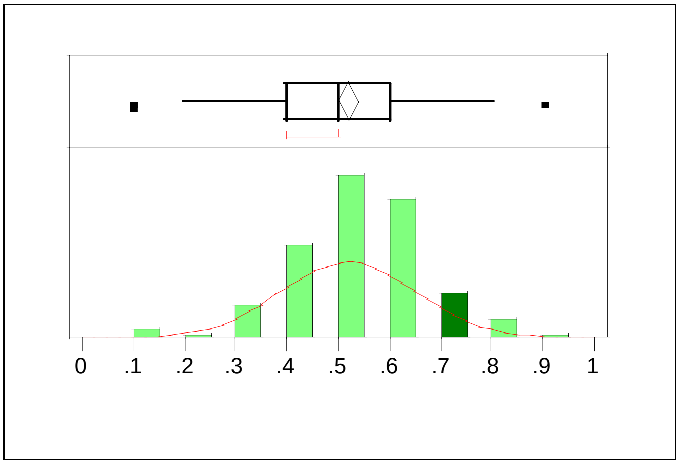

Figure 5. Histogram of Knowledge Pretest Data

The pretest data was then analyzed via the ANOVA statistical method. The initial ANOVA results are presented in Figure 6 below. The rather large ANOVA F-Statistic of 0.5874 gives an initial indication, in line with expectations, that treatment sessions as a variable, do not exert much predictive influence on student learning. In order to test for the statistical validity assumptions of independence of the residuals, normality of the residuals, and constant variance of the residuals, the Studentized Residuals and Cook's D influence factors were saved for further investigation.

Summary of Fit

Rsquare	0.015157
Adj Rsquare	-0.00625
Root Mean Square Error	0.133255
Mean of Response	0.519577
Observations (or Sum Wgts)	189

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Treatment	4	0.0502844	0.012571	0.7080	0.5874
Error	184	3.2672818	0.017757		
C. Total	188	3.3175661			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
1	31	0.545161	0.02393	0.49794	0.59238
2	44	0.497727	0.02009	0.45809	0.53736
3	30	0.533333	0.02433	0.48533	0.58133
4	44	0.511364	0.02009	0.47173	0.55100
5	40	0.522500	0.02107	0.48093	0.56407

Std Error uses a pooled estimate of error variance

Figure 6. Pretest ANOVA

In order to validate the assumption of normality on the residuals, the Studentized Residuals were plotted via a histogram, the output of which is illustrated in Figure 7 below. The plot of the distribution, as highlighted by the overlay of the normal curve, suggests that normality holds true. A Shapiro-Wilk goodness of fit test was then conducted to provide an objective analysis of the initial indications. The result of this test is illustrated in Figure 8 below. Given an assumed alpha of 0.05, the statistical null hypothesis that the residuals follow the normal distribution is not rejected since 0.1034 is not less than the alpha of 0.05. Hence, the assumption of normality for the ANOVA test is satisfied.

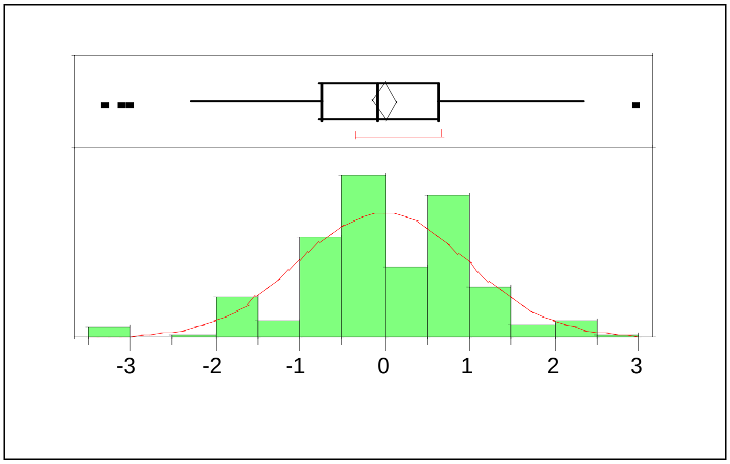

Figure 7. Histogram of Studentized Residuals

Goodness-of-Fit Test

Shapiro-Wilk W Test

W	Prob<W
0.974793	0.1034

Figure 8. Test for Normality

In order to validate the assumption of constant variance of the residuals, a scatter plot of the Studentized Residuals was accomplished to provide an initial indication of any trends inherent in the data (Neter et al., 1996). This plot is illustrated in Figure 9. Since the points on the chart seem to fall within the same general range across all treatment groups, an initial indication is given that the assumption of constant variance holds true. Four separate tests (O'Brien test, Brown-Forsythe test, Levene's test, and Bartlett's test) were conducted to provide an objective statistical basis to further explore the result drawn from Figure 9. The output from these four tests is illustrated in Figure 10 below. Each of these tests confirm the initial indication that the residual variances are equal across the five treatment groups; none of the Prob > F statistical outcomes is under the assumed alpha of 0.05, hence the null hypothesis of equal variance is not rejected. Also, the required assumption of constant variance for the ANOVA test is therefore satisfied.

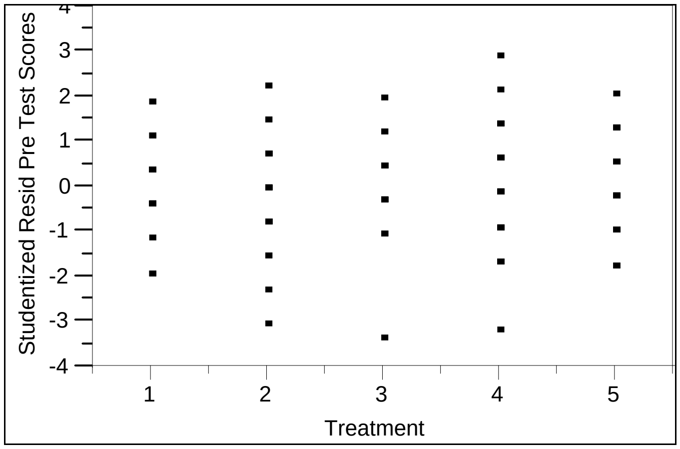

Figure 9. Plot of Residuals

Test	F Ratio	DFNum	DFDen	Prob > F
O'Brien[.5]	0.4830	4	184	0.7482
Brown-Forsythe	0.4574	4	184	0.7669
Levene	0.4736	4	184	0.7551
Bartlett	0.6762	4	.	0.6084

Figure 10. Tests For Equal Variance

Additionally, the saved Cook's D influence statistic was graphically analyzed via a scatter plot in order to examine the ANOVA for the presence of any data entries that exert an inordinate amount of influence on the test results (Neter et al., 1996). The presence of such a statistic could skew analysis results. A Cook's D influence statistic under 10-20% is taken to be acceptable (Neter et al., 1996). As illustrated below in Figure 11, all the Cook's D influence statistics fall within the acceptable range, with a peak level of 7.5%.

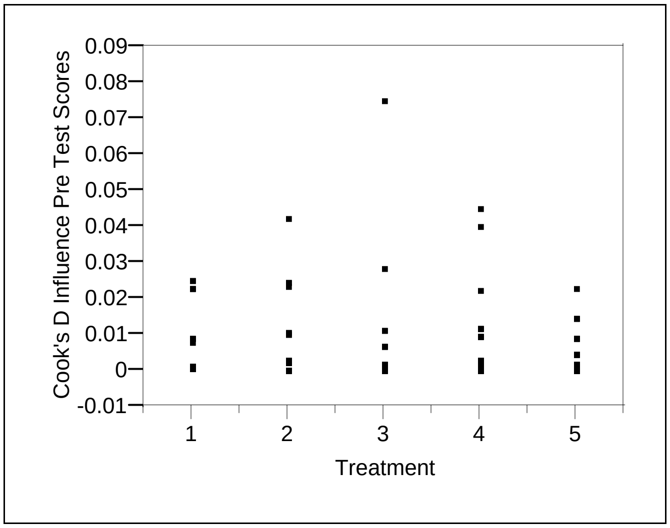

Figure 11. Pretest Cook's D

The initial ANOVA output, as shown in Figure 6, suggests that the means for all treatment groups are roughly identical. Additionally, the F-test p-value—which tests for the equality of factor level means—of 0.5874 leads to the conclusion, in statistical terms of failing to reject the ANOVA's null hypothesis, that none of the variables have any predictive power toward forecasting pretest scores (Neter et al., 1996). This affords an initial indication that the starting condition of all five treatment groups was statistically equivalent. In order to further analyze this conclusion, a Tukey-Kramer HSD was conducted to accomplish a multiple comparison across all the treatment groups. The output is depicted in Figure 12 and visually illustrated in Figure 13 below.

Comparisons for all pairs using Tukey-Kramer HSD

q*	Alpha
2.75502	0.05

Abs(Dif)-LSD

	1	3	5	4	2
1	-0.09325	-0.08219	-0.06519	-0.05229	-0.03865
3	-0.08219	-0.09479	-0.07783	-0.06495	-0.05132
5	-0.06519	-0.07783	-0.08209	-0.06907	-0.05543
4	-0.05229	-0.06495	-0.06907	-0.07827	-0.06463
2	-0.03865	-0.05132	-0.05543	-0.06463	-0.07827

Positive values show pairs of means that are significantly different.

Level		Mean
1	A	0.54516129
3	A	0.53333333
5	A	0.52250000
4	A	0.51136364
2	A	0.49772727

Levels not connected by same letter are significantly different

Figure 12. Pretest Tukey-Kramer HSD

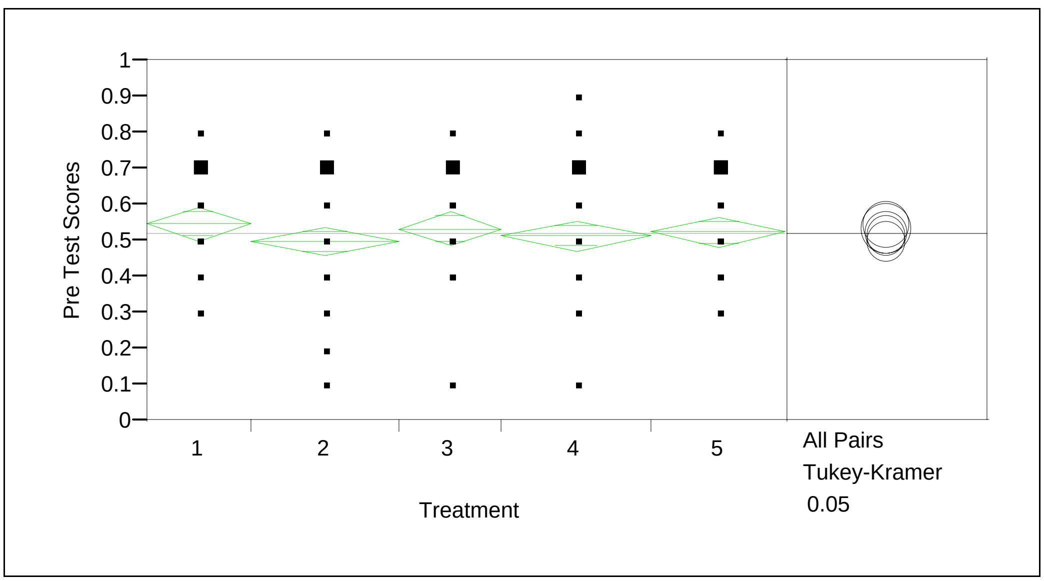

Figure 13. Pretest Tukey-Kramer HSD Graphical Illustration

The results of the Tukey-Kramer HSD confirm the initial indication that none of the treatment groups differed statistically. As shown in Figure 12, given an alpha of 0.05, none of the Tukey-Kramer HSD factors showed up as positive: (positive numbers imply a significant difference between pairs). This statistical inference is visually illustrated in Figure 13, where it should be noted that all the diamonds are within the same general range, and all the circles overlap one another. Restated, no statistical difference in knowledge pretest outcomes across the five treatment groups is evident within the data.

Introductory Knowledge Pretest/Posttest Differential Analysis

The differential between the pretest and posttest knowledge scores was analyzed to look for evidence that different treatments had varying effects on experimental participants. Since the preceding analysis for the knowledge pretest scores established baseline equivalence, finding significant variations in the pretest to posttest score differential would suggest that treatment had an impact on the learning result. This section of the analysis of pretest and posttest score differential forms the basis for the testing of the hypothesis. As alluded to in the previous section, each hypothesis construct has been linked to the change in learning outcomes between two experimental groups. The aim of this phase of the analysis is to statistically assess the presence or absence of such a learning differential. Table 7 below re-summarizes the hypotheses to treatment group construct link.

