THE UNIVERSITY OF CALGARY

# A Comparison of Reference Learning and Comparative Feedback Using Aircraft Recognition Multimedia Software With Military Personnel

by

Mark Owen Wylie

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

CALGARY, ALBERTA

JUNE, 1999

0-612-48366-5

# ABSTRACT

The Purpose of this thesis is to compare reference learning and comparative feedback using aircraft recognition multimedia software with military personnel. As part of an aircraft recognition multimedia software package called *Know Your Combat Jets*, 56 military participants were randomly assigned to one of three groups, model building, reference learning, and comparative feedback. All participants completed the Eysenck Personality Inventory to investigate personality differences in relation to task performance. Each group was allowed to interact with the instructional software with the alternate forms of learner control for a maximum of two hours. Then each soldier was given an aircraft recognition test to determine the effectiveness of each instructional method, followed by an exit survey of attitudes toward the experience.

The results showed significant differences between the comparative feedback and reference learning groups' performance on the test. On the exit questionnaire participants rated the experience as enjoyable and the software as being effective. Possible reasons for these results and implications for future instructional software design were discussed.

## ACKNOWLEDGEMENTS

I would like to thank the following for their assistance in the completion of this thesis:

Dr. Larry Katz for his time, energy, insight, and enthusiasm towards research in educational technology, and his overall assistance as my supervisor in my Masters program.

Dr. Ernie Chang, Dr. Dan McDougall, and Dr. Peter Harasym for their support and feedback as committee members.

# TABLE OF CONTENTS

Approval Page ..... ii

Abstract ..... iii

Acknowledgements ..... iv

Table of Contents ..... v

List of Tables ..... viii

List of Figures ..... x

INTRODUCTION ..... 1

LITERATURE REVIEW ..... 4

Operational Definition of Learner Control ..... 4

Learner Control Research ..... 8

Summary of LC Research ..... 21

Personality and CBI ..... 22

The Need For CBI Air Craft Recognition Software ..... 26

Classification Learning and Categorical Perception ..... 29

An Example Of A Combat Recognition Instructional Software Program ..... 32

Conclusion ..... 36

METHODOLOGY ..... 38

Participants ..... 38

Description of Groups ..... 39

Materials ....................................................................................................40

Procedure ...................................................................................................49

Research Design and Statistical Analysis ..................................................53

RESULTS .....................................................................................................56

Demographics ..............................................................................................56

Eysenck Personality Inventory Results.......................................................57

Questionnaire Results .................................................................................62

Treatment ....................................................................................................65

Inter-rater Reliability ..................................................................................65

Performance Results ...................................................................................66

Test Scores ..................................................................................................68

Compliance .................................................................................................71

Model Building ...........................................................................................72

Reference Learning.....................................................................................72

Comparative Feedback ...............................................................................72

General Interface Usage..............................................................................73

Label Use ....................................................................................................73

Multi-aspect Use .........................................................................................73

Right and Left Interface Use.......................................................................74

Challenge Mode Use...................................................................................75

Additional Features.....................................................................................76

DISCUSSION ..............................................................................................77

Eysenck Personality Inventory ....................77
Participant Questionnaire ....................78
Performance ....................79
Interface Usage ....................82
Learner Control ....................83
Limitations ....................85
Future Considerations ....................87
Summary and Conclusions ....................88
REFERENCES ....................90
Appendix A : Research Cover Letter and Consent for Participation ....................102
Appendix B : Protocol ....................104
Appendix C : Testing Program Sample Output File ....................105
Appendix D : Sample Data Collection Input File ....................106
Appendix E : Bosnian Conflict Aircraft Subset ....................108
Appendix F : Participant Questionnaire ....................109

## LIST OF TABLES


Table 3.1 : Participant Distribution by Gender. 39

Table 4.1 : Measures of Central Tendency and Dispersion for Participant Age and Education. 57

Table 4.2 : Participant Education Level (1-no High School Diploma, 2 – High School Diploma, 3 - Post-Secondary attempted). 57

Table 4.3 : Measures of Central Tendency for Eysenck Personality Inventory Results (E-extraversion, N-neuroticism, L-Lie scale). 57

Table 4.4 : Adjusted Measures of Central Tendency and Dispersion for Eysenck Personality Inventory Results (E-extraversion, N-neuroticism). 58

Table 4.5 : Extraversion-Introversion Frequencies. 59

Table 4.6 : Neuroticism – Stability Frequencies. 59

Table 4.7 : Stable Extravert and non-Stable Extravert Performance Data Comparison... 62

Table 4.8 : Previous Recognition Training and Computer Experience Measures of Central Tendency and Dispersion (1-none, 5-extensive). 63

Table 4.9 : Participant Computer and Recognition Training Experience Frequencies (1-none, 5-extensive). 63

Table 4.10 : Computer Experience and Treatment Group Cross Tabulation. 63

Table 4.11 : Recognition Training Experience and Treatment Group Cross Tabulation. 64

Table 4.12 : Questionnaire Results - Effectiveness and Enjoyment Number of Cases, Range, Minimum, Maximum, Mean, and Standard Deviation 64

Table 4.13 : Effectiveness and Enjoyment Ratings by Scale ........................................... 64

Table 4.14 : Questionnaire Results - Effectiveness and Enjoyment by Treatment Group – Mean and Standard Deviation........................................................................................ 65

Table 4.15 : Inter Rater Reliability correlation using Spearman's Rho. ........................... 66

Table 4.16 : Measures of central tendency and dispersion for Test Score, Time on Test, and Total Clicks for all participants....................................................................... 66

Table 4.17 : Mean, number of cases, and standard deviation of Test Score, Time on Test, and Total clicks by treatment group. ................................................................. 68

Table 4.18 : One-way ANOVA of Test Score, Time on Test, and Total Clicks. ............. 69

Table 4.19 : Performance Variables Post-hoc Scheffe Multiple Comparison Results (1-Model Building, 2-Reference Learning, 3-Comparative Feedback). ...................... 70

Table 4.20 : Valid N, Mean, Standard Deviation, and Percentage of Multi-aspect use by Treatment. ................................................................................................................ 74

Table 4.21 : Left and Right Side utilization of interface - Number of cases, Mean, and Standard Deviation. ................................................................................................ 75

Table 4.22 : Percentage Use of Challenge Mode - Number of Cases, Mean, and Standard Deviation.................................................................................................................. 76

# LIST OF FIGURES

Figure 2.1 : Front-end Interface for *Know Your Combat Jets*. ........................................ 33

Figure 3.1 : Main Interface Screen for *Know Your Combat Jets* Software. ..................... 43

Figure 3.2 : Opening Screen for the Testing Software ................................................... 44

Figure 3.3 : No Participant Number Entered Dialog Box................................................. 44

Figure 3.4 : On-line Test Interface.......................................................................................... 45

Figure 3.5 : Test Completion Dialog Box............................................................................ 46

Figure 3.6 : Data Collection Screen........................................................................................ 47

Figure 3.7 : Protocol .................................................................................................................. 52

Figure 4.1 : Extraversion – Introversion Frequencies. ..................................................... 60

Figure 4.2 : Neuroticism – Stability Frequencies. ............................................................ 60

Figure 4.3 : Participants' combined Neuroticism and Extraversion Scores. .................... 61

Figure 4.4 : Test Score by Treatment Group (1- Model Building, 2- Reference Learning, 3- Comparative Feedback), Participants are Rank-ordered by Test Score. .............. 71

# CHAPTER 1

# INTRODUCTION

Current research concerning the use of technology for instruction suggests significant learning gains are being achieved. Many of these gains have been accomplished through the development of efficient and effective tools to aid in the instructional process, which has also resulted in lower training costs. Results from educational research have led to improved learner performance and provided direction for future developments using instructional technology. One particular area of instructional technology research is that of learner control.

The issue of learner control involves providing the user with greater control over the instructional session. This control can take many forms, but often involves allowing the user control over content, sequencing, and navigation decisions. Research has provided mixed results concerning the overall effectiveness of learner control. Many instructional variables interact with learner control, making it difficult to generalize or compare different research studies. Some of these variables include content, motivation, and design presentation features. Rather than comparing different instructional packages and levels of learner control, research has suggested a need to modify levels of learner control within instructional programs. This type of approach may help to limit intervening variables such as content and presentation format. Many military organizations are realizing the benefits of integrating technology into their instructional programs. In the military, effective and efficient training are vital when dealing with both life and death situations and restricted training budgets.

Military combat vehicle recognition is an area of training that is extremely important for effective defensive and offensive combat operations. Soldiers involved in aircraft-related combat must be able to identify friendly and enemy aircraft to deal effectively with them. The Canadian Military is currently researching ways to effectively train its soldiers in combat vehicle recognition. The use of computer-based instructional packages may offer the ability to train soldiers inexpensively and effectively. Traditional training methods for recognition have often involved static media such as photographs and slides. Current technology can allow soldiers to view multimedia instructional sequences, and to receive instantaneous feedback for responses.

*Know Your Combat Jets*, designed by Axia Multimedia Corporation, is a software program designed to teach military aircraft recognition. This software utilizes video, audio, images, and text to instruct soldiers to effectively recognize both friendly and enemy aircraft. At the heart of this program is the Discrimination Learning Engine (DLE- U.S. Patent no. 5,449,293). The DLE assists the user to classify and categorize combat aircraft. This is accomplished through the use of a multi-aspect dual window display. The user is requested to identify a particular aircraft and when an incorrect response is made, the user is given feedback concerning the incorrect response, and assisted with the exploration of the misidentification space. The user can then explore a variety of aircraft features such as performance capabilities and detailed orthogonal views, in an attempt to discover salient features that may help distinguish between the correct and chosen aircraft.

The main purpose of this study was to evaluate the effectiveness of three forms of learner control while using the *Know Your Combat Jets* software. The three forms of learner control were very similar. The main difference revolved around the type of access participants had to the DLE Effectiveness was evaluated based upon performance on a test, the time it took to complete the test, the number of mouse clicks used during the instructional sequence, and the results of a questionnaire. Based upon the performance and subjective measures, this thesis attempted to determine:

1. Which type of learner control is the most effective?
2. Do personality differences have an effect upon performance?
3. Are learner preferences consistent with performance?

# CHAPTER 2

# LITERATURE REVIEW

The purpose of this section of this thesis is to provide the reader with an in-depth review of literature in the field of learner control using instructional technology. First, learner control is defined, followed by an overview of current research findings and a review of how personality issues affect computer-based instruction. Then, a rationale for using computer-based instruction for military aircraft recognition training is discussed including identification of the need for accurate and effective combat aircraft recognition. A brief introduction to the software used for this study is also given, including a review of classification learning in relation to object recognition which is a major aspect of the software's Discrimination Learning Engine (DLE). The review concludes with a description of how the software and experimental groups in this study are classified with respect to learner control.

## Operational Definition of Learner Control

Learner control (LC) research with instructional technology may have begun as early as 1968 (Williams, 1993) or perhaps as far back as the 1950s with Skinner's work on Teaching Machines (Ross & Morrison, 1989). Computer software and hardware initially presented content in such a way as to strongly resemble a book and were often referred to as 'glorified page turners'. It was soon realized that computers could provide immediacy of feedback, and could also be programmed to adapt based upon user input. This form of adaptation could mean something as simple as basing content upon the learners' responses to particular questions. Goforth (1994) states that it was also soon realized that "diagnosing the learners' state of understanding from their responses to questions is a hard problem..." (p.2). To address this problem, instructional designers began to attempt strategies that allowed the user to retain some of the responsibility for adaptation which in essence forms the basis for learner control (Goforth, 1994). Much of the origin of LC derived from the studies in the early 1960's that attempted to individualize instruction so as to increase performance and motivation, and decrease learning time (Williams, 1993). In his meta-analysis, Williams indicates that LC also accommodated the popular constructivist and cognitive theories as it had the potential to compensate for individual cognitive differences. Clark (1984) also stated that "The distinctive characteristic of cognitive research is the idea that instruction influences achievement through student thought processes." (p.2). Research on instructional technology was also beginning to show that "Learning gains come from adequate instructional design theory and practice, not from the medium used to deliver instruction." (p.3). Computers were not a panacea for learning, but when utilized effectively they could have the potential to enhance learning.

Reeves (1993) states that "Learner control is defined as the design features of computer-based instruction that enable learners to choose freely the path, rate, content, and nature of feedback in instruction." (p. 821). Milheim and Azbell (1988) state that LC is "the degree to which a learner can direct his or her own learning process." (p. 461). Some specific examples of how learners' could take control during computer-based instruction (CBI) are control of content sequence, control of learning activities, and control of pacing and content presentation. Young (1996) states that : "providing learners with control over various instructional decisions individualizes the lesson by allowing learners to choose the amount and/or type of instructional support they require, thus tailoring the lesson to meet their unique needs." (p.17).

Once it was conceived that modifying the amount and type of LC enhanced the learning process, and with the decline in costs of instructional software and hardware, LC research began to steadily increase in popularity (Schnackenberg & Sullivan, 1995; Ross & Morrison, 1989). While commenting upon the difficulties inherent in LC research Reeves (1993) stated that "...at least part of the problem may be insufficient precision in defining just what is meant by "learner control" in various studies, and that learner control probably is not a unitary construct, but rather a collection of strategies that function in different ways depending upon what is being controlled and by whom." (p. 822).

LC is a difficult term to define, particularly because of the multitude of studies that coin the phrase LC, yet are uniquely different in a variety of ways. LC can be thought of as a continuum of how much control a user has over a particular instructional software program. A less technical definition could be that LC is essentially a scale (from no user control to total user control) that measures how much control a person has over a set of instructional materials that reside on some technological device. An important qualifier to this statement is that "...learner control input is implicated in management of the learning experience and is independent of the topic of the tutorial" (Goforth, 1994, p.3). While it is true that the management aspect is vital, it is also important to realize that that the LC aspect is not necessarily mutually exclusive of content. The content of a particular instructional sequence can influence certain learner variables such as motivation and interest, as can the design of the learner interface itself. This could produce results, which do not generalize or are not applicable to other LC computer-based instruction (CBI) designs.

Dependent upon the actual research study in question, the definition of learner control may be affected by a multitude of factors. Certainly the base definition may be identical between separate studies, but there could very easily exist significant differences between the studies that could affect the comparability and generalizability of the results. Take for example research concerning the effects upon performance by allowing the learner to control feedback. Although two separate studies may allow this same type and amount of control, there are other factors not readily apparent that may interfere with the usefulness of the results. Perhaps one study involved video sequences while the other did not. One program could have supplied varying levels of motivation due to content presentation style. These differing factors could interfere with the generalizability of the results. Snow (1980) points out that learner control cannot overcome the fact that "individual characteristics not under control of the individual will determine to a significant extent what and how much that individual will learn in a given instructional setting" (p.1). Another consideration is the purpose of the software itself, the program may provide direct instruction, simulation, indirect learning, or another approach. As discussed below, these and other factors may affect the validity and usefulness of LC research, and enforces the need for a clear definition of not only LC but also the learning environment when conducting LC research.

**Learner Control Research**

To better understand the research results it is important to examine the other constructs directly related to LC. The complete opposite of LC is often considered to be program control (PC). The use of PC suggests that the learner follows a sequential progression through the learning materials, using the order and timing provided by the software. Milheim (1986) stated that most earlier studies on LC and PC "...have dealt with differences in achievement based upon the availability of instructional control" (p.303).

In other words users either received a level of LC or were required to be subject to complete PC. Gay (1986) performed a typical early study on LC verses PC. In her study, students were assigned to either a PC or LC treatment involving computer-assisted video learning. Gay randomly assigned individuals to each treatment group to try and account for individual differences. Her study focussed upon whether learners' prior conceptual understanding would effect the usefulness of either LC or PC. The dependent variables were posttest score and time on task. Her results showed that there appeared to be an interaction between LC and prior familiarity with material presented, and that students with higher previous understanding would perform better under LC. An important finding in her study was that subjects did not appear to use good learning strategies when given control, especially concerning sequencing decisions.

Carrier and Williams (1988) studied task persistence in relation to LC and PC. One hundred and fourteen sixth graders participated in a study that determined the amount of task persistence (in relation to motivation this meant the degree of effort expended in relation to their learning) comparing a learner control group with a program control group. The researchers found that the best performers in PC were those students with medium, not high or low, levels of task persistence. In the LC group the best performers belonged to the high task persistence group.

Schloss, Wisniewski, & Cartwright (1988) stated that "Research to date has, in general, demonstrated that program control is more effective than learner control in aiding student learning." (p.142). However, review of literature prior to 1988 would suggest that there are mixed results. Ross and Morrison (1989) questioned certain aspects of information technology studies that focussed upon LC research. The authors stated that it was important to point out some of the flaws inherent in LC research. The main point discussed was the separation of media from methods in research designs, and improving the use of media replications to improve the generalizability of findings. They pointed out that external validity is at risk when researchers focus upon the interaction of medium and method. In the field of technology, where hardware and software often become obsolete as soon as they are integrated or utilized, this can pose a major problem. They concluded with summaries of their past research findings that showed low ability learners do not benefit from LC when they can modify amount and type of instruction, but could benefit from LC that allowed modification of presentation style.

An advantage of LC is that it allows content designers to utilize the learner to make instructional decisions, rather than attempting to design elaborate monitoring algorithms. Scardamalia, Bereiter, McLean, Swallow, & Woodruff (1989) stated that "The computer environment should not be providing the knowledge and intelligence to guide learning, it should be providing the facilitating structure and tools that enable students to make maximum use of their own intelligence and knowledge." (p.54). Scardamalia et. al.

referred to this process as procedural facilitation. Complementing and furthering this study, Gay (1989) found that learners are better off if they do not have maximum control over sequencing of instructional content, other than that of review. She also pointed out that instructional software designers could benefit from assessing individual characteristics of software users.

Another issue of interest for instructional designers was whether the incorporation of a review component would aid performance during instruction. Allen (1990) examined this question in a study involving 58 Army ROTC cadets that were given a map reading exercise with the option of review added. No significant differences were found in LC and PC performance through the incorporation of this review component. A similar study by Kinzie, Sullivan, and Berdel (1992) found similar results when users were given complete control over how much content review they would receive. Most of the procedures in these LC research studies involved a continual opportunity throughout the lesson for content review rather than just at the end and this may have contributed to the lack of significant differences between LC and PC groups.

Lee and Lee (1991) examined the effectiveness of LC verses PC in relation to learning phase (knowledge acquisition verses knowledge review) and differences in prior knowledge. They concluded that PC is superior to LC for low level domain knowledge, but this difference becomes less significant when the level of domain knowledge increases. That is, the differential effects of PC and LC were more prominent when the learner lacked knowledge relevant to the domain in question. In a related study, Mattoon, Thurman, Klein, and DeVilbiss (1991) found that learners may not be able to effectively control sequence and pace when the subject matter is unfamiliar. Many other studies provided similar observations where prior domain knowledge and level of ability often determined the effectiveness of LC (e.g., Cho, 1995; Henry, 1995; Tennyson & Park, 1985). Learner control with advisement (LCWA) seemed to help compensate for this problem with varying degrees of success.

Sales and Williams (1988) examined the effects of PC with an enhanced or an adaptive form of LC. Adaptive LC involved modifying the users ability to exercise control over sequence, based upon performance. If a student were performing a particular task poorly, they would lose control of feedback options until such a time as their performance improved. This could also be considered as a form of learner control with advisement (LCWA). LCWA helps address the issue raised by Hannafin (1984) and Steinberg, Baskin, and Hofer (1986) that showed learner control of feedback may not be effective (as reported in Williams, 1988). Elaborative feedback (contingent upon student performance) provided right or wrong feedback and additional information to aid the student in correcting misconceptions. Since many students chose the elaborate feedback options, the study was inconclusive in determining a relationship between locus of control and feedback and performance.

LCWA typically meant that although the user had control of certain processes, the computer program would enhance or mediate that control by giving or offering the opportunity to review information at pertinent moments during instruction. LC offered the same amount of control, minus the advisement capabilities. LCWA can be considered as a restricted form of LC due to forced presentation of advisement information. Arnone & Grabowski (1992) undertook a study that examined this variant of LC. The LC allowed participant control over sequencing, pacing, stopping, remediation, and exiting. The advisement treatment group received identical capabilities to the LC group but was supplemented with advisement strategies at any point where navigational changes were initiated that could result in missing important information. If a student chose to miss a section of material, the computer might bring up a dialog box that would display something like "Are you certain that you wish to skip this section?", and also provided motivational statements such as "This next section is very interesting. You might really enjoy it." (p.20). The results of the study showed that the LCWA group scored significantly higher on the achievement posttest than the LC or PC group, while no significant differences were found between the PC and LC group.

Research in learner control has shown that when students attempt to acquire new knowledge in CBI that they often make poor decisions regarding learning needs (Tennyson & Park, 1985). Typical examples of ineffective use of LC involved students who progressed to the evaluative stage before mastery of content was complete. Gay (1985) also points out that "Although learner control seems to be able to adapt to individual differences, research dealing with the variable of learner control has failed to demonstrate that learners can make and carry out good decisions" (p.3). In a recent LCWA study, Shin, Schallert, and Savenye (1994) found that learners with different levels of prior knowledge require different kinds of instructional approaches. This was a result of determining that low prior knowledge students did not make effective use of advisement strategies and seemed to perform better when given less control. This phenomenon of not taking control or not making effective use of that control has occurred in many different studies (Arnone & Grabowski, 1991; Mattoon et. Al., 1991; Chang, Katz, & Wylie, 1997).

Attitudinal factors related to the use of LC have also been examined in a number of studies. Milheim (1986) examined attitude in relation to pacing and sequence control during CBI and discovered that while students indicated a positive attitude towards the use of LC, there did not exist significant differences regarding the LC and PC groups. Crooks, Klein, Jones, & Dwyer (1996) found that individuals expressed positive attitudes when given control during CBI.

Other researchers felt that faults existed with particular factors involved in LC research. In a paper entitled, *Pseudoscience in Computer-based instruction: The case of learner* *control research.* Reeves (1993) detailed problems with definitions of LC, theoretical foundations, treatment duration, outcome measures, sample sizes, and conclusions. Although Reeves (1993) stated that there exists a lack of theoretical foundation for why LC may be useful, many papers reviewed in this thesis have made reference to relevant theories and research that have indicated that in certain circumstances, LC can enhance instruction. For an in-depth review of these and other studies, see the analyses done by Goforth (1994), and Niemiec, Sikorski, & Walberg (1996).

Reeves (1993) also stated that another area of concern was the amount of studies that contain infrequent and brief instructional treatments that "...do not provide learners with sufficient experience for learner control variables to be "actualized."" (p.42). This statement was supported by Cronbach and Snow (1977) who cautioned that ten or more interactive sessions were necessary to acquaint students with innovative instructional treatments. Reeves (1993) mentioned a number of studies that offer an average of approximately thirty minutes of instructional treatment per participant. This seems to contradict Romiszowski's (1986) recommendations that designers provide learner control over micro-instructional experiences occurring within lessons whose duration spans minutes. Again, the duration of instructional treatments are heavily related to content, presentation design, and the user interface in general. Given the variety and levels of instruction it is quite difficult to generalize a minimum time requirement for instructional treatments.

Williams (1992) and Goforth (1994) did meta-analyses* of the state of LC research. In Williams (1992) meta-analysis over seventy different LC research efforts were examined. In most studies no differences were found between LC and PC treatments, and many researchers utilized this 'no-difference' effect to support either a PC or LC model based upon their own position. For example, if a study showed no effect for LC treatment verses PC, but users enjoyed the opportunity for LC, it was often suggested that LC should be used because the participants enjoyed the experience. Time-on-task was also a variable of interest, and the analysis found that the majority of studies showed that LC subjects took less time to complete instructional sequences than PC subjects did. The issue of effective use of time was regarded as non-conclusive, mainly due to confounding of instructional control, time-on-task, and amount of instructional material observed.

Few reports analyzed whether participants actually made effective use of time or simply skipped instructional sequences due to poor decisions. To address the problem of participants choosing insufficient levels of instruction, Williams stated that research in general has shown that LC should only allow participants to select context, sequence, and presentation variables, but not allow them to alter instructional events that could reduce the amount of content support given. The addition of an adaptive module that informs the learner of their current state of learning, yet which phases itself out upon increased levels of mastery, could positively enhance the effects of LC.

Williams (1992) concluded that LC can be effective if:

1. Students have the capacity to make rational choices.
2. Students are provided with a motivating environment.
3. Students' prior knowledge and achievement are considered.

William's (1992) research has found a curvilinear effect between student motivation and learning, but this may very well be related to confounding variables such as medium design and content presentation. Locus of control and emotional state are variables requiring further examination regarding their influence upon learner control and learning with CBI. A suggestion to design feedback responses that attempt to specifically counteract maladaptive attributional patterns is also made, although this should also be adaptive as Williams own research in this area had shown that when applied to those students who did not have maladaptive strategies it had a negative effect (Williams, 1992). Williams (1992) stated that we should be asking the question: "How can I make learner-controlled CBI more effective?" (p.23).

In a more recent meta-analysis Goforth (1994), redefined LC within a general model of decisions and information in instructional systems in an attempt to confirm prior research demonstrating the effectiveness of LC. Goforth (1994) has developed a definition of LC in an attempt to improve the categorization of previous and future experimental systems.