Table 7. Hypothesis Treatment Group Pairs

Hypothesis	Treatment Group Pairs	
H1: Having more media richness within the electronic learning medium will result in significantly higher student learning differentials than media with less media richness.	1	2
H2: An electronic medium with more content flexibility will result in significantly higher student learning differentials than media with less content flexibility.	2	3
H3: The presence of forced engagement of the student with the material will result in significantly higher student learning differentials than media with no forced engagement.	4	5
H4: The sum total of media richness and content flexibility will cumulatively result in significantly higher student learning differentials than that in media with lower levels of media richness and less content flexibility.	1	3
H6: The sum total of content flexibility and forced engagement will cumulatively result in significantly higher student learning differentials than that in media with less content flexibility and no forced engagement.	2	5
H7: The sum total of media richness, content flexibility, and forced engagement will cumulatively result in significantly higher student learning differentials than that in media with lower levels of media richness, less content flexibility, and no forced engagement.	1	5

Analysis in View of Research Hypotheses

In order to test the hypotheses, the same 189 pretest knowledge entries used in the knowledge pretest analysis were used, and each of the participant scores was translated into a composite percentile in the same translation manner already expounded upon in the previous section. The raw data results can be viewed in Appendix H. In order to get an initial understanding of how the data might relate to the hypotheses, the mean and standard deviation was calculated for each treatment group for both their posttest score and for the differential between their posttest and pretest scores. These can be viewed in

Table 8 and Table 9 respectively. As shown, group 5 (as expected) seemed to have the most improvement in its score, giving an initial indication that there was some score variation among groups, suggesting that perhaps some learning improvement may have taken place due to instructional improvement. However, regression analysis and the Tukey HSD Test will help determine statistical significance.

Table 8. Intro Posttest Means and Standard Deviations

Treatment	Mean	Standard Deviation
Group 1	0.7677	0.1600
Group 2	0.7454	0.1606
Group 3	0.7733	0.1413
Group 4	0.7250	0.1587
Group 5	0.8100	0.1766

Table 9. Intro Score Differential Means and Standard Deviations

Treatment	Mean	Standard Deviation
Group 1	0.2226	0.1961
Group 2	0.2477	0.1798
Group 3	0.2400	0.1754
Group 4	0.2136	0.1954
Group 5	0.2875	0.1964

Following the initial mean analysis, the score differentials were visually analyzed for the initial ANOVA statistical assumption of continuity via a histogram. The shape of the histogram suggests that the data is continuous enough to proceed with the ANOVA approach to analysis. The histogram is shown in Figure 14 below.

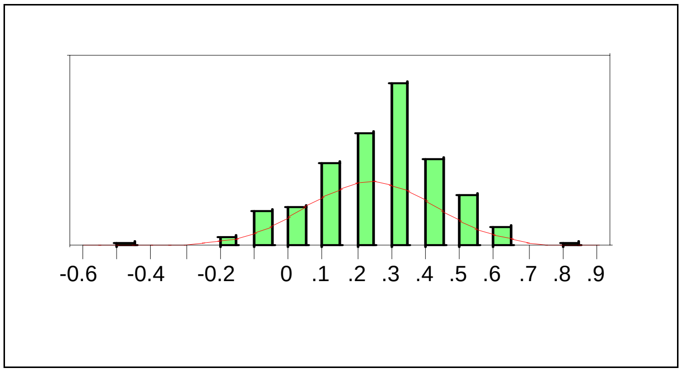

Figure 14. Histogram of Knowledge Pretest/Posttest Differentials

The differential data was then analyzed via the ANOVA. The initial ANOVA results are presented in Figure 15 below. The rather large ANOVA F-Statistic of 0.4539 gives an initial indication, contrary to expectations, that treatment sessions as a variable, do not exert much predictive influence on student learning. In order to test for the statistical validity assumptions of independence of the residuals, normality of the residuals, and constant variance of the residuals, the Studentized Residuals and Cook's D influence factors were saved for further investigation.

Summary of Fit

Rsquare	0.019592
Adj Rsquare	-0.00172
Root Mean Square Error	0.188984
Mean of Response	0.242857
Observations (or Sum Wgts)	189

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Treatment	4	0.1313227	0.032831	0.9192	0.4539
Error	184	6.5715345	0.035715		
C. Total	188	6.7028571			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
1	31	0.222581	0.03394	0.15561	0.28955
2	44	0.247727	0.02849	0.19152	0.30394
3	30	0.240000	0.03450	0.17193	0.30807
4	44	0.213636	0.02849	0.15743	0.26985
5	40	0.287500	0.02988	0.22855	0.34645

Std Error uses a pooled estimate of error variance

Figure 15. ANOVA of Pretest/Posttest Differentials

In order to validate the assumption of normality on the residuals, the Studentized Residuals were plotted via a histogram, the output of which is illustrated in Figure 16 below. The plot of the distribution, as highlighted by the overlay of the normal curve, suggests that normality holds true. A Shapiro-Wilk goodness-of-fit test was then conducted to provide an objective analysis of the initial indications. The result of this test is illustrated in Figure 17 below. Hence, the statistical null hypothesis that the residuals follow the normal distribution is not rejected, since 0.81 is not less than the alpha of 0.05. Hence, the assumption of normality for the ANOVA test is satisfied.

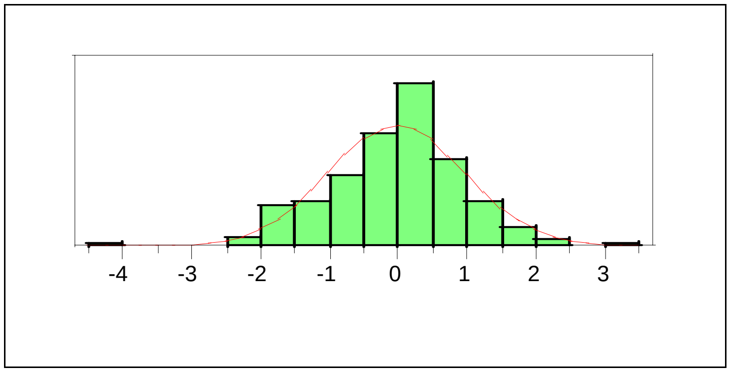

Figure 16. Histogram of Studentized Residuals

Goodness-of-Fit Test

Shapiro-Wilk W Test

W	Prob<W
0.987065	0.8195

Figure 17. Test for Normality

In order to validate the assumption of constant variance of the residuals, a scatter plot of the Studentized Residuals was accomplished to provide an initial indication of any trends inherent in the data. This plot is illustrated in Figure 18. Given that the points on the chart seem to fall within the same general range across all treatment groups, an initial indication is given that the assumption of constant variance holds true. Four separate tests (O'Brien test, Brown-Forsythe test, Levene's test, and Bartlett's test) were conducted to provide an objective statistical basis to further explore the result drawn from Figure 18. The output from these four tests is illustrated in Figure 19 below. Each of these tests confirms the initial indication that the residual variances are equal across the five treatment groups; none of the Prob > F statistical outcomes is under the assumed alpha of 0.05; hence, the null hypothesis of equal variance is not rejected. Also, the required assumption of constant variance for the ANOVA test is therefore satisfied.

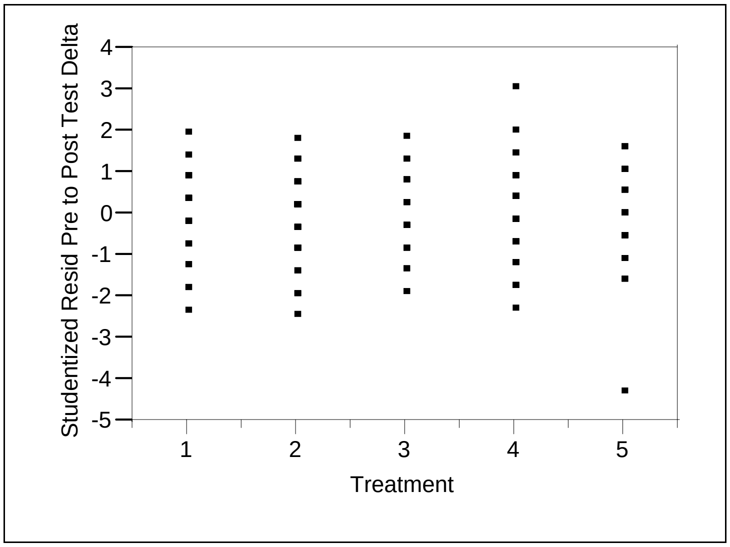

Figure 18. Plot of Residuals

Test	F Ratio	DFNum	DFDen	Prob > F
O'Brien[.5]	0.1169	4	184	0.9764
Brown-Forsythe	0.2260	4	184	0.9236
Levene	0.1656	4	184	0.9556
Bartlett	0.1975	4	.	0.9398

Figure 19. Tests For Equal Variance

Additionally, the saved Cook's D influence statistic was graphically analyzed via a scatter plot in order to examine the ANOVA for the presence of any data entries that exert an inordinate amount of influence on the test results (Neter et al., 1996). The presence of such a statistic could skew analysis results. As stated earlier, a Cook's D influence statistic under 10-20% is taken to be acceptable (Neter et al., 1996). As illustrated below in Figure 20, all the Cook's D influence statistics fall within the acceptable range, with a peak value of 9.1%.

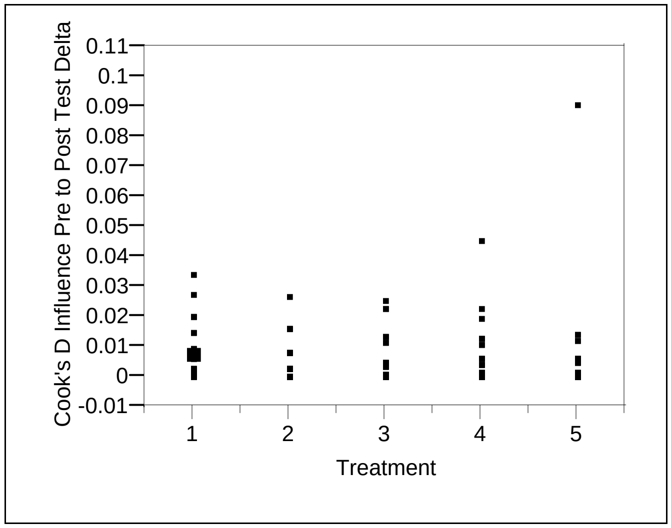

Figure 20. Pretest/Posttest Differential Cook's D

The initial ANOVA output, as shown in Figure 15, suggests that the means for all treatment groups are roughly identical. Additionally, the F-test p-value of 0.4539 leads to the conclusion, in statistical terms, of failing to reject the ANOVA's null hypothesis that none of the variables have any predictive power toward forecasting pretest scores. This affords an initial indication that the starting condition of all five treatment groups was statistically equivalent. In order to further analyze this conclusion a Tukey-Kramer HSD was conducted to accomplish a multiple comparison across all the treatment groups. The output is depicted in Figure 21 and visually illustrated in Figure 22 below.

Comparisons for all pairs using Tukey-Kramer HSD

q*	Alpha
2.75502	0.05

Abs(Dif)-LSD

	5	2	3	1	4
5	-0.11642	-0.07397	-0.07825	-0.05967	-0.03988
2	-0.07397	-0.11100	-0.11555	-0.09694	-0.07691
3	-0.07825	-0.11555	-0.13443	-0.11592	-0.09691
1	-0.05967	-0.09694	-0.11592	-0.13225	-0.11314
4	-0.03988	-0.07691	-0.09691	-0.11314	-0.11100

Positive values show pairs of means that are significantly different.

Level		Mean
5	A	0.28750000
2	A	0.24772727
3	A	0.24000000
1	A	0.22258065
4	A	0.21363636

Levels not connected by same letter are significantly different

Figure 21. Pretest/Posttest Differential Tukey-Kramer HSD

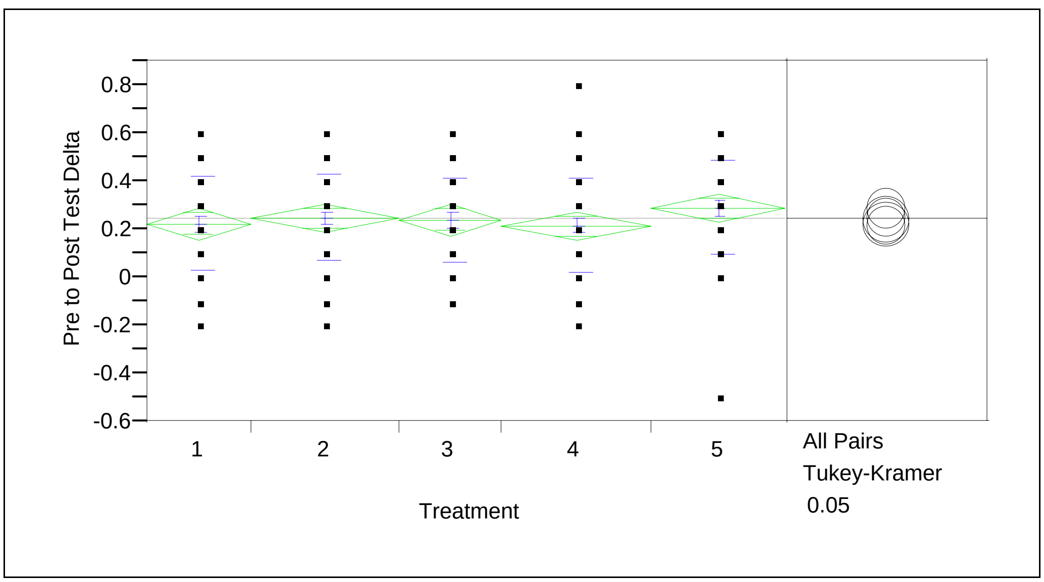

Figure 22. Pretest/Posttest Differential Tukey-Kramer HSD Graphical Illustration

The results of the Tukey-Kramer HSD confirm the initial indication that none of the treatment groups differed statistically. As shown in Figure 21, given an alpha of 0.05, none of the Tukey-Kramer HSD factors showed up as positive: (positive numbers imply a significant difference between pairs). This statistical inference is visually illustrated in Figure 22, where it should be noted that all the diamonds are within the same general range, and all the circles overlap one another. Restated, no statistical difference in knowledge pretest/posttest differentials across the five treatment groups is evident within the data.

Analysis Outcomes in Terms of Experimental Hypotheses

In accordance with the analysis architecture which was previously illustrated in Chapter 3, the statistical outcomes of the Tukey-Kramer HSD were translated into practical outcomes in view of the experimental hypotheses. Each of the hypotheses was rationally linked as a comparative relationship between two treatment groups: the first formed the experimental baseline from which the second treatment group was expected to improve upon, because of the addition of factors such as increased media richness, content flexibility, and forced engagement. The exception to this, as already fully explained in Chapter 3, was the inability to test original hypothesis five. Given the cooperative nature of the experiment and the limited number of treatment groups, a baseline from which this hypothesis could be measured was not conducted; therefore, this hypothesis must be excluded from the analysis. As explained above, the Tukey-Kramer HSD statistical comparison of the means highlights different means by assigning them with different significance levels—these along with the treatment group comparisons and hypotheses results are summarized in Table 10 below. As shown, none of the original seven hypotheses are supported by data.

Table 10. Hypotheses Outcomes on Introductory Session Analysis

Hypothesis	Treatment Groups		Tukey HSD Mean Level		Result
H1	1	2	A	A	Unsupported
H2	2	3	A	A	Unsupported
H3	4	5	A	A	Unsupported
H4	1	3	A	A	Unsupported
H6	2	5	A	A	Unsupported
H7	1	5	A	A	Unsupported

Cues Knowledge Pretest Analysis

The same standard of rigor that was applied to the analysis of the introductory pretest scores carried over into the analysis of the cues pretest data analyzed to establish baseline equivalence across all five treatment group's. All groups, as formerly discussed, were assigned in a pseudorandom manner and were thus expected to be equivalent subject groups. Since the experiment conducted was longitudinal in nature, some participant attrition did occur (see Chapter 3). Four participants did not return for the cues treatment. The remaining 185 pretest knowledge answers were translated into a composite percentile score in the same way as was accomplished for the knowledge pretest data analysis. A summary of the post-translation data set can be viewed in Appendix I. Additionally, the initial mean and standard deviation calculation results are summarized in Table 11 below. The means, as expected, all seem to be very similar ranging from 0.52-0.60. However, to ensure statistical equivalence the same procedures and statistical methodology used to examine session 1 data will also be implemented on the cues dataset.

Table 11. Cues Pretest Means and Standard Deviations

Treatment	Mean	Standard Deviation
Group 1	0.5750	0.1818
Group 2	0.5568	0.1946
Group 3	0.5172	0.2054
Group 4	0.6000	0.1775
Group 5	0.5605	0.2184

Following brief look at pretest score means, the pretest scores were visually analyzed for the initial ANOVA statistical assumption of continuity via a histogram. The shape of the histogram suggests that the data is continuous enough to proceed with the ANOVA approach to analysis. The histogram is shown in Figure 23 below.

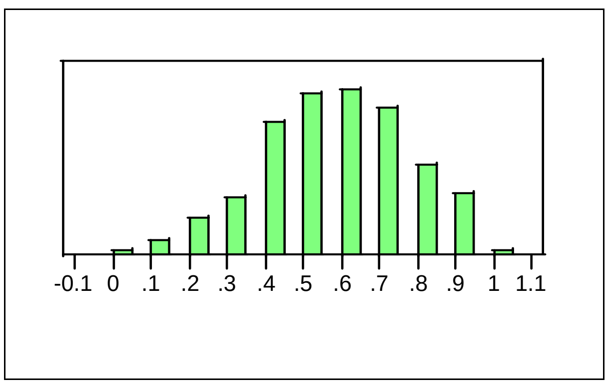

Figure 23. Histogram of Cues Pretest Data

The pretest data was then analyzed via the ANOVA statistical method. The initial ANOVA results are presented in Figure 24 below. The rather large ANOVA F-Statistic of 0.5312 gives an initial indication, in line with expectations, that treatment sessions as a variable, do not exert much predictive influence on student learning. In order to test for the statistical validity assumptions of independence of the residuals, normality of the residuals, and constant variance of the residuals, the Studentized Residuals and Cook's D influence factors were saved for further investigation.

Summary of Fit

Rsquare	0.017316
Adj Rsquare	-0.00452
Root Mean Square Error	0.196704
Mean of Response	0.563784
Observations (or Sum Wgts)	185

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Treatment	4	0.1227268	0.030682	0.7930	0.5312
Error	180	6.9646246	0.038692		
C. Total	184	7.0873514			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
1	28	0.575000	0.03717	0.50165	0.64835
2	44	0.556818	0.02965	0.49830	0.61533
3	29	0.517241	0.03653	0.44517	0.58932
4	41	0.600000	0.03072	0.53938	0.66062
5	43	0.560465	0.03000	0.50127	0.61966

Std Error uses a pooled estimate of error variance

Figure 24. Pretest ANOVA

In order to validate the assumption of normality on the residuals, the Studentized Residuals were plotted via a histogram, the output of which is illustrated in Figure 25 below. The plot of the distribution, as highlighted by the overlay of the normal curve, suggests that normality holds true. A Shapiro-Wilk goodness-of-fit test was then conducted to provide an objective analysis of the initial indications. The result of this test is illustrated in Figure 26 below. Given an assumed alpha of 0.05, the statistical null hypothesis that the residuals follow the normal distribution is not rejected, since 0.1338 is not less than the alpha of 0.05. Hence, there the assumption of normality for the ANOVA test is satisfied.

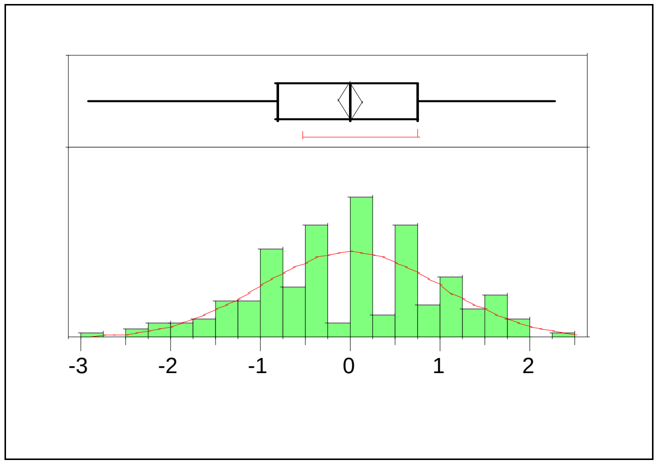

Figure 25. Histogram of Studentized Residuals

Goodness-of-Fit Test	
Shapiro-Wilk W Test	
W	Prob<W
0.975618	0.1338

Figure 26. Test for Normality

In order to validate the assumption of constant variance of the residuals, a scatter plot of the Studentized Residuals was accomplished to provide an initial indication of any trends inherent in the data. This plot is illustrated in Figure 27. Given that the points on the chart seem to fall within the same general range across all treatment groups, an initial indication is given that the assumption of constant variance holds true. Four separate tests (O'Brien test, Brown-Forsythe test, Levene's test, and Bartlett's test) were conducted to provide an objective statistical basis to further explore the result drawn from Figure 27. The output from these four tests is illustrated in Figure 28 below. Each of these tests confirm the initial indication that the residual variances are equal across the five treatment groups. None of the Prob > F statistical outcomes is under the assumed alpha of 0.05; hence, the null hypothesis of equal variance is not rejected. Also, the required assumption of constant variance for the ANOVA test is therefore satisfied.

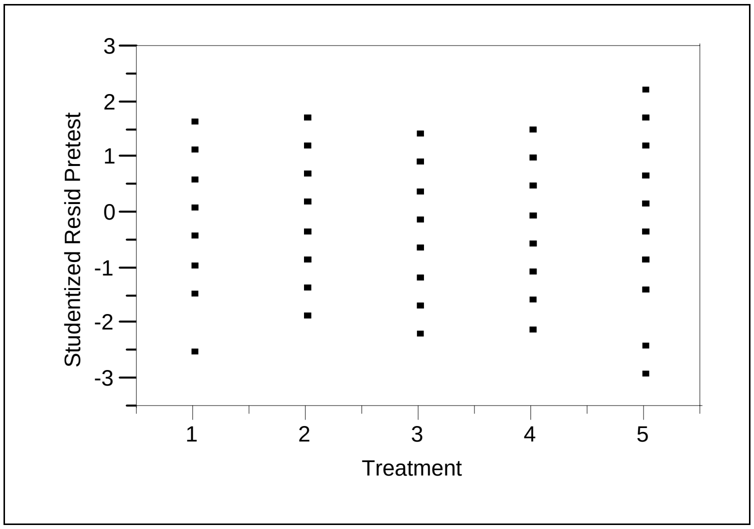

Figure 27. Plot of Residuals

Test	F Ratio	DFNum	DFDen	Prob > F
O'Brien[.5]	0.6656	4	180	0.6167
Brown-Forsythe	0.7991	4	180	0.5272
Levene	0.9924	4	180	0.4131
Bartlett	0.5465	4	.	0.7016

Figure 28. Tests For Equal Variance

Additionally, the saved Cook's D influence statistic was graphically analyzed via a scatter plot in order to examine the ANOVA for the presence of any data entries that exert an inordinate amount of influence on the test results (Neter et al., 1996). The presence of such a statistic could skew analysis results. As stated earlier, a Cook's D influence statistic under 10-20% is taken to be acceptable (Neter et al., 1996). As illustrated below in Figure 29, all the Cook's D influence statistics fall within the acceptable range, with a peak value of 4.47%.

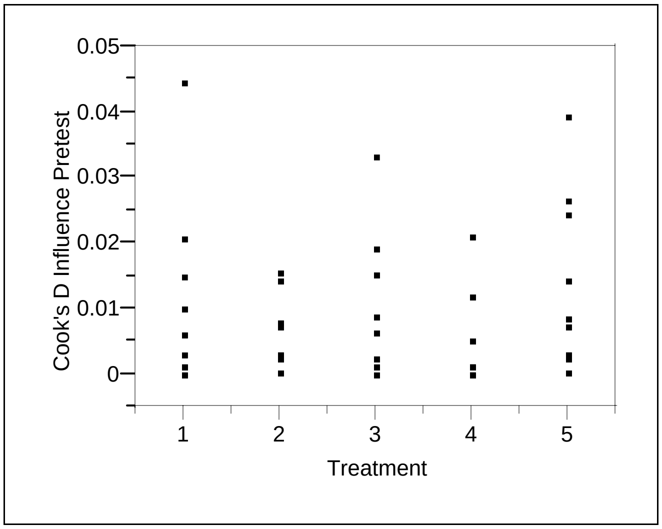

Figure 29. Cues Pretest ANOVA Cook's D Influence Factors

The initial ANOVA output, as shown in Figure 24, suggests that the means for all treatment groups are roughly identical. Additionally, the F-test p-value of 0.5312 leads to the conclusion, in statistical terms, of failing to reject the ANOVA's null hypothesis (which assumes that none of the variables have any predictive power toward forecasting pretest scores). This affords an initial indication that the starting condition of all five treatment groups was statistically equivalent. In order to further analyze this conclusion, a Tukey-Kramer HSD was conducted to accomplish a multiple comparison across all the treatment groups. The output is depicted in Figure 30 and visually illustrated in Figure 31 below.

Abs(Dif)-LSD

	4	1	5	2	3
4	-0.11972	-0.10789	-0.07878	-0.07448	-0.04876
1	-0.10789	-0.14487	-0.11709	-0.11286	-0.08585
5	-0.07878	-0.11709	-0.11690	-0.11259	-0.08702
2	-0.07448	-0.11286	-0.11259	-0.11556	-0.09007
3	-0.04876	-0.08585	-0.08702	-0.09007	-0.14235

Positive values show pairs of means that are significantly different.