This definition focuses upon the responsibility entrusted to the learner in each environment, particularly the kind and amount of control. Out of the forty-eight studies in the analysis, thirty-three exhibit a ceiling effect, where subjects in LC are only offered the choice of receiving what the PC groups received and nothing more. It is more difficult for LC groups in these studies to outperform the PC groups since they receive no more content assistance regardless of LC decisions made during instruction. Therefore, in those cases Goforth (1994) regarded non-significant results between PC and LC as a favourable condition for LC. Support for this position can be found in research studies which have shown that regardless of achievement effects, LC groups often spend less time and show more improved attitudes toward instruction.

Goforth's model of control is as follows:

i. Current conditions represent the level of mastery already achieved as a ratio of problems successfully solved to problems tried and a record of activity in the unit.

ii. Subgoals are (i) understand the procedures, (ii) understand example problems, (iii) do sample problems.

iii. Methods available to reach the subgoals are (i) text descriptions of procedures, (ii) video sequence of procedures, (iii) several completed example problems, (iv) several problems for interactive solution.

iv. Evaluation of outcomes is minimal for the first three methods - all that is known is that the learner observed them for how long; for the final method, the outcome is either right, wrong, or no answer. Whatever the outcome, it is incorporated in a new determination of current conditions. (p.7)

To summarize, current learner state is determined; subgoals are identified; methods are selected and applied; and finally the outcome is evaluated (how close is the current state to the desired state?).

In Goforth's (1994) model, the next stage is to determine where control resides. Four types of control are utilized: (i) private decisions by the learner, (ii) private decision by the software, (iii) announced (input) decision from the learner, and (iv) transparent decision (output) from the computer. This categorization of control allows each of the studies to be sorted in accordance with the types of control, and the cycle of state readiness (state, goal, method, and finally evaluation). This is an extremely useful model in that it encompasses a variety of studies due to the "definition" of control. As an example of how a drill and practice program would fit the model, the author describes it as being one where the computer has compete control. The current state and sub-goals are determined by the software (and not explicitly stated to the user), while the methods and evaluation are also determined by the software, but are stated in the form of output to the user. A situation where the user has control would be one where the user determines the initial state of their knowledge, and thus their path towards a goal state is self-determined. Then the user chooses the amount of verification problems, and the computer performs the final evaluation of this state of readiness by monitoring the output.

Goforth's use and subsequent analysis of this model has produced some interesting (although definitely exploratory) results. First, it appears to be important that the learner have some form of control, although it is not clear where that control must exist within the process of state, goal, method, and evaluation. This may involve considerations of content, learner, and presentation style. Goforth does not find any evidence to suggest the importance of providing the learner with information about current conditions. Although Goforth states that his model is crude, it does contribute a new way to analyze and design CBI software which utilizes LC. Goforth concludes with the statement "What is needed now is a shift in focus to analysis of how control is exercised and how learning strategies are constructed from sequences of decisions." (p.20).

Goforth's model points out a particular rationale for poor performance under LC conditions. For the user to determine a goal path once the current state has been determined he/she needs to have an effective idea of possible strategies or goal paths. If knowledge of the capabilities or familiarity with the software is limited, the user may make ineffective decisions regarding goal attainment.

**Summary of LC Research**

Research has produced inconsistent results when attempting to determine the effectiveness of learner control (LC), particularly due to learner differences, varied instructional objectives, and the context of the computer-based instruction (Young, 1996). Research has also shown that increasing LC may reduce learning time and increase certain positive learner variables such as interest, attitude, effort, and motivation within the particular knowledge domain (Young, 1996; Milheim & Martin, 1991).

In summary, it would appear that to be effective, a CBI program should involve LC, but that LC should dynamically change according to learner, content, and program interactions. The result would be a piece of software that reacts differently with different individuals (allowing alternate levels of LC), and ultimately will have varying levels of effectiveness upon those individuals. Careful and detailed analysis within a particular instructional program could assist with the design of this form of LC module.

There have been a number of summary analyses conducted concerning LC research efforts, but according to Niemiec, Sikorski, & Walberg (1996), Goforth's meta-analysis appears to be one of the more significant. In addition, the majority of reviews (with the obvious exception of Goforth's and possibly one conducted by Ross and Morrison (1989)) tend not to consider the previously mentioned problems with definitions and generalizability. The summary of results from Niemeic, Sikorski, & Walberg (1996) generally support the review conducted by Goforth (1994), particularly in regard to his suggestions for future research.

**Personality and CBI**

LC research has made significant gains towards accommodating certain individual differences such as learning style and speed. Another area of concern to CBI developers is that of personality. Personality issues such as aggression and introversion/extraversion may interact with the effectiveness of CBI and should also be taken into consideration when designing CBI modules, or evaluating the effects of learner control and CBI (Scott, 1995).

Eysenck's model of personality maintains that personality differences may be best expressed in terms of primary and lower order traits (Francis, Katz, & Evans, 1996). Eysenck's theory proposes three main factors, namely, neuroticism, extraversion and psychoticism. Eysenck created the Eysenck Personality Questionnaire (EPI) which, together with a lie scale, rate individuals in regard to their responses on a survey of situational questions (Eysenck & Eysenck, 1968). This questionnaire (and alternate forms based upon it) have been translated into several languages, and have had a significant influence on the study of personality and individual differences (Francis, Katz, & Evans, 1996). In a recent review of the literature surrounding personality and CBI, Pocius (1991) states that "The personality dimension most extensively studied in the human-computer interaction literature is introversion-extraversion." (p.105). Research has shown varying results regarding how these personality issues affect individual performance while using CBI.

The EPI Testing Manual states that a highly neurotic individual is one who is an anxious, worrying individual, who is moody and frequently depressed (Eysenck & Eysenck, 1968). The manual also states that this person may be considered a worrier, often demonstrating higher levels of anxiety. Conversely, the typical extravert "craves excitement, takes chances, acts on the spur of the moment, is carefree, easy-going, optimistic,..." (Francis, Katz, & Evans, 1996, p.165). "The typical introvert is a quiet, retiring sort of person, introspective, fond of books...." (Eysenck & Eysenck, 1968, p.6). The testing manual states that average scores are said to be those between the 31st and 70th percentiles (Eysenck & Eysenck, 1968).

Thackray, Jones, and Touchstone (1974) provided sixty college males with a monotonous, minor physical fatigue computer task requiring continuous discrimination to determine whether introverts behaved differently than extraverts. They discovered that extraverted subjects showed increasing lapses of attention, while introverted subjects failed to show any evidence of a decline. Similar results were discovered by Hoffman and Waters (1982) who found that military subjects using CBI that were classified as extraverts were likely to not complete the training program.

Contradicting those previous studies is a study performed by Goh and Moore (1977) who found that although the EPI introversion/extraversion scale correlated highly with performance, the participants with higher extraversion scores performed better on the academic tasks than the introverts. It is important to note that these tasks were not necessarily computer related and there was no consistency regarding tasks between participants as the correlation was made between the EPI and the participants grade point average. Goh and Moore (1977) also state the need for distinct educational groups when examining the relationship between performance and personality.

Eysenck (1979) found that anxiety was correlated with neurotic introverts, and low-anxiety with stable extraverts. In addition he found that high anxiety subjects tended to engage in task-irrelevant processing which would often reduce the capabilities of working memory. These high anxiety subjects would tend to compensate through increased effort. In a study to determine the effects of anxiety on computer-assisted learning, Spielberger (1970) found that anxiety impaired computer task performance. It is important to note that although students were using CBI materials, the curriculum was mathematics based. Similarly Tallmadge, Shearer, and Greenberg (1968) examined the effects of anxiety on participants involved in a computer-based aircraft recognition task and found that lower anxiety subjects performed better than high.

In an investigation to determine the relationship between personality and computer-related attitudes, Francis, Katz, and Evans (1996), gave 298 female university students the EPI (Hebrew Version), and a computer attitudes scale. The results showed that introverts had a more positive attitude towards computers than extraverts. The authors conclude the study with a discussion of the importance of exploring other samples and conditions where personality effects may differ from those presented in their study.

In his review of studies investigating the relation between personality dimensions and human-computer interaction, Pocius (1991) found that introversion-extraversion and traits characterizing introversion-extraversion are related to many aspects of human-computer interaction. Pocius (1991) defined human-computer interaction as "any process in which the user and computer engage in a communicative dialogue whose purpose is the accomplishment of some task" (p.103). Pocius' review found that introverts perform better on conventional linear tasks, when given maximal structure and guidance, while extraverts performed better with discovery programs requiring minimal structure and guidance.

Clearly research has shown that individual personality differences can have a significant effect upon performance while using CBI. The difficult task is then determining whether these results are consistent when used within different content domains, instructional designs, and participants demographics.

**The Need For CBI Air Craft Recognition Software**

Since the introduction of aircraft into the military there has existed a need to provide accurate detection of both friendly and enemy forces aircraft (DND, Land Force Western Area AFV & Aircraft Home Study Guide, LFW00999, 1995, p.11). Due to the speed and threat inherent in a military aircraft attack, efficiency and effectiveness are essential factors for aircraft recognition (Secrist & Hartman, 1992). In addition, the ability to correctly identify both friendly and enemy aircraft is necessary to avoid attacking friendly aircraft during high activity combat situations. The ability to detect the type of aircraft could also assist in determining threat posture, and possible engagement tactics. With these factors in mind, it is necessary to effectively train military personnel to be able to quickly and accurately recognize military aircraft.

This training has predominantly existed in the form of lectures with books, photographs, and slides (DND, Land Force Western Area AFV & Aircraft Home Study Guide, LFW00999, 1995). A typical training session would involve a lecturer standing in front of the class showing slides of aircraft and discussing their salient features and threat capabilities. After the session, the group would be shown a slide of a particular aircraft and be expected to determine whether it was friend or foe, and to determine the type of aircraft. This approach to recognition training is slowly being adapted to incorporate technology into the presentation and comprehension tasks, particularly as computer training has been identified as having the potential to provide better instruction than that of using traditional means in certain content domains (Matoon & Klein, 1993; Bridgeman & Post, 1985).

A programmed instruction package (PIP) in the military consists of a book which presents information to the learner and which is normally followed with a comprehension test. This learning format often involves the use of slides or still images for recognition training (DND, Land Force Western Area AFV & Aircraft Home Study Guide, LFW00999, 1995, p.4). This type of learning could be easily incorporated into a technological model as computers work well at providing sequential information presentation and performance feedback. Due to previously high costs of technology and limited software, the military has, until recently, used these PIP's for training purposes (DND, Land Force Western Area AFV & Aircraft Home Study Guide, LFW00999, 1995).

The Canadian military recognized problems with their recognition training methods, and "...training endeavors in many cases were diminished by relatively unskilled instructors, insufficient training aids, and a lack of direction." (DND, Land Force Western Area AFV & Aircraft Home Study Guide, LFW00999, 1995, p.3). To address this problem and others, the Land Forces Command Headquarters undertook a feasibility study of computer assisted learning options for future aircraft recognition training in 1992 (DND, Land Force Western Area AFV & Aircraft Home Study Guide, LFW00999, 1995, p.3). The results of this study recommended that consideration be given for the possibility of incorporating computer assisted learning into future learning module development. Allen (1990) points out that the US army recognized a major source for achieving better quality of training was through instructional technology.

Current technology, particularly the use of instructional multimedia, has the capability of providing better instruction than the PIP method. Rieber, Boyce, and Assah (1989) state that computer animations (not available in traditional recognition training packages) can enhance learning, thus supporting the use of CBI for recognition training instruction. In one of his papers on computer-based instruction, Clark (1984) stated that "Highly directed instruction produces the kind of behaviorally-based learning often desired by military and business clients." (p.4).

Computer technology currently has the capability of integrating a variety of multimedia sequences (video and audio for example) into an instructional program. Software designers have also been incorporating a variety of learning models into their programs. These models often include the ability to monitor progress, provide feedback, and control pacing (Allessi & Trollip, 1991). With these advances in capability and the relatively low costs of computer power, the learner can now interact with a software program that can provide much more than the traditional lecture and slide presentation. However, in a study of computer-based and paper-based recognition training performance, Frederico and Aiken (1989) found that there were no significant differences favoring either paper or computer based recognition training models.

**Classification Learning and Categorical Perception**

The task of object recognition can be described as a qualitative difference in how similar things look depending on whether or not they are co-classified in the same category (Harnad, 1987). Harnad (1987) referred to this phenomenon as categorical perception (CP), "the compression and expansion of similarities and differences between things, depending upon whether they belong to the same category (in which case they look similar) or to different categories (in which case they look different)" (p.2). Tarr (1995) states that "The mechanisms used in human object recognition are almost certainly a product of many factors, including the task, the learning and retrieval contexts, and the functional and visual relationships between objects both encoded in memory and observed in the environment." (p.73).

In a discussion of how cognitive psychology findings can raise the productivity of CBI, Hodges (1986) stated that "people need to make mistakes in classifying examples before they can discriminate accurately among positive examples and near misses." (p.4). Hodges (1986) also described the following as effective ways to modify CBI to enhance learning:

1. Provide practice in order to build discrimination;
2. Add new concepts, one or two at a time, to the set of concepts to be learned; and
3. Provide the user with control over the program. (p.6).

In his review of categorization, Goldstone (1994) stated that "Interactions between compared objects, and task and stimulus variables, also serve to resolve uncertainties about the basis of similarity." Consideration must be given for alternate aspects of similar concepts, such as examining different properties of similar objects (Goldstone, 1994). For example, when attempting to differentiate between two similar aircraft, one should have the capability of exploring different features of both aircraft (such as three-dimensional and orthogonal views) to assist in the categorization and classification task.

Harnad (1987) stated that "successful categorization depends upon finding the critical features on the basis of which reliable, correct performance can occur." (p.9).

Siegel and DiBello (1980) performed a recognition training study with simple letter computer-based recognition tasks. It was found that adaptive feedback with discrimination training significantly reduced the number of discrimination errors on the posttest. Similar results were found in a later study by Siegel and Misselt (1984), with the added finding that the adaptive feedback group completed the posttest significantly faster than any other group.

Siegel and Misselt (1984) describe discrimination feedback as that which "tells what stimulus the student responded to as well as the answer to the stimulus that was actually presented" (p.311). Discrimination training involves the simultaneous presentation of the missed item and the item with which it was confused to allow the student to study their similarities and differences (Siegel & Misselt, 1984). Park (1984) examined three experimental studies involving discrimination learning and technology and found that "if instruction continued without facilitating the discrimination learning between the two confused concepts, the confusion would interfere not only with learning the two concepts, but with learning other coordinate concepts" (p.15). This presentation of items to be compared should be presented simultaneously to facilitate discrimination (Park, 1984). Bavelier (1994) found that by not using a multi-window view to simultaneously view concepts, one increased the possibility of being susceptible to what she terms, repetition blindness (RB). RB can be defined as the failure to see or recall the second of two visually similar or identical items in rapid serial visual presentation.

**An Example Of A Combat Recognition Instructional Software Program**

There exists a broad base of research concerning the use of technology in the learning milieu. Often these studies involved different instructional content domains and learner tasks. However, few studies have examined the effectiveness of computer technology for aircraft recognition training.

Axia Multimedia Corporation has developed one of the few programs currently available for identifying military aircraft called, *Know Your Combat Jets* (Chang, Katz, & Wylie, 1997). This CBI program involves the use of a variety of multimedia including text, still images, audio, and video. It also incorporates a high degree of learner control, where the user can decide his/her own pacing, and even the size of the learning subset, by predetermining the amount and types of aircraft to be used in the session. The key component of this particular software application is Axia's patented Discrimination Learning Engine or DLE.

> "...the Discrimination Learning Engine supports multi-dimensional comparison of a set of Items by using side-by-side windows, with a set of Aspect buttons between them. These Aspects are multiple representations of the Item set, while multiple Instances of an Item-

Aspect, with secondary voice-over audio, labels and test, provide levels of detail which are also a powerful scaffolding mechanism." (Chang, Katz, & Wylie, 1997, p.2).

When using the DLE in Challenge Mode, the learner is shown an aircraft (either still picture, or video with audio), and then is asked to identify the aircraft. If the user makes the correct choice then the user is informed that he or she has made the correct identification. This is not necessarily different from traditional training methods, but when a user performs a misidentification, the program displays, in a side by side window (see Figure 2.1) both the aircraft to be identified, and the aircraft believed to be correct.

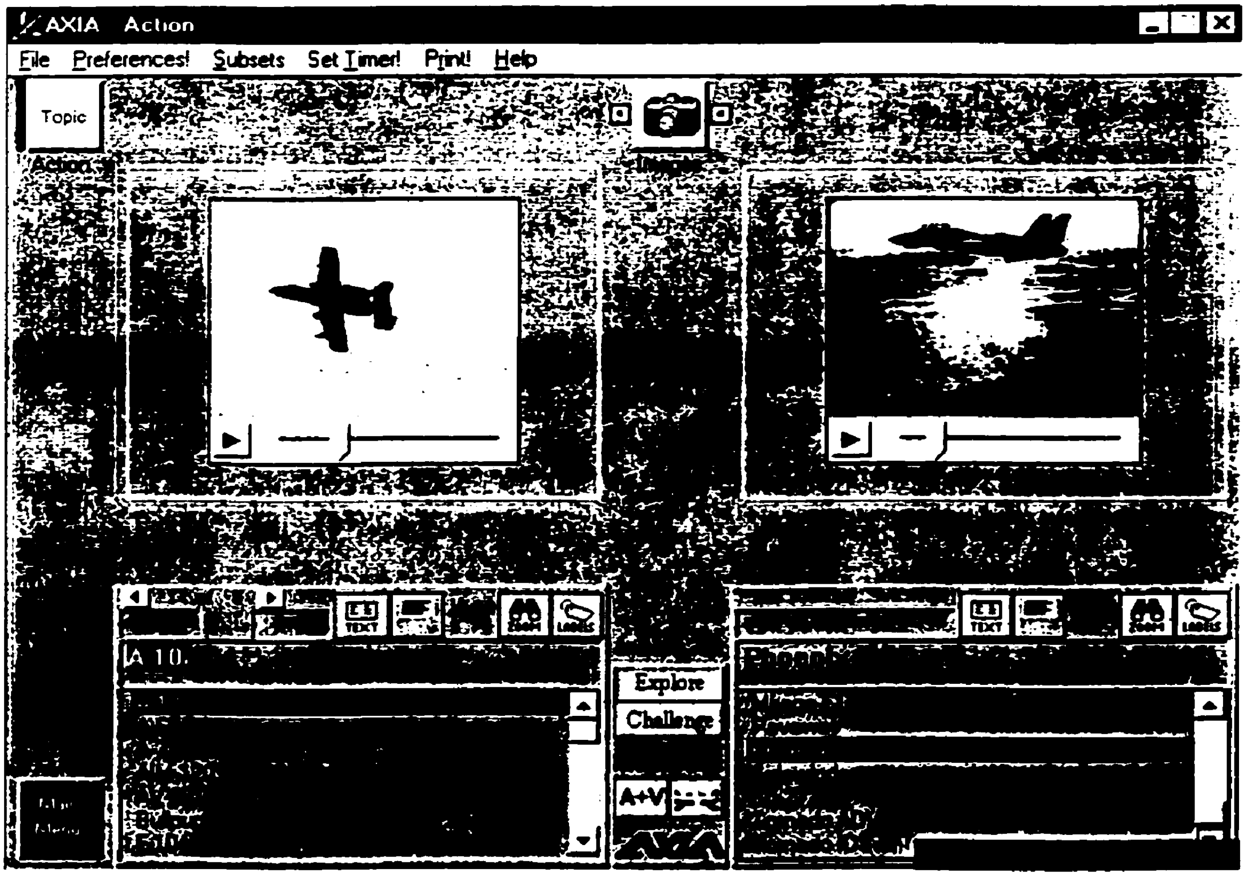


Figure 2.1 : Front-end Interface for *Know Your Combat Jets*.

The user is then allowed to view various aspects of each aircraft such as silhouettes, photographs, performance charts, and video sequences to allow them to identify salient features without which misidentification may still occur. This process of taking an incorrect answer and allowing the user to explore his/her inner cognitive structure regarding the classification task is the basis of the DLE. In support of this model, Scardamalia, Bereiter, McLean, Swallow, & Woodruff (1989) suggested that "Wherever possible the accurate identification of knowledge lacks should have positive consequences within the functioning of the program, resulting in enhanced possibilities for achieving the goals that are motivating the use of the program in the first place." (p.57).

In this study, three modes are used: Challenge mode as described above, Reference Mode, where the user has the same capabilities as Challenge mode, but receives no assistance in the process of answer analysis (the user would have to manually explore the misidentification space), and finally model building in which the learner is restricted from receiving the LCWA provided by the DLE. In all groups each learner also has the capability to restrict the size of the learning subset, and to set time limits on identification tasks.

By restricting the ways in which the user can interact with the DLE, various forms of LC can be analyzed. In Reference Mode, the user can decide upon the current conditions (what he/she does or does not know in relation to the task), the current sub-goal(s), and the method selection (how to use the software to improve his/her knowledge base). The software provides the evaluation of performance and immediate feedback in the form of a yes or no answer. This represents a rather high degree of LC where the learner manages sequence and amount of instruction.

While in Challenge Mode, the user determines his/her current state of knowledge (has the option of removing him/herself from Challenge Mode), but the goals, and methods are determined by the software, with the resulting evaluation process being co-determined by both the software which provides support for the analysis of response, and the user who determines whether multiple aspect ratio investigation is required. Essentially this means that the software provides the user with the opportunity to explore his/her knowledge about a specific concept. This model could be considered as a variant of the model suggested in Goforth's analysis, where a form of LCWA is introduced in the final evaluation stage of the learning cycle. In this form of advisement, the software supplies the user with misidentification information, which the user has the option of acting upon.

Some additional components of the *Know Your Combat Jets* (KYCJ) interface that are useful for assisting the user are the ability to limit the learning subset (incremental learning), the ability to view labels that indicate salient recognition features, and the timing feature that forces the user to respond to questions within a user-specified time frame.

## Conclusion

From the literature review on learner control in CBI it appears that there exists a need for subject-specific learner control research to help understand differences due to content, and purpose (Tennyson & Park, 1985; Shin, Schallert, & Savenye, 1994; Fitzgerald, 1995; Lee & Lee, 1991; and Goforth, 1994). Determining the effect that personality may have upon performance may also provide useful information for object recognition training software design. Many studies have shown the limited way in which users take advantage of the opportunities provided by learner control. The need for effective recognition training and the capabilities of multimedia software to enhance instruction have also been noted. *Know Your Combat Jets* has the potential to be an extremely useful tool both in aircraft recognition training and LC research. In an event-driven program such as *Know Your Combat Jets* (KYCJ), the user has a high degree of control over content sequence, amount, and pacing. By restricting how the user interacts with the DLE it is possible to analyze the effects of these varying types of learner control. Subsequent analysis of this approach to LC can possibly provide future direction for recognition software design and LC research.

With these issues in mind, this study represents an analysis of *Know Your Combat Jets*, with a particular focus upon learner control issues within the software environment using participants who have a need to know the information, which should enhance motivation. Three forms of LC are examined. The first form of LC is the Model Building group where the user has total control over the software determining the current, sub-goal and method states, but does not use the evaluation states and does not receive LCWA. In the second group (Reference Learning) the software controls the sub-goal, method, and evaluation states, but does not provide the user with LCWA related to performance. The third group, Comparative Feedback, is identical to the Challenge group, except that LCWA (comparison feedback) is provided. The three groups are provided with identical software features until a decision is made by the user to enter Challenge mode. Personality effects in relation to performance are also investigated. Finally, a qualitative analysis of user rated effectiveness and enjoyment is performed.

# CHAPTER 3

# METHODOLOGY

This chapter describes the sample population and the materials used in the experiment. Also the protocols given to the participants are described. The chapter concludes with an explanation of the data collection procedures, the research design, and the statistical analysis.

## Participants

This study involved the analysis of the software program *Know Your Combat Jets* (Axia International Inc., 1994). Therefore, the participants for this study came from a population consisting entirely of soldiers belonging to the Canadian Department of National Defense (DND). Members of three Canadian Forces Bases, Base Calgary, Base Suffield, and the Lethbridge Air Defense Regiment, were asked to participate. The soldiers included in this study all belong to the Combat Arms branch of the DND, which indicates that they could very possibly be involved in combat on the battlefield. The members were given no incentive to participate other than to allow them an opportunity to engage in some relatively new aircraft recognition training. The material studied in the experiment focussed upon current potential threat – all aircraft in the study were potential Bosnian Conflict participants. It was hoped that this would also increase the motivational levels of participants during the study. Participants were also each given $5.00 for a light meal upon completion of the study due to the travel requirements for participation.

In all, 56 military personnel volunteered for the study. The mean age of the participants was 29 years, with a minimum age of 19 and a maximum of 41. Fifty-one of the participants were male and 5 were female.

Participants were randomly assigned to one of three groups, Model Building, Reference, and Comparative Feedback. The distribution of participants is shown in Table 3.1:

| Group | Number of Participants | Number of Males | Number of Females |
|---|---|---|---|
| Model Building | 19 | 18 | 1 |
| Reference Learning | 18 | 17 | 1 |
| Comparative Feedback | 19 | 16 | 3 |

Table 3.1 : Participant Distribution by Gender.