Level		Mean
4	A	0.60000000
1	A	0.57500000
5	A	0.56046512
2	A	0.55681818
3	A	0.51724138

Levels not connected by same letter are significantly different

Figure 30. Pretest Tukey-Kramer HSD

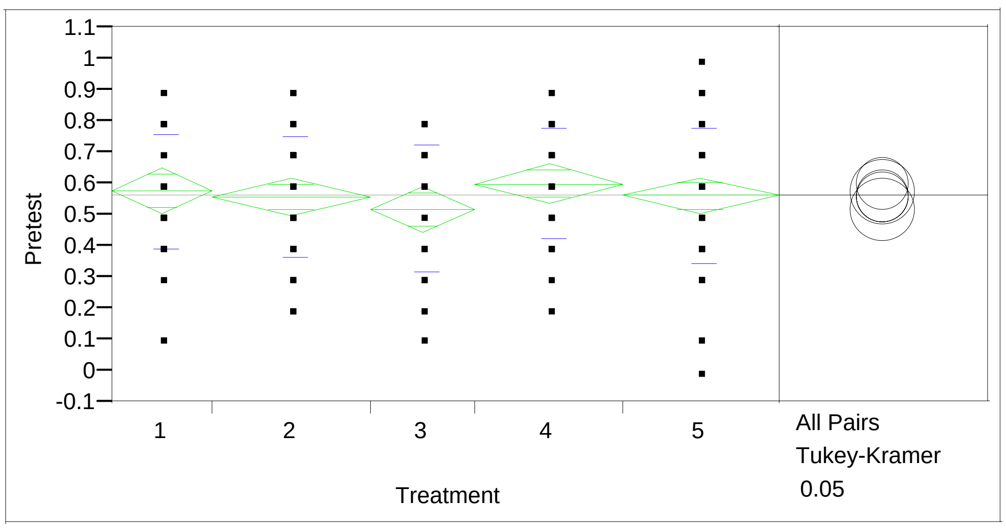

Figure 31. Pretest Tukey-Kramer HSD Graphical Illustration

The results of the Tukey-Kramer HSD confirm the initial indication that none of the treatment groups differed statistically. As shown in Figure 30, given an alpha of 0.05, none of the Tukey-Kramer HSD factors showed up as positive (positive numbers imply a significant difference between pairs). This statistical inference is visually illustrated in Figure 31, where it should be noted that all the diamonds are within the same general range, and all the circles overlap one another. Restated, no statistical difference in knowledge pretest outcomes across the five treatment groups is evident within the data.

Cues Knowledge Pretest/Posttest Differential Analysis

The differential between the cues pretest and posttest knowledge scores was analyzed to look for evidence that different treatments had varying effects on experimental participants. Since the preceding analysis for the cues knowledge pretest scores established baseline equivalence, finding statistically significant variations in the pretest to posttest score differential would suggest that treatment had an impact on the learning result. The same 185 pretest knowledge entries used in the knowledge pretest analysis were used, and each of the participant scores was translated into a composite percentile in the same translation manner already expounded upon in the previous section. The raw data results can be viewed in Appendix I. Additionally, the mean and standard deviation for each treatment groups posttest score and for the differential between its posttest and pretest score were calculated and are presented in Table 12 and Table 13 below. The score means range from 0.60 to 0.69, with treatment group 5 having the highest score. In order to determine statistical significance the same tests used in the previous sections were executed on this dataset.

Table 12. Cues Posttest Means and Standard Deviations

Treatment	Mean	Standard Deviation
Group 1	0.6036	0.1426
Group 2	0.6659	0.1584
Group 3	0.6793	0.1473
Group 4	0.6024	0.1651
Group 5	0.6967	0.1683

Table 13. Cues Score Differential Means and Standard Deviations

Treatment	Mean	Standard Deviation
Group 1	0.0286	0.1843
Group 2	0.1091	0.0269
Group 3	0.1621	0.2094
Group 4	0.0024	0.1810
Group 5	0.1372	0.2161

Following the initial look at the means, the score differentials were visually analyzed for the initial ANOVA statistical assumption of continuity via a histogram. The shape of the histogram suggests that the data is continuous enough to proceed with the ANOVA approach to analysis. The histogram is shown in Figure 32 below.

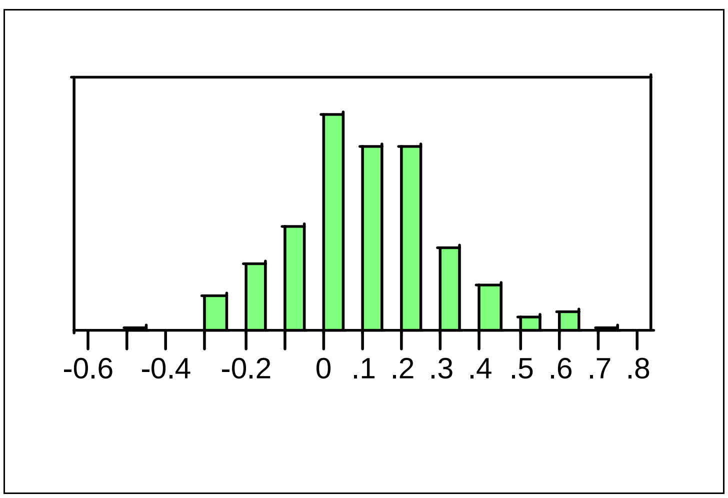

Figure 32. Histogram of Knowledge Pretest/Posttest Differentials

The differential data was then analyzed via the ANOVA statistical method. The initial ANOVA results are presented in Figure 33 below. The low ANOVA F-Statistic of 0.0017 gives an initial indication, in line with expectations, that at least one of the treatment sessions exerts some predictive influence on student learning In order to test for the statistical validity assumptions of independence of the residuals, normality of the residuals, and constant variance of the residuals, the Studentized Residuals and Cook's D influence factors were saved for further investigation.

Summary of Fit

Rsquare	0.091229
Adj Rsquare	0.071035
Root Mean Square Error	0.194251
Mean of Response	0.088108
Observations (or Sum Wgts)	185

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Treatment	4	0.6818343	0.170459	4.5175	0.0017
Error	180	6.7920036	0.037733		
C. Total	184	7.4738378			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
1	28	0.028571	0.03671	-0.0439	0.10101
2	44	0.109091	0.02928	0.0513	0.16688
3	29	0.162069	0.03607	0.0909	0.23325
4	41	0.002439	0.03034	-0.0574	0.06230
5	43	0.137209	0.02962	0.0788	0.19566

Std Error uses a pooled estimate of error variance

Figure 33. ANOVA of Cues Pretest/Posttest Differentials

In order to validate the assumption of normality on the residuals, the Studentized Residuals were plotted via a histogram, the output of which is illustrated in Figure 34 below. The plot of the distribution, as highlighted by the overlay of the normal curve, suggests that normality holds true. A Shapiro-Wilk goodness-of-fit test was then conducted to provide an objective analysis of the initial indications. The result of this test is illustrated in Figure 35 below. Hence, the statistical null hypothesis that the residuals follow the normal distribution is not rejected, since 0.3326 is not less than the alpha of 0.05. Again, the assumption of normality for the ANOVA test is satisfied.

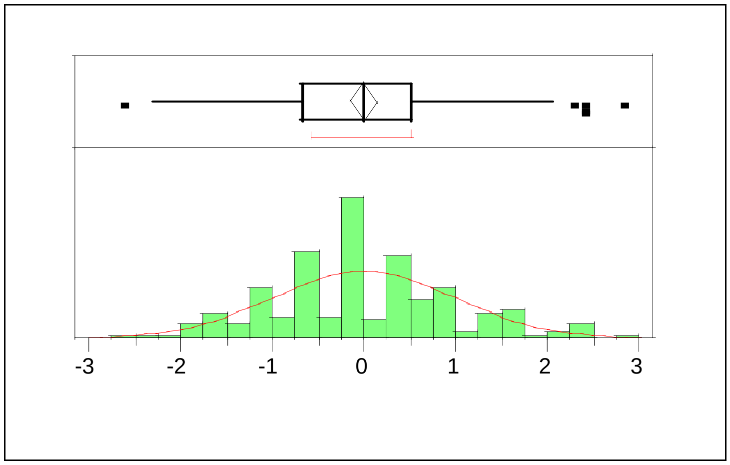

Figure 34. Histogram of Studentized Residuals

Goodness-of-Fit Test

Shapiro-Wilk W Test

W	Prob<W
0.979791	0.3326

Figure 35. Test for Normality

In order to validate the assumption of constant variance of the residuals, a scatter plot of the Studentized Residuals was accomplished to provide an initial indication of any trends inherent in the data. This plot is illustrated in Figure 36. Given that the points on the chart seem to fall within the same general range across all treatment groups, an initial indication is given that the assumption of constant variance holds true. Four separate tests (O'Brien test, Brown-Forsythe test, Levene's test, and Bartlett's test) were conducted to provide an objective statistical basis to further explore the result drawn from Figure 36. The output from these four tests is illustrated in Figure 37 below. Each of these tests confirms the initial indication that the residual variances are equal across the five treatment groups. None of the Prob > F statistical outcomes is under the assumed alpha of 0.05; hence, the null hypothesis of equal variance is not rejected. Also, the required assumption of constant variance for the ANOVA test is therefore satisfied.

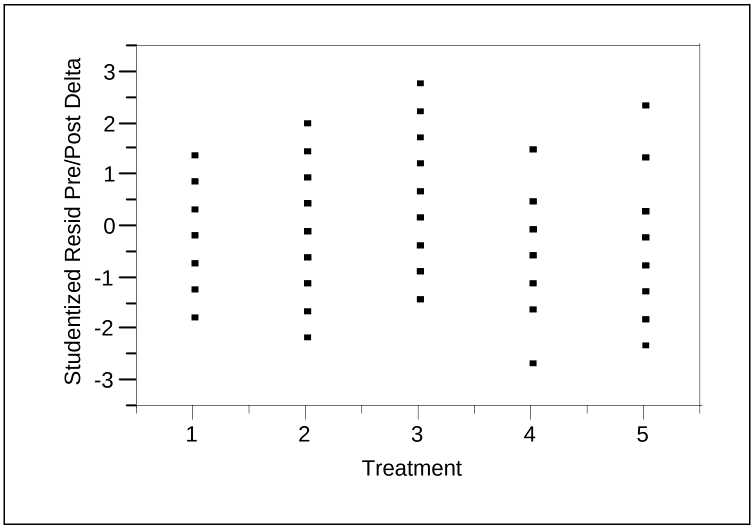

Figure 36. Plot of Residuals

Test	F Ratio	DFNum	DFDen	Prob > F
O'Brien[.5]	0.6269	4	180	0.6439
Brown-Forsythe	0.6562	4	180	0.6233
Levene	0.8099	4	180	0.5203
Bartlett	0.6018	4	.	0.6613

Figure 37. Tests For Equal Variance

Additionally, the saved Cook's D influence statistic was graphically analyzed via a scatter plot in order to examine the ANOVA for the presence of any data entries that exert an inordinate amount of influence on the test results (Neter et al., 1996). The presence of such a statistic could skew analysis results. As stated earlier, a Cook's D influence statistic under 10-20% is taken to be acceptable (Neter et al., 1996). As illustrated below in Figure 38, all the Cook's D influence statistics fall within the acceptable range, with a peak value of 5.67%.

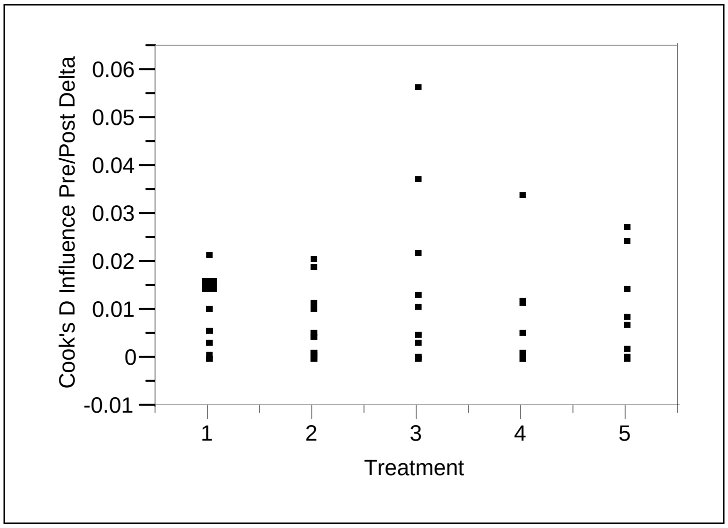

Figure 38. Cues Pretest/Posttest Differential ANOVA Cook's D Influence Factors

The initial ANOVA output, as shown in Figure 33, suggests that the means for all treatment groups are roughly identical. Additionally, the F-test p-value of 0.0017 leads to the conclusion, in statistical terms to reject the ANOVA's null hypothesis that none of the variables have any predictive power toward forecasting pretest scores. This affords an initial indication that least one of the five treatment groups was statistically different than the other four. In order to further analyze this conclusion, a Tukey-Kramer HSD was conducted to accomplish a multiple comparison across all the treatment groups. The output is depicted in Figure 39 and visually illustrated in Figure 40 below.

Abs(Dif)-LSD

	3	5	2	1	4
3	-0.14057	-0.10376	-0.07505	-0.00832	0.02975
5	-0.10376	-0.11544	-0.08667	-0.02135	0.01793
2	-0.07505	-0.08667	-0.11412	-0.04888	-0.00954
1	-0.00832	-0.02135	-0.04888	-0.14306	-0.10510
4	0.02975	0.01793	-0.00954	-0.10510	-0.11822

Positive values show pairs of means that are significantly different.

Level			Mean
3	A		0.16206897
5	A		0.13720930
2	A	B	0.10909091
1	A	B	0.02857143
4		B	0.00243902

Levels not connected by same letter are significantly different

Figure 39. Cues Pretest/Posttest Differential Tukey-Kramer HSD

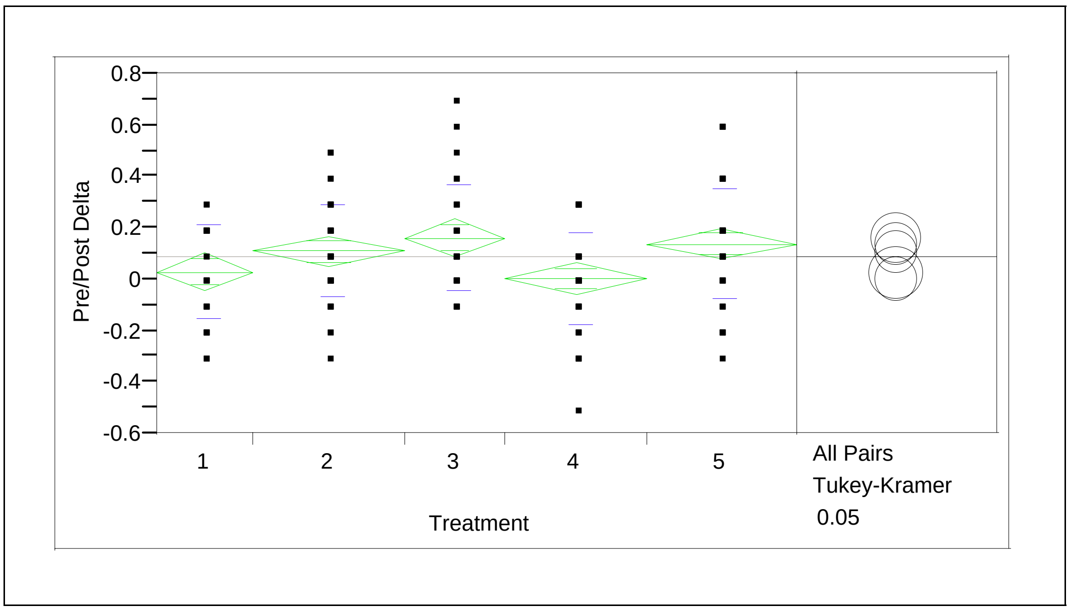

Figure 40. Cues Pretest/Posttest Differential Tukey-Kramer HSD Graphical Illustration

The results of the Tukey-Kramer HSD confirm the initial indication that there exists some statistically significant variation among score differential outcomes between the treatment groups. As shown in Figure 39, given an alpha of 0.05, the Tukey-Kramer HSD factors associated with the paired comparison of treatments 3 and 4 and the paired comparison of treatments 4 and 5 showed up as positive; these positive numbers imply a significant difference between pairs. This statistical inference is visually illustrated in Figure 40, where it should be noted that the diamond and circle associated with treatment 4 does not overlap the diamonds or circles associated with treatments 3 and 5. Restated, the only statistical difference in cues knowledge pretest/posttest differentials across the five treatment groups is evident between treatment 4 and both treatments 3 and 5.

Analysis Outcomes in Terms of Experimental Hypotheses

Following the same analysis methodology as in the introductory session analysis, each of the hypotheses was rationally linked as a comparative relationship between two treatment groups—the first of which formed the experimental baseline from which the second treatment group was expected to improve upon, because of the addition of factors such as increased media richness, content flexibility, and forced engagement. As with the introductory session, hypothesis five must be excluded from the analysis for the afore-mentioned reasons. As before, the Tukey-Kramer HSD statistical comparison of the means highlights significantly different means by assigning them with different significance levels; these, along with the treatment group comparisons and hypotheses results, are summarized in Table 14 below. As shown, only hypothesis three is supported by the data. This suggests that the introduction of forced engagement had an effect on student learning. The implications are discussed in greater detail in Chapter 5.

Table 14. Hypotheses Outcomes on Cues Session Analysis

Hypothesis	Treatment Groups		Tukey HSD Mean Level		Result
H1	1	2	A,B	A,B	Unsupported
H2	2	3	A,B	A	Unsupported
H3	4	5	B	A	**Supported**
H4	1	3	A,B	A	Unsupported
H6	2	5	A,B	A	Unsupported
H7	1	5	A,B	A	Unsupported

Summary of Statistical Analysis Results

The pretreatment condition of all groups was found to be statistically equivalent for all treatment groups for both the introductory and cues treatment sessions. This assertion can be made because none of the Tukey-Kramer HSD scores for each pretest showed any statistical difference at an alpha of 0.05. Within the introductory treatment, the data failed to support any evidence that any of the treatments had a different impact on the pretest/posttest score differential, and none of the seven hypotheses were supported (the Tukey-Kramer HSD scores failed to showed any statistical difference at an alpha of 0.05). However, within the cues treatment, a statistically significant difference was found to be evident between treatments 3 and 4, and treatments 5 and 4 at an alpha of 0.05. This translated into statistical support for hypothesis three. The implications and limitations of these results will be thoroughly discussed in the following chapter.

V. Conclusion, Recommendations, and Lessons Learned

Summary of Findings

The underlying goal of this experimental research was to investigate how the application of various changes to the presentation of information within an electronic medium affected participant learning. The variations applied to the medium, and the resulting hypotheses that were developed, were based on learning theory, learning styles, and other past research findings. The results of the rigorous in-depth analysis of the research data, within the framework of the underlying research hypotheses, are summarized in Table 15 below. The aim of this chapter is to rationally discuss the limitations of the research results and the implications thereof.

Table 15. Summary of Findings

Hypothesis	Result (Intro)	Result (Cues)
H1: Having more media richness within the electronic learning medium will result in significantly higher student learning differentials than media with less media richness.	Not Supported	Not Supported
H2: An electronic medium with more content flexibility will result in significantly higher student learning differentials than media with less content flexibility.	Not Supported	Not Supported
H3: The presence of forced engagement of the student with the material will result in significantly higher student learning differentials than media with no forced engagement.	Not Supported	**Supported**
H4: The sum total of media richness and content flexibility will cumulatively result in significantly higher student learning differentials than that in media with lower levels of media richness and less content flexibility.	Not Supported	Not Supported
H6: The sum total of content flexibility and forced engagement will cumulatively result in significantly higher student learning differentials than that in media with less content flexibility and no forced engagement.	Not Supported	Not Supported
H7: The sum total of media richness, content flexibility, and forced engagement will cumulatively result in significantly higher student learning differentials than that in media with lower levels of media richness, less content flexibility, and no forced engagement.	Not Supported	Not Supported

As shown in the table, the only hypothesis that was supported was hypothesis three, that the presence of pop up quizzes that force the engagement of the student with the material will result in significantly higher student learning differentials. It should be noted, as is evident in Chapter 3, that the support for this hypothesis was only partial, as it was only supported in the cues session of the experiment and not in the introductory session. Additional limitations of this finding are discussed below.

What the Results Mean

The lack of variation in learning outcomes across the five treatment groups seems to suggest that the influence of the three primary research constructs of media richness, content flexibility, and forced engagement have less impact on learning than previously thought. The implication may be that it is just as effective to use a simple taped video lecture and present it across an electronic medium as it is to present a more elaborate and intricate medium, which may require more resources in the form of time and money to create. In reference to the Air Force, this may mean that a simple instructional video on safety, harassment, and security training may be enough—and the time and money spent on creating more elaborate training media is unnecessary and unfruitful. However, the research does have some limitations and more research is called for to confirm this basis across other subject areas and subject groups, as is discussed in more detail in the following section.

Limitations to Research Findings

Considering the experimental results summarized above, several limitations of the experimental effort should be considered. Firstly, while from a pure statistical framework there is strong support for concluding at least partial support for hypothesis three, the actual causal effect is questionable. It is questionable whether or not the statistical support for the hypothesis stems from an actual benefit derived from the pop-up quizzes or from an unusually low baseline improvement of treatment four—against which treatment five with the pop-up quizzes was compared. An impartial viewing of the data, as available in Appendices H and I, seems to imply that the latter may be true. This is also suggested by the low score improvement of treatment group 4, as already illustrated in Table 13, in Chapter four. As shown, treatment group 4 almost had no score improvement, with a mean score change of only 0.0024—well below that of the other groups. This is visually illustrated in Figure 41 below.

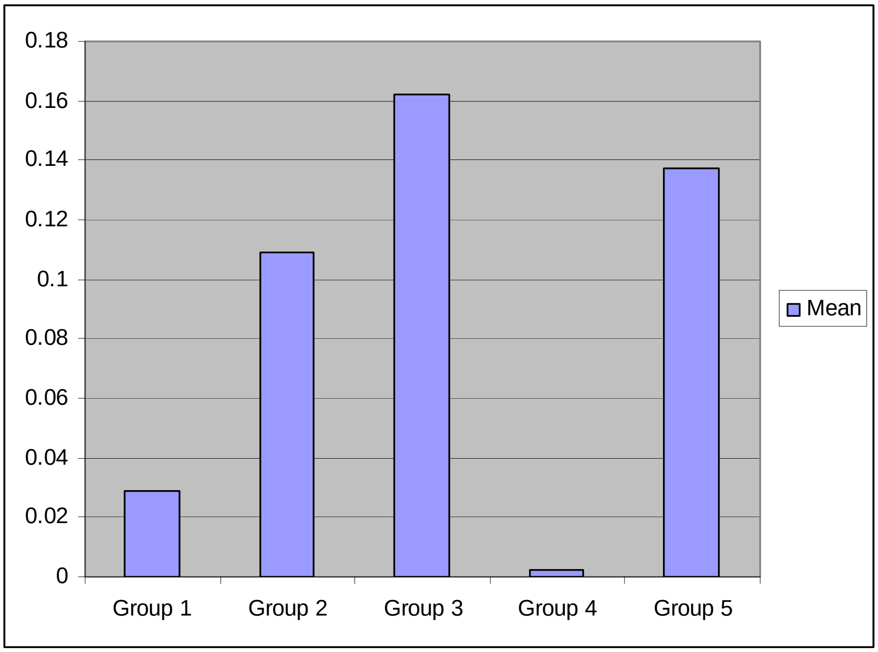

Figure 41. Pretest/Postest Differential Mean by Treatment Group

This suspicion fails to be attenuated by rationally considering and bringing to bear the lack of statistical support for hypotheses six and seven, which also involved the construct of forced participant engagement. Hence, it becomes feasible to suspect that the statistical support for hypothesis three may actually have stemmed from a low score in the control baseline rather than from an increase that stemmed from forced engagement.

An additional limitation, which may in part be to blame, is the short length of the pretest and posttests used in the study. Both the pretest and the posttest were only ten questions in length. Hence, just a deviation of one additional correct or incorrect response from a participant had a 10-point effect on that participant's measured learning outcome. In future studies, a more precise outcome may be achieved by increasing the number of test questions on the pre-experimental and post-experimental observations.

Limitations to Implications: May be Topic Specific

A limitation to the application of the implications of this study should also be considered. It can be conceptualized that a particular electronic learning medium, while being effective in teaching a particular topic, may not produce the same learning result given a different topic. In this experimental study, the instruction of the topic of detection of deceptive communication was the focus. However, there are many other topics of training interest—Figure 42 illustrates the conceptualization that the hypotheses findings of this study may not hold true for other instructional topics of interest. Depicted in this figure is the rationalization that an individual who uses a particular electronic learning medium to engage in study of a number of arbitrary topics may find that the particular medium in use is not equally conducive to all topics, and therefore yields not a comparative result but rather a stratification of learning outcomes.

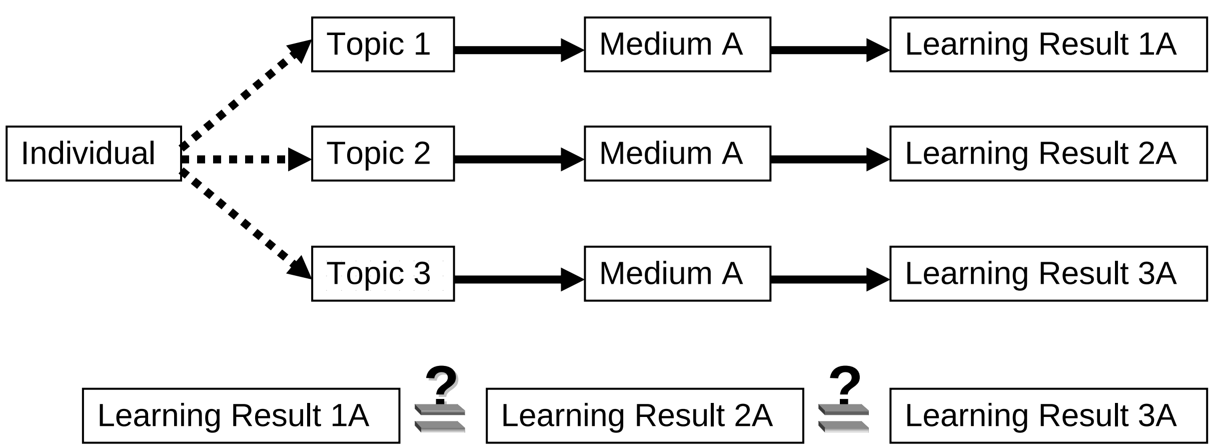

Figure 42. Three Topics and One Medium

While the single topic focus (on the detection of deceptive communication) of this study was insightful, testing these learning theories in a similar fashion against other topic areas of interest is an area in which more study is recommended.