Description of Groups

Each of the three groups were allowed various functionality within the software program, *Know Your Combat Jets*. The amount and type of Learner Control varied for each group and thus represents the main independent variable in the experiment. The Model Building group was not allowed to use the challenge mode button, but could interact with all other aspects of the software. The Reference Learning group was allowed to use the challenge mode, but in a restricted fashion. They could only use challenge mode to have the software pose a question, and then they would be required to make their guess and automatically press the answer button to determine whether they were correct or not. This approach would prevent the reference learning group from receiving computer generated advisement related to any misidentification. The Comparative Feedback Group had no restrictions while using the software and they were shown how to use the Discrimination Learning Engine (DLE) to help discriminate between different stimuli. This was accomplished by showing participants what happened when a misidentification occurred and then the steps one could take to attempt to understand the source of the misidentification.

Due to the limited availability of video equipment and the scheduling process, the maximum number of participants in the lab at any given time varied between five and eight.

**Materials**

Eysenck Personality Inventory (EPI) (Form A). The EPI measures two independent dimensions of personality: 1) extraversion-introversion; and 2) neuroticism-stability (Eysenck, & Eysenck, 1968). Each of these traits is measured by 24 questions to which the respondent must answer either yes or no. A response distortion (Lie Scale) is also included to detect attempts to falsify responses. In total, the participants must answer 57 questions to complete the inventory. Empirical data strongly supports the view that introversion and extraversion are primary measurable components of individual differences (The Personality and Individual Differences Study Guide to Personality: An Integrative Approach, http://www-hcs.derby.ac.uk/psychlogy/pidstud.htm). The rationale for the use of the EPI is to demonstrate clearly the relationship between the sample and the population from whence it came. This can be accomplished through a comparison of well established means and the means on the EPI for the participants in this study. Test-retest and internal consistency reliability scores were found to be between .84 and .94 for subjects tested and re-tested one-year later (Eysenck & Eysenck, 1968).

Computers. Each participant used both the instructional software and took the online test on a Pentium 100 MHz personal computer. Each computer had a 15 inch multi-sync monitor set at a resolution of 640x480 pixels, with a color-depth of 16 bits-per-pixel. Each computer also had a 12 X Speed internal CD-ROM drive. Input devices included a standard two-button PS/2 style mouse, and a 101- key enhanced keyboard.

Video Cameras. A total of eight different cameras were used (due to the lack of materials, each camera was a different make and model), and the recording media consisted of VHS, Beta, and 8mm tape. The 8mm tape results were then copied onto either VHS or Beta format for later viewing convenience. All cameras were placed upon tripods and positioned so that the entire computer screen could be seen, but not the participant.

"*Know Your Combat Jets*," (Axia International Inc., 1994). This program was used for the content acquisition (learning) phase of the experiment. This software program allows the user to interact with a database (stored on CD-ROM) consisting of a series of multimedia still images, video, sound, and text. The program utilizes a special interface that attempts to enhance the visual discrimination learning process. The general interface consists of a set of side-by-side windows that contain one of a series of multimedia information datasets (see figure 3.1 below).

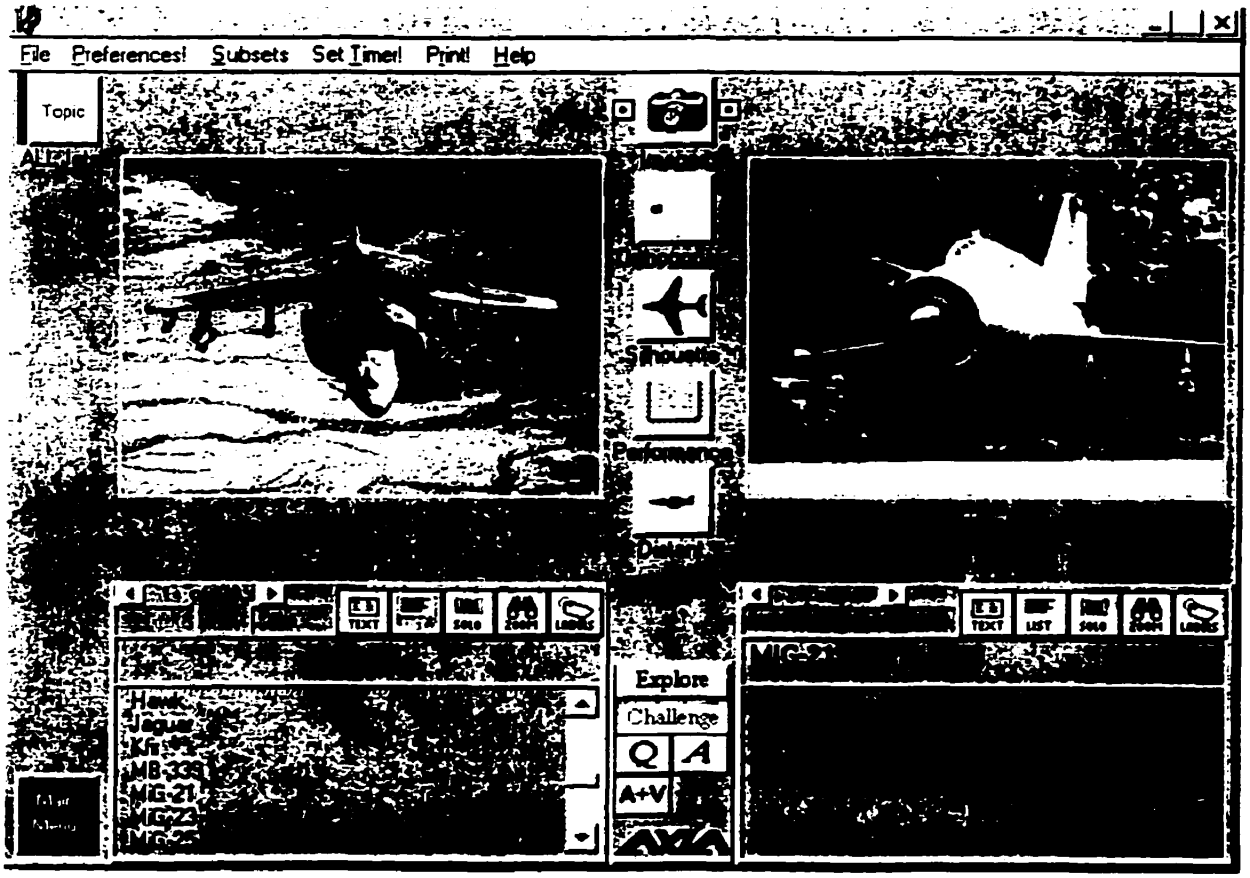


Figure 3.1 : Main Interface Screen for *Know Your Combat Jets* Software.

Aircraft Recognition Test (online test). Due to the need for a testing module, the author designed and programmed the testing software. The software was written in Microsoft Visual Basic (Microsoft, 1996) as this software allows for rapid prototyping with a graphical user interface (GUI). The testing software would output user responses to file for data collection. Users were also shown a progress meter to indicate how close to completion there are while taking the test. Screenshots of the user interface are shown in Figures 3.2 through 3.4. Upon launching the testing software the user would first see the screen in Figure 3.2 below.

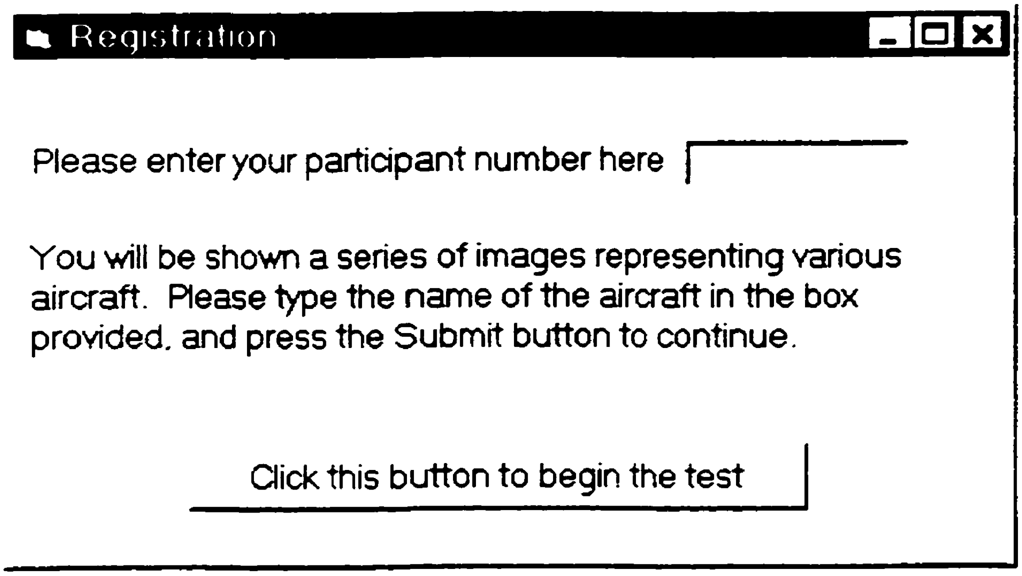


Figure 3.2 : Opening Screen for the Testing Software

If the participant failed to enter a participant coded number, and then attempted to proceed with the test he/she would see the following dialog box (Figure 3.3).

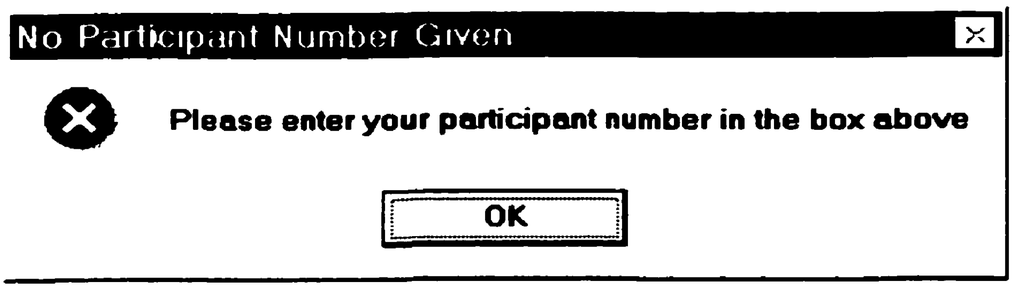


Figure 3.3 : No Participant Number Entered Dialog Box.

The actual test screen can be seen in Figure 3.4 below (not shown to scale).

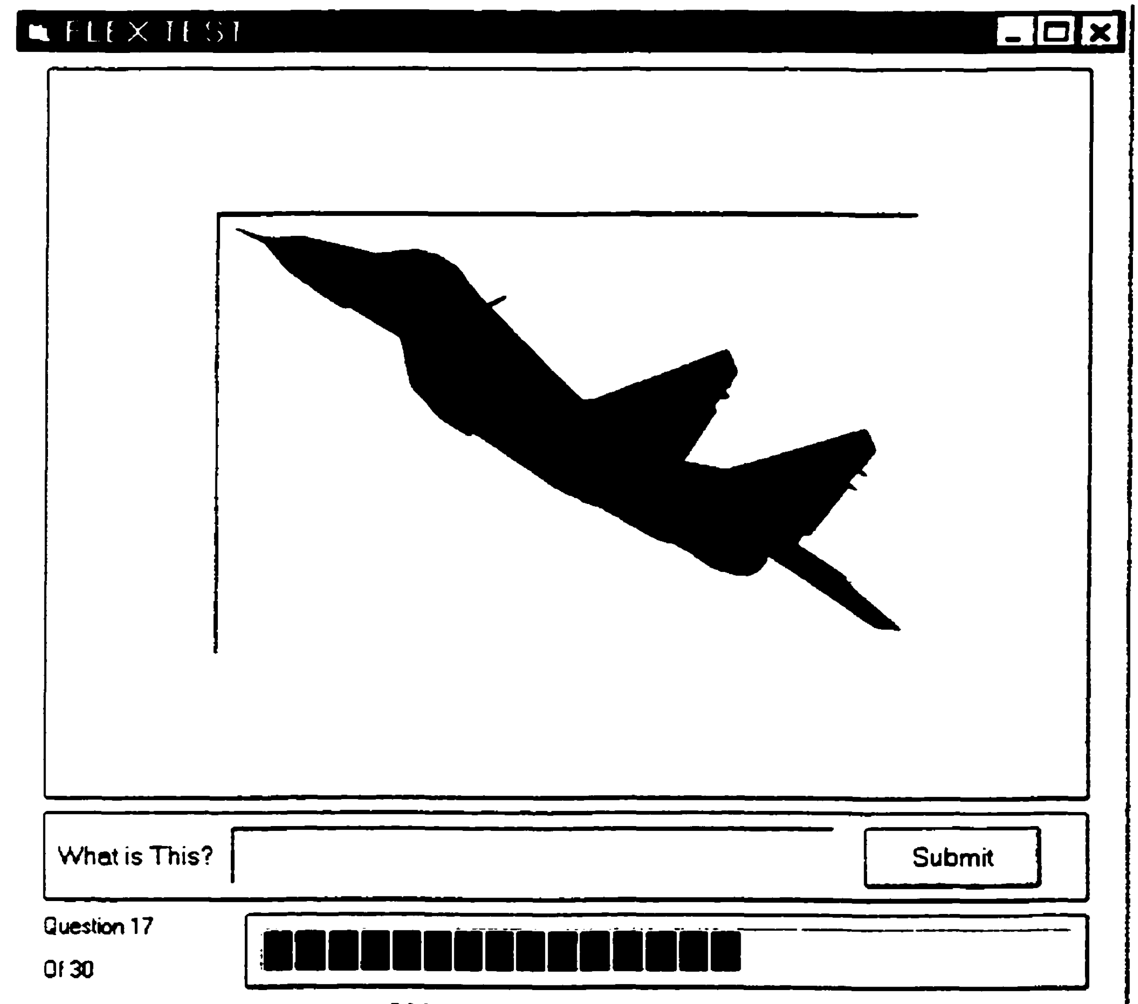


Figure 3.4 : On-line Test Interface.