Implications to Practice

The importance of proper training and education of military personnel will continue to be of high importance to the United States Air Force. Electronic learning media provide an attractive way in which training can be conducted following an "anywhere" and "anytime" methodological approach. As the use of such electronic training media continues to develop, the importance of focusing on teaching productivity and rate of return on time invested by military personnel in these training forums becomes increasingly important. This experimental research was based on sound learning theory and modes of learning constructs to explore the feasibility of tailoring electronic instruction in such a way that knowledge transfer and productivity while engaged in electronic learning are maximized.

Although within this experimental research no venues for improving electronic research productivity were discovered, given the general lack of support for the research hypotheses of this study it can be suggested that given the current empirical evidence the use of a video lecture as an electronic training medium can be just as effective as a more complicated training medium which takes more time and money to construct—at least within the realm of detecting deceptive communication. This lends itself to potential savings in cost and time in the development of future training regiments.

Conclusion

The results from this experimental research suggest that the variations in media richness, content flexibility, and forced engagement had a negligible effect on participant learning in the area of detecting deception detection. While partial support for the use of pop up quizzes was found, the interpretative relevance of its impact is questionable because the statistical relevance seems to originate from a low baseline score, rather than from a large learning increase of the treatment group. However, the lack of support for the research hypotheses suggests that at least within the realm of deception detection a pure video lecture can be just as effective as a more complicated media-rich electronic medium at facilitating knowledge transfer and learning.

Appendix A. Definition of Terms

AFOSR	Air Force Office of Scientific Research
HSD	Honestly Significant Difference; part of the Tukey-Kramer test

Appendix B. AFRL Approval Letter

DEPARTMENT OF THE AIR FORCE
AIR FORCE RESEARCH LABORATORY (AFMC)
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

2
September 2003

MEMORANDUM FOR AFIT/ENV
ATTN: Jachin Sakamoto

FROM: AFRL/HEH

SUBJECT: Approval for the Use of Volunteers in Research

1. Human experimentation as described in Protocol 03-80, "Experiment on Abstract Learning Stratification & Knowledge Transfer in an Electronically Mediated environment Driven by Leadership", may begin.

2. In accordance with AFI 40-402, this protocol was reviewed and approved by the Wright Site Institutional Review Board (WSIRB) on 25 August 2003, the AFRL Chief of Aerospace Medicine on 27 August 2003

3. Please notify the undersigned of any changes in procedures prior to their implementation. A judgment will be made at that time whether or not a complete WSIRB review is necessary.

Signed 2 September 2003
HELEN JENNINGS
Human Use Administrator

Appendix C. Introduction Pretest

1. Typically, people successfully detect deception about ______ of the time.
    a) 20%
    b) 50%
    c) 80%
    d) 90%

2. An example of adaptor behavior would be:
    a) blinking
    b) shaking one's fist
    c) increased voice pitch
    d) scratching one's face

3. All of the following are properties of language that might serve as deception cues EXCEPT:
    a) the use of pronouns
    b) submissive language
    c) number of words
    d) nasality of the voice

4. A truthful message is more likely to contain:
    a) larger words
    b) smaller words
    c) simple sentence structure
    d) lack of emotion

5. Which of the following would NOT directly lead to better detection accuracy?
    a) familiarity with the communicative context
    b) experience with the communicative medium
    c) familiarity with the topic of conversation
    d) experience in high-risk situations

6. Applying lie detection skills and staying focused for long periods of time is known as:
    a) leakage
    b) arousal
    c) vigilance
    d) truth bias

7. Deception is
    a) the same thing as lying
    b) the same thing as social engineering
    c) misleading others through means other than communication
    d) any intentional form of communication or behavior used to mislead others

8. Most people assume that
    a) messages from other people are truthful
    b) other people lie most of the time
    c) people only lie about really important things
    d) people only lie about things that are in their own self-interest

9. Which of the following is NOT a reliable visual indicator of deception?
    a) fake smiles
    b) amount of eye contact
    c) amount of gesturing
    d) self-touching

10. Deception is unlikely to be present when people are communicating by:
    a) cell phone
    b) e-mail
    c) videoconferencing
    d) None. Deception occurs in any mode.

Appendix D. Introduction Posttest

1. A simple way to define *deception* is:
   a. a message that is inaccurate in its content and assumptions
   b. a message that is purposely used to foster a false conclusion by others
   c. a message that contradicts the beliefs of the majority of society
   d. a message that blatantly breaks the norms of a society's culture

2. The tendency for most human beings to believe other people are honest by default is known as the __________.
   a. trust bias
   b. truth bias
   c. lie bias
   d. gullibility bias

3. The type of deception in which a person uses ambiguous language to answer a question is
   a. equivocation
   b. fabrication
   c. evasion
   d. concealment

4. Interpersonal Deception Theory views deception as being a __________ process.
   a. fixed, unchanging
   b. dynamic, changing
   c. psychological
   d. mostly unpredictable

5. In terms of detecting deception, the downside of being suspicious is that it might lead to:
   a. more truth bias
   b. excessive cognitive processing
   c. more false alarms
   d. less accuracy detecting deception

6. The concept that deceivers are not able to control indicators pointing to their dishonesty is the idea behind:
   a. **leakage theory**
   b. interpersonal deception theory
   c. truth bias
   d. immediacy theory

7. All of the following are considered ways to strategically alter a message EXCEPT
   a. reduce the quantity of information
   b. change the truthfulness of the information
   c. make the language less clear
   d. use the same language repeatedly

8. Lies contain more:
   a. emotional language
   b. definite details
   c. imagery
   d. simple words

9. What would be a reliable vocal indicator of deception?
   a. slowed rate of speech
   b. relaxed laughter
   c. few pauses in speech
   d. lower voice pitch

10. Training works best when it includes all of the following EXCEPT
   a. specific examples
   b. immediate feedback on judgment success
   c. practice
   d. single exposures

Appendix E. Cues Lecture Pretest

1. The theory that suggests deceivers will be unable to control all of their behavior while lying is:
    a. interpersonal deception theory
    b. indicator theory
    c. cognitive effort theory
    d. leakage theory

2. Lying to others is less difficult mentally when deceivers are able to:
    a. perceive the consequences of being caught as severe
    b. rush through their presentation
    c. rehearse their deceptive message in their mind or out loud
    d. spontaneously produce the deceptive message

3. Deceivers are apt to display **_______** -based cues if the consequences of having a lie detected are perceived to be severe.
    a. arousal
    b. memory
    c. cognitive
    d. haptic

4. If asked "Have you seen Joe's missing wallet?", a deceiver using the delay tactic of tag questions would respond with:
    a. "What are you implying?"
    b. "That's too bad for Joe, isn't it?"
    c. "Who are you to ask me such a question?"
    d. "Why should I have seen it? Of course not."

5. Deceivers tend to use**________** in their messages than truthtellers.
    a. fewer personal pronouns
    b. faster speaking tempo
    c. more detailed explanations
    d. more formal names and places

6. The use of qualifying terms like "maybe, perhaps, could have":
    a. are more likely under deception
    b. are more likely under truth
    c. are equally likely under both truth and deception
    d. are poor cues to distinguish truth from deception

7. Which of the following is more likely among deceivers than truth tellers at the beginnings of conversations?
    a. vocal pleasantness
    b. active gesturing
    c. monotone speaking
    d. unusual details

8. When relating a past event, an honest communicator is less likely to:
    a. report on his or her emotional state at the time of the event
    b. report on unusual details about the event
    c. report on the verbatim discussion of those at the event
    d. leave out the names of people at the event

9. Truthful messages are more likely to contain:
    a. fewer personal pronouns (I, we, etc.)
    b. definitive words like "definitely" and "absolutely"
    c. use of vague verbs like "could" and "would"
    d. qualifiers like "perhaps" and "possibly"

10. All of the following are major classes of behavioral deception indicators EXCEPT
    a. emotion
    b. memory
    c) cognitive effort
    d) physiognomy

Appendix F. Cues Lecture Posttest

1. What would be a reliable kinesic indicator of deception?
    a. the speaker is leaning forward
    b. the speaker has stiff, wooden posture
    c. a highly expressive face
    d. relaxed posture

2. Under what conditions are deceivers likely to produce longer messages than truthtellers?
    a. when they have time to plan, rehearse or edit their communication
    b. when they are trying to be persuasive
    c. when the communication medium has time delays such as with email
    d. all of the above

3. A physiological indicator that might tip off a deceiver to a polygraph (lie detection machine) would be:
    a. decreased blinking
    b. increased pulse rate
    c. negative speech
    d. increased stuttering

4. The theory that suggests that deceivers strategically and intentionally alter their messages to avoid detection is:
    a. **interpersonal deception theory**
    b. indicator theory
    c. leakage theory
    d. cognitive effort theory

5. Which of the following would NOT be a reliable cue pointing toward deception?
    a. lower voice pitch
    b. poor detail in a particular message
    c. non-ah nonfluencies
    d. fewer pauses

6. It is possible that deceivers are having a difficult time lying if we notice that they:
    a. respond immediately after being asked a question
    b. behave in a normal manner
    c. drop the names of others into conversation
    d. stop gesturing

7. A ________ may be used by deceivers to mask their negativity toward others.
    a. non-ah nonfluency
    b. self-grooming behavior
    c. feigned smile
    d. long response latency

8. The increased difficulty associated with lying while conveying a consistent story to others is known as _____________.
    a. cognitive effort
    b. leakage theory
    c. arousal
    d. deceptive stress and strain

9. Truthful messages are more likely to contain:
    a. fewer personal pronouns (I, we, etc.)
    b. use of vague verbs like "could" and "would"
    c. qualifiers like "perhaps" and "possibly"
    d. definitive words like "definitely" and "absolutely"

10. If writing or e-mailing a truthful message to others, an honest person is likely to use:
    a. more punctuation
    b. simple sentence structure
    c. limited vocabulary
    d. more misspelled words

Appendix G. Pre- and Post-Judgment Tests

Test / Question	Answer	Source	Length		Variable Name
POST1 1	Deception	POST1 D MT 825 Q15-16	12;12	V1 A1-2	IntrJO1
POST1 2	Truth	POST1 T MT 866 Q17	34;18	V1 A1-2	IntrJO2
POST1 3	Deception	POST1 D FH 45 Q5	30;00	V1	IntrJO3
POST1 4	Truth	POST1 T EXP5 29 Q10	22:09	V1 A1-2	IntrJO4
POST1 5	Truth	POST1 T FH 54 QY	30;00	V1	IntrJO5
POST1 6	Truth	POST1 T MT 801 Q6	19;28	V1 A1-2	IntrJO6
POST1 7	Truth	POST1 T EXP5 13 Q2 AUD	50;28	V1 A1-2	IntrJO7
POST1 8	Deception	POST1 D EXP5 53 Q9	44;27	V1 A1-2	IntrJO8
POST1 9	Deception	POST1 D MT 859 Q17 AUD	12;01	V1 A1-2	IntrJO9
POST1 10	Truth	POST1 T EXP5 23 Q6	45;12	V1 A1-2	IntrJO10
POST2 1	Truth	POST2 T EXP5 48 Q11	54;14	V1 A1-2	CueJO1
POST2 2	Truth	POST2 D MT 805 Q6-7			CueJO2
POST2 3	Deception	POST2 D MT 887 Q19	16;24	V1 A1-2	CueJO3
POST2 4	Truth	POST2 T MT 864 Q17	38;07	V1 A1-2	CueJO4
POST2 5	Deception	POST2 D EXP5 48 Q3 AUD	35;24	V1 A1-2	CueJO5
POST2 6	Truth	POST2 T EXP5 55 Q4	38;06	V1 A1-2	CueJO6
POST2 7	Deception	POST2 D FH 45 QX	30;00	V1	CueJO7
POST2 8	Truth	POST2 T MT 804 Q20	21;11	V1 A1-2	CueJO8
POST2 9	Truth	POST2 T EXP5 35 Q6 AUD	1:02:13	V1 A1-2	CueJO9
POST2 10	Deception	POST2 D MT 861 Q15-16	16;19	V1 A1-2	CueJO10
PRE1 1	Deception	PRE1 D EXP5 22 Q3	34;26	V1 A1-2	IntrJE1
PRE1 2	Deception	PRE1 D MT 881 Q20	14;07	V1 A1-2	IntrJE2
PRE1 3	Truth	PRE1 T MT 810 Q17	50;26	V1 A1-2	IntrJE3
PRE1 4	Truth	PRE1 T EXP5 25 Q2	38;06	V1 A1-2	IntrJE4
PRE1 5	Deception	PRE1 D FH 54 Q8	30;00	V1	IntrJE5
PRE1 6	Truth	PRE1 T MT 864 Q5-6 AUD	27;19	V1 A1-2	IntrJE6