The user was shown a progress meter and was required to type the answer in the "What is This?" box and then either click submit or press the enter key on the keyboard.

Questions 1 through 10 show actual images, questions 11 through 20 show silhouettes with the original image background intact, and Questions 21 through 30 show only the silhouette with a blue background, taken from the original instructional software. All questions were randomly presented (within their categories) to each participant to reduce order effects. The program also provided participants with a progress meter on the bottom of the screen to indicate how close they were to completion. The software also kept track of the amount of time taken to complete each question. Upon completion of the testing software, the Participants were shown the dialog box presented in Figure 3.5.

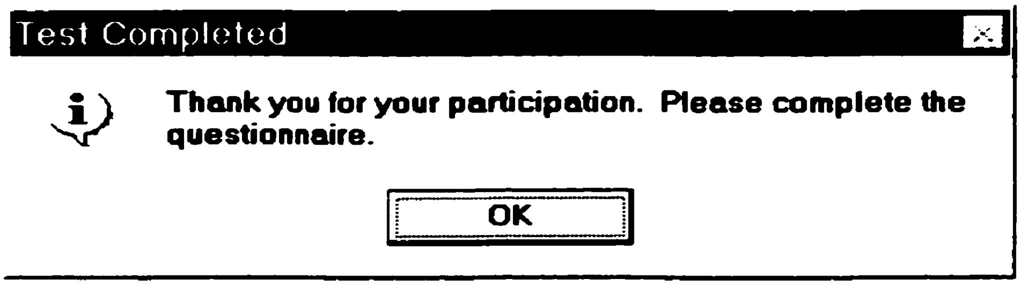


Figure 3.5 : Test Completion Dialog Box.

For an example of the output provided by the testing program for the subsequent analysis, see Appendix C.

Video Data Collection Software. Due to the large amount of data collected from videotaping participants, an economic, accurate, and efficient method was needed for transcribing the video data into numerical results. The author also wrote another program in Visual Basic to accommodate this need. A program which mimicked the

*Know Your Combat Jets* main screen was designed to allow the data entry operator to observe the video record and then transcribe the data to a computer file. Performance activities such as objects clicked, sections viewed, and the summary of total clicks were then automatically written to the local hard drive in the form of a text file for future analysis. The data recording screen can be seen in Figure 3.6 (not to scale).

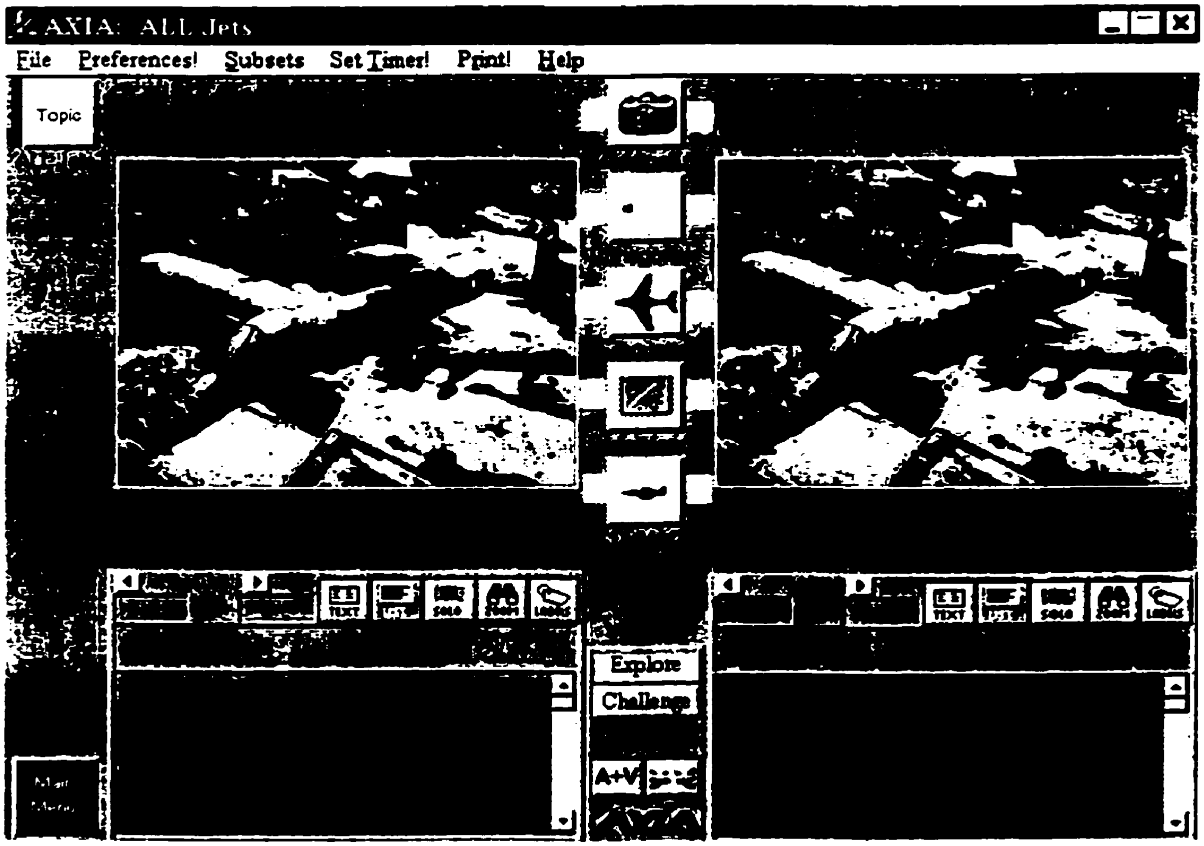


Figure 3.6 : Data Collection Screen.

The major difference between this screen and the *Know Your Combat Jets* screen are the boxes to the left and right of the optional viewing buttons which simply assist the data recorder by providing him/her with a spot to click the mouse to respond to user input. It was felt that this would increase accuracy and inter-rater reliability. The end result is that the data collector only needed to observe the video screen to watch for the participants mouse pointer and then mimic their clicks. This would result in an output file that records all objects clicked and the sequence of clicks. The data entry software also maintains a summary of total clicks on each object. Due to the increased probability of timing errors, the software did not attempt to record the amount of time taken between clicks. See Appendix D for an example of the information that was output to a file for data collection purposes.

Questionnaire. This questionnaire was designed to collect participant demographics and feelings about the overall learning process. Age, sex, education level, military trade, previous computer experience, and previous recognition training were the demographic variables. Two 5-point Likert scales were used to determine: (1) how effective the participants felt the software was for training purposes, and (2) their level of enjoyment in using this method over traditional means. A copy of the participant questionnaire is provided in Appendix F.

Data Analysis Software. The statistical software package SPSS for Windows (SPSS Inc., 1998) was used to perform the data analysis.

**Procedure**

All 56 participants were randomly assigned to one of three groups. At any given time there were no less than five and no more than eight participants in the lab. Sessions progressed over a four-week period. Participants were given a consent form and all soldiers asked, agreed to participate. Participants were then instructed to complete the EPI inventory. Next a short tutorial on the basic features of the *Know Your Combat Jets* software program, and on basic Graphical User Interface operating techniques was given. The tutorial involved the use of the following software features: labels, text, multiple representations, comparison, incremental learning, and the image presentation timer. Participants were also shown basic graphical interface functions such as moving the mouse pointer and clicking the mouse button (Refer to Appendix B for a complete list of items presented during the tutorial session). All participants were then told that they would be able to use the software for a maximum of two hours to try and learn to identify the subset of 40 jets contained in the Bosnian recognition scenario (See Appendix E for a complete list of these planes). Then each group was further subdivided into their assigned treatment group and given individualized instruction concerning how they were to interact with the software. During individual instruction, the remaining two groups waited outside of the lab. Finally, all participants in the session returned to the lab. Each participant was then seated in front of a computer with the software loaded and the camera positioned. Participants were asked to avoid moving their chairs as they might obstruct the video camera which was taping the screen. Upon completion of the learning phase, each participant was to raise his/her hand and the researcher would turn off the camera, and launch the testing software. Participants would then take the online test, and fill in the questionnaire.

Participants were also asked to refrain from talking during the session, other than with the researcher. They were also informed that they would be able to take as much time as necessary to complete the on-line test. Once the participant had taken the tutorial, taken the test, and completed the questionnaire, they were given five dollars and left the room. During the video data collection phase, the recorder could only record clicks that caused some interface reaction. In other words, after a click, if some object changed, or some new media was presented, the recorder would record a click on the object that the user clicked on. This implies that the user may have pressed the mouse button many more times than that recorded, but not in a useful fashion such that they were interacting with the KYCJ software (such as clicking on the title bar or some other object that does not respond to mouse clicks).

During the preliminary testing of procedures and equipment it was discovered that occasionally the computers would stop working during the learning phase and would need to be restarted and the software reloaded. The lab supervisor would provide assistance where required if the software crashed while using the tutorial program. There were no computer crashes during the testing phase. All results were encoded and stored in a locked file cabinet. The video data was collected by means of an aide who would watch each video and record the results using the data collection software mentioned previously. Due to the large amount of data within the video segments, and the likelihood of error when inputting video data into binary format, an inter-rater reliability check was performed on the data collection process.

Figure 3.7 describes the protocol in detail.

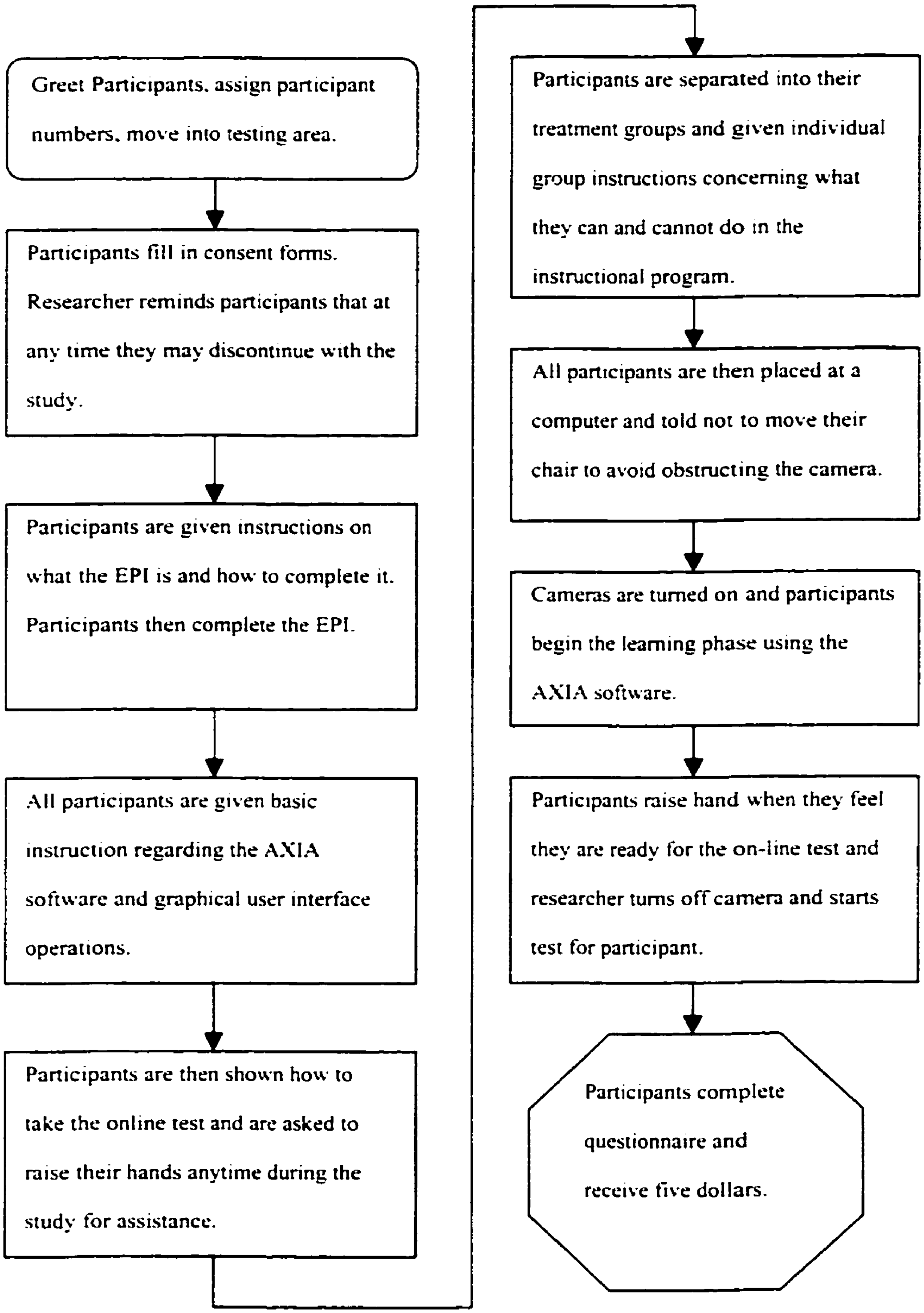


Figure 3.7 : Protocol

**Research Design and Statistical Analysis**

This study utilized a between subjects design with random assignment to reduce individual difference effects. The independent variable was learner control, which had three levels. The purpose of this study was to determine the overall effectiveness of these three different types of learner control when using the *Know Your Combat Jets* software. The performance measures of test score, time on test, and total number of clicks on objects during the instructional phase were chosen as dependent measures of overall effectiveness. With three dependent measures a MANOVA is the appropriate analysis method. However, missing data from the video sequences due to malfunctioning equipment resulted in different sample sizes for each dependent variable. Thus a series of one-way between subjects ANOVA with fixed effects was used to analyze the data.

An investigation of the effects of personality was performed based upon the participants' results on the Eysenck Personality Inventory (EPI). Those participants who demonstrate either high or low scores on the extraversion-introversion scale were classified as either introverts or extraverts. Those participants scoring either high or low on the neuroticism scale were classified as either neurotic or stable. A correlational analysis was performed on participant performance and attitude in relation to either extraversion or introversion. A series of one-way ANOVAs were performed to investigate significant differences, followed by post-hoc tests when necessary.

Two raters entered the data for three different video records (i.e. one soldier's tape from each treatment group) and then a correlation analysis was performed using Spearman's Rho. This is a more conservative test than Pearson's as it first ranks the data before computing the correlational coefficient.

Finally, data from the user questionnaire was analyzed to determine user rated effectiveness and enjoyment of the KYCJ software. The analysis consisted of an ANOVA to explore significant differences across treatments.

It is difficult to provide a firm theoretical basis for the hypotheses in this study, mainly due to the lack of prior research in the combined areas of LC, CBI, and Aircraft Recognition training. Prior research in the area of LC has provided varying results, giving rise to a need for an examination of specific effects of LC models within specific content domains. This exploratory study will hopefully provide researchers with recommendations for future research with this model, and with the design of future models. Based upon the research literature, it was hypothesized that:

Performance Results:

1. The Comparative Feedback group would:
    a. Have significantly higher test scores than the other LC groups.
    b. Spend significantly less time on the test than the other LC groups.

c. Require less mouse clicks in the instructional phase than the other LC groups.

2. The Reference Learning group would:
   a. Have significantly higher test scores than the Model Building group.
   b. Spend significantly less time on the test than the Model Building group.
   c. Require less mouse clicks in the instructional phase than the Model Building group.

Eysenck Personality Inventory

1. Extraverts would use significantly less mouse clicks than introverts.
2. Introverts would outperform extraverts on the on-line test.

Questionnaire Results

1. The majority of users would rate the instructional software as being both effective and enjoyable.
2. Neurotic Introverts would rate the instructional method higher than would Stable Extraverts.

# CHAPTER 4

# RESULTS

Initially, demographic information on the military personnel is described. Then the results of the Eysenck Personality Inventory (EPI) are presented. Next, the results of the participant questionnaire responses are described. Inter-rater reliability and video data results are then provided. Finally the performance data, number of mouse clicks, time on test, and score on test are presented. This section concludes with a presentation of participant behavior while using the interface.

## Demographics

In all, 56 military personnel volunteered for the study. The mean age of the participants was 29 years, with a minimum age of 19 and a maximum of 41. Of the 51 males and 5 females that participated, 14% had not achieved a high school diploma, while 70% had achieved a high school diploma, and 16% had attempted some form of post-secondary studies. 88% of all participants rated themselves as having little to average computer ability. Tables 4.1 and 4.2 provide summary information regarding participant demographics.

| | Mean | Median | Standard Deviation | Range | Minimum | Maximum |
|---|---|---|---|---|---|---|
| Educational Level | 2.02 | 2.00 | .56 | 2 | 1 | 3 |
| Age | 29.34 | 29 | 5.47 | 22 | 19 | 41 |

Table 4.1 : Measures of Central Tendency and Dispersion for Participant Age and Education.

| Valid | Frequency | Percent | Cumulative Percent |
|---|---|---|---|
| 1 | 8 | 14.3 | 14.3 |
| 2 | 39 | 69.6 | 83.9 |
| 3 | 9 | 16.1 | 100.0 |
| Total | 56 | 100.0 | |

Table 4.2 : Participant Education Level (1-no High School Diploma, 2 – High School Diploma, 3 - Post-Secondary attempted).

**Eysenck Personality Inventory Results**

The results of the Eysenck Personality Inventory are shown in Table 4.3.

| Scale | N | Range | Minimum | Maximum | Mean | Standard Deviation |
|---|---|---|---|---|---|---|
| E | 55 | 15 | 4 | 19 | 13.11 | 3.55 |
| N | 55 | 14 | 0 | 14 | 6.84 | 3.60 |
| L | 55 | 7 | 0 | 7 | 3.29 | 1.65 |

Table 4.3 : Measures of Central Tendency for Eysenck Personality Inventory Results (E-extraversion, N-neuroticism, L-Lie scale).

Four participants exceeded the allowable range for the lie-detection scale (L score cut off of 6 or greater) and thus were not included in the summary. One other participant failed to complete the entire inventory and thus the results were also not included in the analysis. The adjusted scores are presented in Table 4.4.

| Scale | N | Range | Minimum | Maximum | Mean | Standard Deviation |
|---|---|---|---|---|---|---|
| E | 51 | 15 | 4 | 19 | 13.10 | 3.63 |
| N | 51 | 14 | 0 | 14 | 7.04 | 3.59 |

Table 4.4 : Adjusted Measures of Central Tendency and Dispersion for Eysenck Personality Inventory Results (E-extraversion, N-neuroticism).

Tables 4.5 and 4.6 show both the extraversion-introversion and neuroticism – stability scale frequencies. Figures 4.1 and 4.2 provide a graphical representation of the same data. Eysenck's Inventory Manual states that scores either below the 31$^{st}$ percentile or above the 70$^{th}$ percentile represent extreme values and thus a score higher than the 70$^{th}$ percentile (15 or higher) on the extraversion scale would imply a highly extraverted personality, while a score less than the 31$^{st}$ percentile (9 or lower) would imply a highly introverted personality. Similarly, on the neuroticism scale, a 13 or greater would imply a highly neurotic personality, while a score of 6 or lower would imply a stable personality.

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| Valid | 4 | 1 | 2.0 | 2.0 | 2.0 |
| | 7 | 1 | 2.0 | 2.0 | 3.9 |
| | 8 | 1 | 2.0 | 2.0 | 5.9 |
| | 9 | 8 | 15.7 | 15.7 | 21.6 |
| | 10 | 2 | 3.9 | 3.9 | 25.5 |
| | 11 | 7 | 13.7 | 13.7 | 39.2 |
| | 12 | 5 | 9.8 | 9.8 | 49.0 |
| | 13 | 1 | 2.0 | 2.0 | 51.0 |
| | 14 | 4 | 7.8 | 7.8 | 58.8 |
| | 15 | 6 | 11.8 | 11.8 | 70.6 |
| | 16 | 5 | 9.8 | 9.8 | 80.4 |
| | 17 | 2 | 3.9 | 3.9 | 84.3 |
| | 18 | 5 | 9.8 | 9.8 | 94.1 |
| | 19 | 3 | 5.9 | 5.9 | 100.0 |
| | Total | 51 | 100.0 | 100.0 | |
| Total | | 51 | 100.0 | | |

Table 4.5 : Extraversion-Introversion Frequencies.

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| Valid | 0 | 1 | 2.0 | 2.0 | 2.0 |
| | 1 | 2 | 3.9 | 3.9 | 5.9 |
| | 2 | 2 | 3.9 | 3.9 | 9.8 |
| | 3 | 3 | 5.9 | 5.9 | 15.7 |
| | 4 | 5 | 9.8 | 9.8 | 25.5 |
| | 5 | 7 | 13.7 | 13.7 | 39.2 |
| | 6 | 6 | 11.8 | 11.8 | 51.0 |
| | 7 | 3 | 5.9 | 5.9 | 56.9 |
| | 8 | 5 | 9.8 | 9.8 | 66.7 |
| | 9 | 1 | 2.0 | 2.0 | 68.6 |
| | 10 | 4 | 7.8 | 7.8 | 76.5 |
| | 11 | 5 | 9.8 | 9.8 | 86.3 |
| | 12 | 4 | 7.8 | 7.8 | 94.1 |
| | 13 | 2 | 3.9 | 3.9 | 98.0 |
| | 14 | 1 | 2.0 | 2.0 | 100.0 |
| | Total | 51 | 100.0 | 100.0 | |
| Total | | 51 | 100.0 | | |

Table 4.6 : Neuroticism – Stability Frequencies.

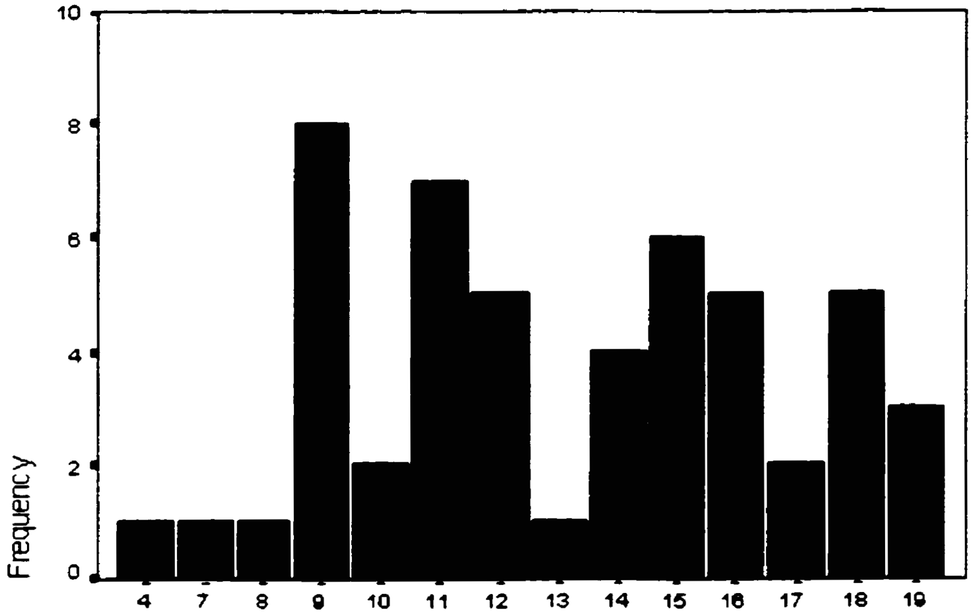


Figure 4.1 : Extraversion – Introversion Frequencies.

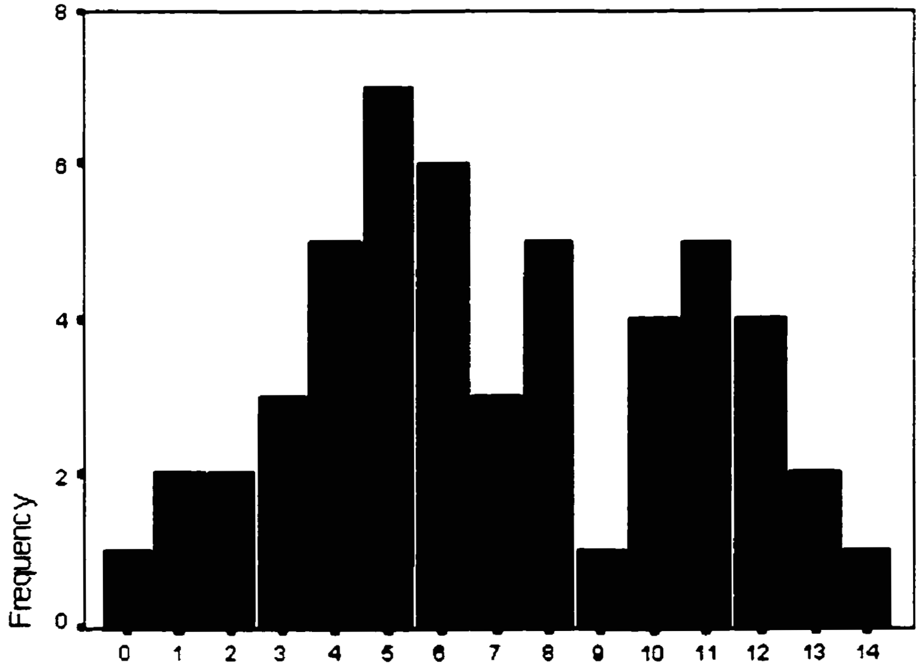


Figure 4.2 : Neuroticism – Stability Frequencies.