PRE1 7	Deception	PRE1 D EXP5 36 Q2	55;23	V1 A1-2	IntrJE7
PRE1 8	Deception	PRE1 D MT 888 Q12	33;29	V1 A1-2	IntrJE8
PRE1 9	Truth	PRE1 T EXP5 55 Q11	52;02	V1 A1-2	IntrJE9
PRE1 10	Truth	PRE1 T MT 860 Q20	25;12	V1 A1-2	IntrJE10
PRE2 1	Truth	PRE2 T MT 859 Q3-6	36;21	V1 A1-2	CueJE1
PRE2 2	Truth	PRE2 T MT 816 Q16	17;13	V1 A1-2	CueJE2
PRE2 3	Deception	PRE2 D EXP5 21 Q12 AUD	1:28:20	V1 A1-2	CueJE3
PRE2 4	Deception	PRE2 D MT 821 Q19	32;04	V1 A1-2	CueJE4
PRE2 5	Truth	PRE2 T FH 51 QZ	30;00	V1	CueJE5
PRE2 6	Deception	PRE2 D MT 828 Q19	31;08	V1 A1-2	CueJE6
PRE2 7	Deception	PRE2 D EXP5 24 Q3 AUD	33;17	V1 A1-2	CueJE7
PRE2 8	Deception	PRE2 D MT 819 Q12	30;18	V1 A1-2	CueJE8
PRE2 9	Truth	PRE2 T MT 824 Q20 AUD	23;05	V1 A1-2	CueJE9
PRE2 10	Deception	PRE2 D MT 887 Q12-13	30;15	V1 A1-2	CueJE10

Appendix H. Introductory Session Translated Raw Data

id	Treatment	Score PreTest	Score Posttest	Score Delta
1000	1	0.5	0.9	0.4
1001	1	0.6	0.7	0.1
1002	1	0.5	0.6	0.1
1003	1	0.5	0.6	0.1
1004	1	0.7	0.6	-0.1
1005	1	0.6	0.9	0.3
1006	1	0.6	0.8	0.2
1007	1	0.5	0.8	0.3
1008	1	0.5	0.9	0.4
1009	1	0.8	0.7	-0.1
1010	1	0.7	0.7	0
1011	1	0.3	0.5	0.2
1012	1	0.6	0.9	0.3
1013	1	0.5	0.9	0.4
1015	1	0.3	0.9	0.6
1017	1	0.8	0.9	0.1
1018	1	0.5	0.9	0.4
1019	1	0.6	0.4	-0.2
1021	1	0.5	0.5	0
1025	1	0.5	0.8	0.3
1032	1	0.6	0.5	-0.1
1033	1	0.5	0.7	0.2
1034	1	0.3	0.8	0.5
1035	1	0.5	0.9	0.4
1036	1	0.5	1	0.5
1037	1	0.5	0.7	0.2
1038	1	0.6	0.8	0.2
1039	1	0.6	0.9	0.3
1040	1	0.4	0.7	0.3
1041	1	0.6	0.9	0.3
1080	1	0.7	1	0.3
1016	2	0.8	0.7	-0.1
1022	2	0.4	0.9	0.5
1023	2	0.4	0.8	0.4
1024	2	0.6	0.8	0.2
1026	2	0.5	0.8	0.3
1027	2	0.5	0.6	0.1
2000	2	0.4	0.7	0.3
2001	2	0.5	0.9	0.4
2002	2	0.4	0.6	0.2
2003	2	0.7	0.6	-0.1
2005	2	0.5	0.8	0.3

2006	2	0.1	0.4	0.3
2007	2	0.4	0.9	0.5
2008	2	0.6	0.6	0
2009	2	0.6	0.4	-0.2
2010	2	0.4	0.7	0.3
2011	2	0.6	0.9	0.3
2012	2	0.6	0.7	0.1
2014	2	0.6	0.7	0.1
2015	2	0.6	0.9	0.3
2016	2	0.4	0.8	0.4
2017	2	0.5	0.8	0.3
2018	2	0.6	1	0.4
2019	2	0.5	0.8	0.3
2020	2	0.6	0.9	0.3
2021	2	0.4	0.7	0.3
2022	2	0.4	0.9	0.5
2023	2	0.4	0.8	0.4
2024	2	0.4	1	0.6
2025	2	0.5	0.8	0.3
2026	2	0.2	0.6	0.4
2027	2	0.5	1	0.5
2028	2	0.5	0.8	0.3
2029	2	0.3	0.4	0.1
2030	2	0.5	0.8	0.3
2031	2	0.8	0.7	-0.1
2032	2	0.5	0.7	0.2
2033	2	0.5	0.7	0.2
2034	2	0.7	0.8	0.1
2035	2	0.4	0.7	0.3
2036	2	0.6	0.7	0.1
2037	2	0.6	0.9	0.3
2038	2	0.3	0.3	0
2039	2	0.6	0.8	0.2
3000	3	0.4	0.7	0.3
3001	3	0.5	1	0.5
3002	3	0.6	0.7	0.1
3003	3	0.7	0.8	0.1
3004	3	0.5	0.9	0.4
3005	3	0.5	0.7	0.2
3006	3	0.4	0.8	0.4
3007	3	0.5	0.9	0.4
3008	3	0.5	0.7	0.2
3009	3	0.6	0.8	0.2
3010	3	0.4	0.9	0.5
3011	3	0.6	0.8	0.2
3012	3	0.6	0.6	0

3013	3	0.6	0.7	0.1
3014	3	0.6	0.8	0.2
3015	3	0.6	1	0.4
3016	3	0.5	0.8	0.3
3017	3	0.8	0.8	0
3018	3	0.5	0.6	0.1
3019	3	0.6	0.9	0.3
3020	3	0.6	0.8	0.2
3021	3	0.5	0.7	0.2
3022	3	0.1	0.7	0.6
3023	3	0.5	1	0.5
3024	3	0.5	0.8	0.3
3025	3	0.5	0.8	0.3
3026	3	0.4	0.3	-0.1
3033	3	0.7	0.6	-0.1
3040	3	0.6	0.8	0.2
3042	3	0.6	0.8	0.2
4000	4	0.4	0.7	0.3
4001	4	0.6	0.8	0.2
4002	4	0.6	0.6	0
4003	4	0.5	0.8	0.3
4004	4	0.4	0.6	0.2
4005	4	0.6	0.7	0.1
4006	4	0.7	0.9	0.2
4007	4	0.6	0.8	0.2
4008	4	0.5	0.7	0.2
4009	4	0.3	0.9	0.6
4010	4	0.7	0.8	0.1
4011	4	0.5	0.8	0.3
4013	4	0.7	0.7	0
4014	4	0.1	0.9	0.8
4015	4	0.4	0.8	0.4
4016	4	0.7	1	0.3
4017	4	0.3	0.4	0.1
4018	4	0.5	0.8	0.3
4019	4	0.5	0.8	0.3
4020	4	0.5	0.5	0
4021	4	0.4	0.2	-0.2
4022	4	0.4	0.8	0.4
4023	4	0.6	0.8	0.2
4024	4	0.4	0.8	0.4
4025	4	0.6	0.5	-0.1
4026	4	0.4	0.6	0.2
4027	4	0.9	0.9	0
4028	4	0.4	0.7	0.3
4029	4	0.4	0.7	0.3

4032	4	0.3	0.8	0.5
4033	4	0.5	0.7	0.2
4034	4	0.6	0.8	0.2
4035	4	0.4	0.8	0.4
4036	4	0.3	0.6	0.3
4037	4	0.6	0.9	0.3
4038	4	0.6	0.5	-0.1
4039	4	0.8	0.9	0.1
4040	4	0.5	1	0.5
4041	4	0.7	0.6	-0.1
4042	4	0.5	0.6	0.1
4043	4	0.5	0.8	0.3
4044	4	0.5	0.6	0.1
4045	4	0.5	0.6	0.1
4046	4	0.6	0.7	0.1
5000	5	0.5	0.9	0.4
5001	5	0.7	0.9	0.2
5002	5	0.4	0.5	0.1
5003	5	0.5	0.5	0
5004	5	0.4	0.8	0.4
5008	5	0.6	0.7	0.1
5009	5	0.4	0.8	0.4
5010	5	0.6	0.9	0.3
5011	5	0.4	1	0.6
5012	5	0.3	0.8	0.5
5013	5	0.8	1	0.2
5014	5	0.4	0.9	0.5
5015	5	0.4	0.7	0.3
5016	5	0.7	0.7	0
5017	5	0.6	0.9	0.3
5018	5	0.5	0	-0.5
5019	5	0.6	0.8	0.2
5020	5	0.5	0.9	0.4
5021	5	0.5	0.7	0.2
5022	5	0.5	0.9	0.4
5023	5	0.5	1	0.5
5024	5	0.6	0.9	0.3
5025	5	0.7	0.9	0.2
5026	5	0.4	0.8	0.4
5027	5	0.5	0.6	0.1
5028	5	0.5	0.8	0.3
5029	5	0.3	0.8	0.5
5030	5	0.7	0.9	0.2
5031	5	0.3	0.9	0.6
5032	5	0.5	0.8	0.3
5033	5	0.7	0.8	0.1

5034	5	0.6	0.9	0.3
5035	5	0.4	0.9	0.5
5036	5	0.5	0.8	0.3
5037	5	0.5	0.8	0.3
5038	5	0.6	1	0.4
5039	5	0.5	0.9	0.4
5040	5	0.6	0.8	0.2
5041	5	0.6	0.9	0.3
5042	5	0.6	0.9	0.3

Appendix I. Cues Session Translated Raw Data

id	Treatment	Score Pretest	Score Posttest	Score Delta
1000	1	0.9	0.6	-0.3
1001	1	0.5	0.4	-0.1
1002	1	0.8	0.6	-0.2
1003	1	0.1	0.4	0.3
1004	1	0.6	0.8	0.2
1005	1	0.6	0.6	0
1006	1	0.8	0.8	0
1007	1	0.4	0.6	0.2
1008	1	0.4	0.6	0.2
1010	1	0.8	0.6	-0.2
1012	1	0.3	0.5	0.2
1013	1	0.5	0.7	0.2
1014	1	0.6	0.8	0.2
1015	1	0.5	0.7	0.2
1017	1	0.6	0.6	0
1018	1	0.5	0.8	0.3
1019	1	0.4	0.6	0.2
1020	1	0.9	0.6	-0.3
1032	1	0.6	0.4	-0.2
1033	1	0.5	0.6	0.1
1034	1	0.5	0.4	-0.1
1035	1	0.5	0.5	0
1036	1	0.7	0.6	-0.1
1037	1	0.6	0.6	0
1038	1	0.8	0.9	0.1
1039	1	0.5	0.7	0.2
1040	1	0.5	0.3	-0.2
1041	1	0.7	0.6	-0.1
1016	2	0.6	0.5	-0.1
1021	2	0.8	0.5	-0.3
1022	2	0.3	0.8	0.5
1023	2	0.4	0.5	0.1
1024	2	0.4	0.6	0.2
1027	2	0.4	0.8	0.4
2000	2	0.3	0.3	0
2001	2	0.6	0.7	0.1
2002	2	0.7	0.7	0
2003	2	0.4	0.4	0
2004	2	0.8	0.7	-0.1
2005	2	0.6	0.7	0.1
2006	2	0.2	0.3	0.1

2007	2	0.5	0.8	0.3
2009	2	0.6	0.8	0.2
2010	2	0.4	0.6	0.2
2011	2	0.2	0.7	0.5
2012	2	0.5	0.8	0.3
2014	2	0.7	0.9	0.2
2015	2	0.9	0.9	0
2016	2	0.7	0.6	-0.1
2017	2	0.6	0.7	0.1
2018	2	0.4	0.7	0.3
2019	2	0.7	0.7	0
2020	2	0.5	0.6	0.1
2021	2	0.8	0.6	-0.2
2022	2	0.6	0.7	0.1
2023	2	0.7	0.9	0.2
2024	2	0.9	0.7	-0.2
2025	2	0.5	0.7	0.2
2026	2	0.4	0.7	0.3
2027	2	0.9	0.9	0
2028	2	0.6	0.7	0.1
2029	2	0.5	0.7	0.2
2030	2	0.7	0.7	0
2031	2	0.5	0.4	-0.1
2032	2	0.4	0.7	0.3
2033	2	0.7	0.7	0
2034	2	0.5	0.7	0.2
2035	2	0.5	0.5	0
2036	2	0.9	0.9	0
2037	2	0.3	0.7	0.4
2038	2	0.2	0.3	0.1
2039	2	0.7	0.8	0.1
3000	3	0.1	0.8	0.7
3001	3	0.7	0.7	0
3002	3	0.8	0.7	-0.1
3003	3	0.7	0.8	0.1
3004	3	0.8	0.7	-0.1
3005	3	0.3	0.6	0.3
3006	3	0.7	0.9	0.2
3007	3	0.7	0.7	0
3008	3	0.4	0.6	0.2
3009	3	0.2	0.5	0.3
3010	3	0.3	0.9	0.6
3011	3	0.7	0.7	0
3012	3	0.4	0.3	-0.1
3013	3	0.5	0.6	0.1
3014	3	0.3	0.7	0.4

3015	3	0.8	0.8	0
3016	3	0.5	0.9	0.4
3017	3	0.7	0.8	0.1
3018	3	0.5	0.6	0.1
3019	3	0.7	0.8	0.1
3020	3	0.6	0.6	0
3021	3	0.4	0.7	0.3
3022	3	0.2	0.7	0.5
3023	3	0.6	0.8	0.2
3024	3	0.4	0.5	0.1
3025	3	0.6	0.5	-0.1
3033	3	0.6	0.6	0
3040	3	0.6	0.8	0.2
3042	3	0.2	0.4	0.2
4001	4	0.5	0.6	0.1
4002	4	0.6	0.5	-0.1
4003	4	0.7	0.7	0
4004	4	0.5	0.8	0.3
4005	4	0.9	0.4	-0.5
4006	4	0.9	0.8	-0.1
4007	4	0.7	0.8	0.1
4008	4	0.6	0.7	0.1
4009	4	0.2	0.5	0.3
4010	4	0.7	0.8	0.1
4011	4	0.8	0.5	-0.3
4012	4	0.5	0.8	0.3
4013	4	0.5	0.5	0
4014	4	0.7	0.4	-0.3
4015	4	0.4	0.3	-0.1
4016	4	0.8	0.8	0
4017	4	0.4	0.4	0
4018	4	0.6	0.6	0
4019	4	0.3	0.6	0.3
4020	4	0.5	0.5	0
4022	4	0.6	0.7	0.1
4023	4	0.8	0.8	0
4025	4	0.6	0.6	0
4026	4	0.7	0.7	0
4027	4	0.5	0.5	0
4028	4	0.9	0.6	-0.3
4029	4	0.6	0.4	-0.2
4032	4	0.8	0.9	0.1
4034	4	0.6	0.7	0.1
4035	4	0.6	0.5	-0.1
4036	4	0.4	0.7	0.3
4037	4	0.6	0.6	0

4038	4	0.2	0.2	0
4039	4	0.8	0.8	0
4040	4	0.7	0.5	-0.2
4041	4	0.7	0.7	0
4042	4	0.4	0.7	0.3
4043	4	0.8	0.6	-0.2
4044	4	0.4	0.3	-0.1
4045	4	0.6	0.7	0.1
4046	4	0.5	0.5	0
1011	5	0	0.2	0.2
4024	5	0.4	0.6	0.2
5000	5	0.3	0.9	0.6
5001	5	1	0.9	-0.1
5002	5	0.6	0.5	-0.1
5003	5	0.9	0.7	-0.2
5004	5	0.6	0.8	0.2
5005	5	0.7	0.9	0.2
5006	5	0.4	0.8	0.4
5007	5	0.4	0.8	0.4
5008	5	0.3	0.9	0.6
5009	5	0.7	0.7	0
5010	5	0.6	0.7	0.1
5011	5	0.3	0.7	0.4
5014	5	0.4	0.6	0.2
5015	5	0.8	0.6	-0.2
5016	5	0.6	0.5	-0.1
5017	5	0.9	0.7	-0.2
5018	5	0.1	0.1	0
5019	5	0.7	0.7	0
5020	5	0.5	0.6	0.1
5021	5	0.5	0.7	0.2
5022	5	0.7	0.9	0.2
5023	5	0.4	0.6	0.2
5024	5	0.6	0.6	0
5025	5	0.4	0.8	0.4
5026	5	0.6	0.7	0.1
5027	5	0.4	0.6	0.2
5028	5	0.4	0.8	0.4
5029	5	0.5	0.6	0.1
5030	5	0.6	0.7	0.1
5031	5	0.5	0.6	0.1
5032	5	0.7	0.9	0.2
5033	5	0.7	0.8	0.1
5034	5	0.8	0.9	0.1
5035	5	0.7	0.8	0.1
5036	5	0.3	0.9	0.6

5037	5	0.5	0.7	0.2
5038	5	0.5	0.7	0.2
5039	5	0.5	0.7	0.2
5040	5	0.8	0.8	0
5041	5	0.9	0.6	-0.3
5042	5	0.9	0.7	-0.2

Bibliography

13111, E. O. (1999). *Using Technology to Improve Training Opportunities for Federal Government Employees*. Retrieved 12 Jan, 2004, from http://www.opm.gov/hrd/lead/Policy/EO13111.asp

Bredo, E. (1994). Reconstructing Educational Psychology: Situated Cognition and Deweyan Pragmatism. *Educational Psychologist, 29*(1), 23-35.

Brown, J., Collins, A., & Duguid, P. (1989). Situated Cognition and the Culture of Learning. *Educational Researcher, 18*(1), 32-42.

Bruner, J. (1990). *Acts of Meaning*. Cabridge, MA: Harvard University Press.

Carter, V. (1996). Do Media Influence Learning? Revisiting the Debate in the Context of Distance Education. *Open Learning, 11*, 31-40.

Chen, S. (2002). A Cognitive Model for Non-Linear Learning in Hypermedia Programmes. *British Journal of Educational Technology, 33*(4), 449-460.

Daft, R. L., & Lengel, R. H. (1986). Organizational Information Requirements, Media Richness and Structural Design. *Managment Science, 31*(5), 554-571.

Daft, R. L., Lengel, R. H., & Trevino, L. K. (1987). Message Equibocality, Media Selection, and Manager Performance: Implications for Informaiton Systems. *MIS Quarterly, 11*, 355-366.

Department of Defense Strategic Plan for Advanced Distributed Learning. (1999). Washington, DC: U.S. Government Printing Office: Undersecretary of Defense (Readiness).

Dooley, D. (2001). *Social Research Methods* (6th ed.). Upper Saddle River, New Jersey: Prentice Hall.

Ex-FBI Spy Hanssen Sentenced to Life, Apologizes. (2002). Retrieved 12 Aug, 2003, from http://www5.cnn.com/2002/LAW/05/10/hanssen.sentenced/index.html

Goodgride, E. (2002). E-Learning Offers Employee Performance Boost. *Information Week*, 884,842.

Greeno, J. (1991). Number sense as a situated knowing in a conceptual domain. *Journal for Research in Matematics Education, 22*(3), 170-218.

Hansen, E. J. (1989). Interactive Video for Reflection: Learning Theory and a New Use of the Medium. *Educational Technology, 29*, 8.

Hung, D. (2001). Theories of Learning and Computer-Mediated Instructional Technologies. *Educational Meida International, 38*(4), 7.

Kraidy, U. (2002). Digital Media and Education: Cognitive Imnpact of Information Visualization. *Journal of Educational Media, 27*(3), 95-106.

Kumar, D. D. (1991). A Meta-Analysis of the Relationship between Science Instruction and Student Engagment. *Educational Review, 43*(1), 49-62.

Lengel, R. H., & Daft, R. L. (2001). The Selection of Communicaiton Media as an Executive Skill. *The Academy of Management Executive, 11*(3), 225-232.

Lim, K. H., & Benbasat, I. (2000). The Effect of Multimedia on Perceived Equivocality and Perceived Usefulness of Information Systems. *MIS Quarterly, 24*(3), 23.

Maturana, H., & Varela, F. (1987). *The Tree of Knowledge: The Biological Roots of Knowledge*. Boston, MA: Shambhala.

Miller, G. R., & Stiff, J. B. (1993). *Deceptive Communicaiton*. Newbury Park, CA: Sage.

Neter, J., Kunter, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). *Applied Linear Statistical Models* (4th ed.). Irwin: Chicago.

Nolen, J. L. (2003). Multiple Intelligences in the Classroom. *Education, 124*(1), 5.

Pastore, M. (2001). *Companies, Universities Moving Toward E-Learning.* Retrieved 7 Aug, 2003, from http://cyberatlas.internet.com/markets/education/article/0,,5951_737341,00.html

Phipps, R., & Merisotis, J. (1999). *What's the Difference?: A Review of Contemporary Research on the Effectiveness of Distance Learning in Higher Education.* Washington: Free Hand Press: The Institute for Higher Education Policy.

Piaget, J. (1960/1981). *The Psychology of intelligence*. New Jersey: Littlefield, Adams & Co.

Prewat, R. (1996). Constructivism, Modern and Postmdern. *Educational Psychologis, 31*(3/4), 215-225.

Rathis, J. (2002). Improving Instruction. *Theory Into Practice, 41*(4), 5.

Robson, J. (2000). Evaluating On-line Teaching. *Open Learning, 15*(2), 22.

Rosenberg, M. (2001). *E-Learning: Strategies for Delivering Knowledge in the Digital Age*. New York: McGraw-HIll.

Rowe, H. (1991). Introduction: Paradigm and Context. In H. Rowe (Ed.), *Intelligence: Reconceptualizationa and Measurement*. Lawrence Erlbaum Associates: Hillsdale, NJ.

SECAF. (1 Sep 2000). *AFI 40-402: Protection of Human Subjects in Biomedical and Behavioral Research*. Retrieved 5 Feb, 2004, from www.e-publishing.af.mil/pubfiles/af/40/afi40-402/afi40-402.pdf

Skinner, B. F. (1974). *About behaviorism*. Penguin, London.

Thurston, P. W., & Earnhardt, C. C. (2003). *Going Further With E-Learning: Factors Influencing Persistence, Transfer and Intentions to Continue.* Paper presented at the Draft.

Thurston, P. W., & Reyolds, K. A. (2002). *Going the Distance with E-Learning: Motivational Factors that Influence Course Completion.* Paper presented at the

Proceedings for the Interservice/Industry Training, Simulation and Education Conference.

Vrij, A., Kneller, W., & Mann, S. (2000). The effect of informing liars about Criteria Based Content Analysis on their ability to deceive CBCA-raters. *Legal and Criminological Psychology*, *5*, 57-70.

Watson, B. W. e. a. (1991). *Military Lessons of the Gulf war*. Novato, CA: Presidio Press.

Winograd, T., & Flores, F. (1986). *Understanding computers and cognition.* New York: Addison Wesley.

Zielinski, D. (2000). Can You Keep Learners Online? *Training*, *37*, 64-71.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and Nonverbal Communications of Deception. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 14, pp. 1-59). New York, NY: Academic Press.

Vita

Captain Jachin Sakamoto graduated from Hilo High School in Hilo, Hawaii in June of 1995. Within a month following graduation, he entered into basic training at the United States Air Force Academy, from which he graduated with a Bachelor of Science degree in Operations Research and minors in German and Mathematics, in June of 1999. He was commissioned and entered active duty on 02 June 1999.

Since graduation, he has been assigned to Vandenberg Air Force Base in California and Schriever Air Force Base in Colorado. His primary duties have included satellite operations crew commander, satellite operations instructor, and satellite operations evaluator. In 2002 he was competitively selected by the National Reconnaissance Organization to attend the Graduate School of Engineering and Management, Air Force Institute of Technology, at Wright-Patterson Air Force Base in the degree track of Information Systems Management. Following graduation, he will be assigned to National Reconnaissance Headquarters in Chantilly, Virginia.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 074-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of the collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to an penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
23-03-2004	**Master's Thesis**	Aug 2002 – Mar 2004

4. TITLE AND SUBTITLE AN EXPERIMENTAL INVESTIGATION OF PRODUCTIVITY STRATIFICATION & KNOWLEDGE TRANSFER IN AN ELECTRONICALLY MEDIATED ENVIRONMENT	**5a. CONTRACT NUMBER** **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER**
6. AUTHOR(S) Sakamoto, Jachin, Captain, USAF	**5d. PROJECT NUMBER** ENR # 2002-093 **5e. TASK NUMBER** **5f. WORK UNIT NUMBER**
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way, Building 641 WPAFB OH 45433-7765	**8. PERFORMING ORGANIZATION REPORT NUMBER** AFIT/GIR/ENV/04M-19
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) AFMC/AFRL/Air Force Office of Scientific Research (AFOSR)/NM 801 n. Randolph St., Rm 732 Arlington, VA 22203-1977	**10. SPONSOR/MONITOR'S ACRONYM(S)** **AFOSR/NM** **11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION/AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

The idea of anywhere and anytime learning is enticing from a military standpoint, given the high deployment rates in the current operational environment. Electronic-based learning is seen as an answer to this requirement. Currently there are many variations in electronic-based instructional media, and little has been done to determine which format or combination of formats is most conducive to facilitating knowledge transfer and learning. The research project explores, through the use of an experiment, three primary constructs of media richness, content flexibility, and forced engagement, in their relation to effectiveness or productivity in facilitating learning in the experimental participants. The instructional subject matter of choice in this experimental research was the art of detecting deceptive communication. Within the confines of this study, little empirical evidence was found to support the idea that any of the specific variations of electronic training medium outperformed the others in facilitating knowledge transfer.

15. SUBJECT TERMS
Experiment; Agent99; learning; media richness; content flexibility; forced engagement

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					David P. Biros, Lt Col, USAF
REPORT	ABSTRACT	c. THIS PAGE			**19b. TELEPHONE NUMBER** *(Include area code)*
U	**U**	**U**	**UU**	123	(703) 601-4504; e-mail: david.biros@pentagon.af.mil

Standard Form 298 (Rev: 8-98)
Prescribed by ANSI Std. Z39-18