Figure 4.3 presents the combined scores of participants on both the Extraversion and Neuroticism scales. The two shaded regions refer to the Neurotic Introverts (NI) and the Stable Extraverts (SE) respectively. The vertical and horizontal lines represent the mean scores for each scale.

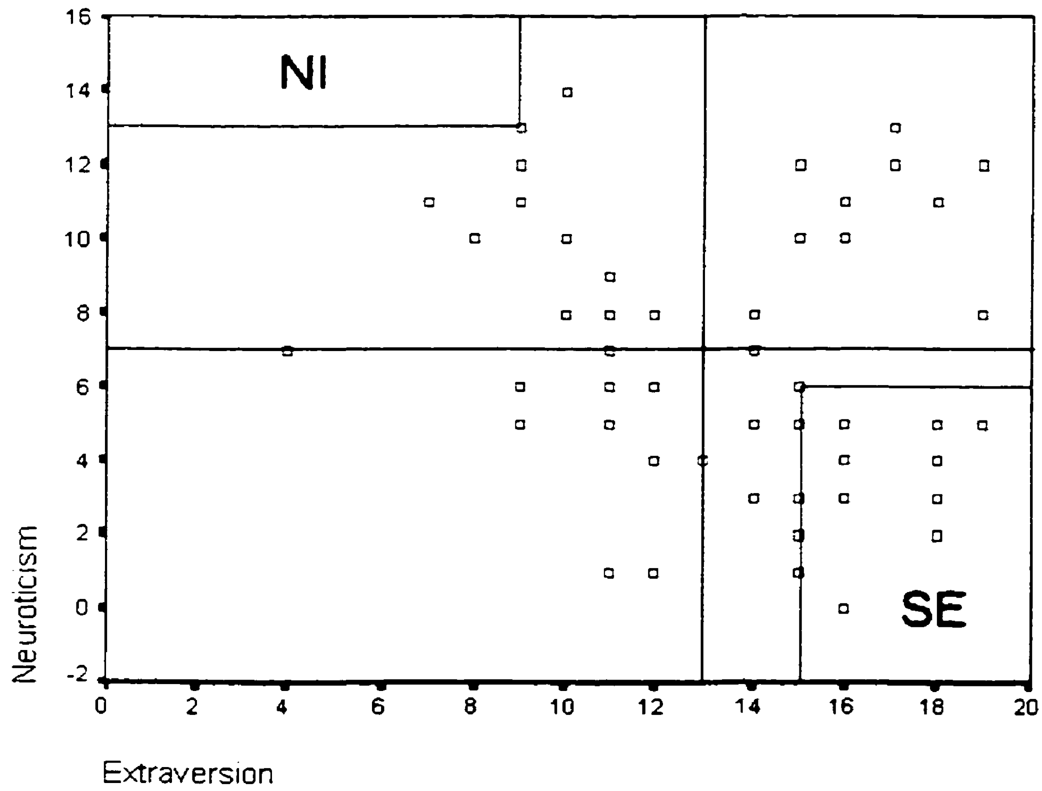


Figure 4.3 : Participants' combined Neuroticism and Extraversion Scores.

Subsequent ANOVAs showed no significant results relating either extraversion or neuroticism to performance (time on test, total clicks, and test score).

None of the participants had a combination of extreme neuroticism and extreme introversion; in other words none of the participants were classified as neurotic introverts. Twelve participants were rated as being stable extraverts (extreme scores on both the extraversion and neuroticism scales). Table 4.7 shows the performance of the "stable extraverts" on the on-line test, total clicks, and time on test in comparison to those participants not labeled as stable extraverts. Stable extraverts did not perform significantly different than non-stable extraverts.

| Stable Extravert | Mean | Standard Deviation |
|---|---|---|
| Total Clicks | 613.33 | 325.04 |
| Time on Test | 362.33 | 120.13 |
| Test Score | 9.33 | 4.66 |
| Non-Stable Extravert | | |
| Total Clicks | 800.15 | 338.77 |
| Time on Test | 444.59 | 195.68 |
| Test Score | 7.72 | 4.54 |

Table 4.7 : Stable Extravert and non-Stable Extravert Performance Data Comparison.

**Questionnaire Results**

Participants rated themselves according to how much previous military recognition training and previous computer experience they had (1-none, 5-extensive). The results are shown in Tables 4.8 and 4.9.

| | Minimum | Maximum | Mean | Standard Deviation |
|---|---|---|---|---|
| Recognition Training Experience | 1 | 5 | 3.02 | 1.10 |
| Computer Experience | 1 | 5 | 2.30 | 0.99 |

Table 4.8 : Previous Recognition Training and Computer Experience Measures of Central Tendency and Dispersion (1-none, 5-extensive).

| | Computer Experience | | Recognition Training Experience | |
|---|---|---|---|---|
| Rating | Total | Percentage | Total | Percentage |
| 1 | 12 | 21.4 | 3 | 5.4 |
| 2 | 23 | 41.1 | 17 | 30.4 |
| 3 | 14 | 25.0 | 19 | 33.9 |
| 4 | 6 | 10.7 | 10 | 17.9 |
| 5 | 1 | 1.8 | 7 | 12.5 |
| Totals | 56 | 100 | 56 | 100 |

Table 4.9 : Participant Computer and Recognition Training Experience Frequencies (1-none, 5-extensive).

The distribution of ratings between treatment groups follows in Tables 4.10 and 4.11. A subsequent Chi-square analysis found no significant difference between observed and expected frequencies.

| Rating | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Model Building | 4 | 8 | 5 | 1 | 1 |
| Reference Learning | 3 | 8 | 4 | 3 | 0 |
| Comparative Feedback | 5 | 7 | 5 | 2 | 0 |

Table 4.10 : Computer Experience and Treatment Group Cross Tabulation.

| Rating | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Model Building | 1 | 4 | 6 | 5 | 3 |
| Reference Learning | 0 | 9 | 7 | 2 | 0 |
| Comparative Feedback | 2 | 4 | 6 | 3 | 4 |

Table 4.11 : Recognition Training Experience and Treatment Group Cross Tabulation.

Enjoyment and effectiveness. two 5-point Likert Scale questions. are the remaining questionnaire data. Enjoyment and effectiveness of this type of learning method were the actual questions asked. A score of one indicated a poor rating while five indicated an excellent rating (Appendix F). In support of this type of instruction. the means for both questions were above average. The results of the questionnaire are shown in Table 4.12 through Table 4.14.

| | N | Range | Minimum | Maximum | Mean | Standard Deviation |
|---|---|---|---|---|---|---|
| Effectiveness | 56 | 3 | 2 | 5 | 3.86 | .92 |
| Enjoyment | 56 | 4 | 1 | 5 | 3.86 | 1.02 |

Table 4.12 : Questionnaire Results - Effectiveness and Enjoyment Number of Cases. Range. Minimum. Maximum. Mean. and Standard Deviation.

| | Effectiveness | | Enjoyment | |
|---|---|---|---|---|
| Rating | Total | Percentage | Total | Percentage |
| 1 | 0 | 0 | 1 | 1.8 |
| 2 | 4 | 7.1 | 5 | 8.9 |
| 3 | 16 | 28.6 | 12 | 21.4 |
| 4 | 20 | 35.7 | 21 | 37.5 |
| 5 | 16 | 28.6 | 17 | 30.4 |
| Totals | 56 | 100 | 56 | 100 |

Table 4.13 : Effectiveness and Enjoyment Ratings by Scale

| Treatment | | Effectiveness | Enjoyment |
|---|---|---|---|
| Model Building | Mean | 4.16 | 4.16 |
| | N | 19 | 19 |
| | Standard Deviation | 0.90 | 0.90 |
| Reference Learning | Mean | 3.72 | 3.67 |
| | N | 18 | 18 |
| | Standard Deviation | .96 | 1.24 |
| Comparative | Mean | 3.68 | 3.74 |
| Feedback | N | 19 | 19 |
| | Standard Deviation | .89 | .87 |

Table 4.14 : Questionnaire Results - Effectiveness and Enjoyment by Treatment Group – Mean and Standard Deviation.

As can be seen in Table 4.14, it appears that the Model Building group had the highest self-rated level of both effectiveness and enjoyment, but no significant differences between groups were found on these two variables (Effectiveness: $F(2,53)=1.57$; Enjoyment: $F(2,53)=1.29$). A strong positive correlation was found between enjoyment and effectiveness (rho = 0.695, $p<0.01$).

**Inter-rater Reliability**

Two raters entered the data for three different video records (i.e. one soldier's tape from each treatment group) and then a correlation analysis was performed using Spearman's Rho. Even though this is a more conservative test than Pearson's as it first ranks the data before computing the correlational coefficient, the results indicated a strong positive (nearly perfect positive) correlation. The results can be seen in Table 4.15.

| | Video One | Video Two | Video Three |
|---|---|---|---|
| Correlation Coefficient | 0.942 ** | 1.000 ** | 0.928 ** |
| ** p<0.01 | | | |

Table 4.15 : Inter Rater Reliability correlation using Spearman's Rho.

Malfunctioning video equipment restricted the amount of data gathered during the video data collection process. It was discovered that fourteen of the fifty-six videos were either blank or unclear. As a result, seven participants from group one, two from group two, and five from group three, had missing video data. This effectively reduced the number of cases for video analysis to 42, but did not affect online test results as this data was recorded into a computer-generated file during testing.

**Performance Results**

Performance variable descriptive totals are provided in Table 4.16.

| | N | Minimum | Maximum | Mean | Standard Deviation |
|---|---|---|---|---|---|
| Test Score | 56 | 1 | 17 | 8.09 | 4.59 |
| Time on Test (seconds) | 56 | 179.00 | 1000.00 | 426.45 | 178.76 |
| Total Clicks | 42 | 143 | 1982 | 733.52 | 327.15 |

Table 4.16 : Measures of central tendency and dispersion for Test Score, Time on Test, and Total Clicks for all participants.

The total number of correct answers out of twenty questions on the online test was recorded. Each participant actually attempted thirty questions on the on-line test, but due to an error in the presentation algorithm that randomly presented questions to the user, questions 21 through 30 had to be eliminated. The testing program presented the questions to the user in the normal random fashion, but the results were incorrectly assigned to a data array, which meant incorrect scoring had occurred. This algorithmic error only occurred in questions 21 through 30, so questions 1 through 20 were scored correctly. Test questions 21 through 30 involved the random presentation of silhouette only images, with blue backgrounds. These were considered the most difficult questions as each picture presented the least amount of visual information as compared to the other questions on the test.

The total time taken (in seconds) to complete the online test was also recorded. It is important to note that this time includes the time to complete all questions, including 21 through 30, since it is still a valid performance variable, regardless of whether the answers given were coded correctly. Time in the learning task would also have been a useful performance variable to analyze, but due to both video and computer problems, the data could not be reliably collected.

The total number of clicks made by each participant during the learning phase using the *Know Your Combat Jets* (KYCJ) software was assessed by reviewing the video.

Next, the same set of dependent variables is categorized by treatment (Table 4.17).

| Treatment Group | | Test Scores | Time on Test (seconds) | Total Clicks |
|---|---|---|---|---|
| Model Building | Mean | 8.16 | 511.00 | 677.33 |
| | N | 19 | 19 | 12 |
| | Standard Deviation | 4.07 | 232.82 | 239.20 |
| Reference Learning | Mean | 6.11 | 396.44 | 795.81 |
| | N | 18 | 18 | 16 |
| | Standard Deviation | 4.48 | 123.70 | 447.91 |
| Comparative Feedback | Mean | 9.89 | 370.32 | 710.50 |
| | N | 19 | 19 | 14 |
| | Standard Deviation | 4.63 | 131.19 | 222.24 |

Table 4.17 : Mean, number of cases, and standard deviation of Test Score, Time on Test, and Total clicks by treatment group.

No significant correlation was found for either effectiveness or enjoyment when compared to total clicks or aircraft recognition test scores. Also, no significant correlation was found between effectiveness and time on test, however, a small positive correlation was found to exist between enjoyment and time spent completing the recognition test (rho = 0.282, $p<0.05$).

To test the significance of differences between treatments and prior hypotheses, a series of one-way ANOVAs were performed on each dependent performance variable. The results are summarized in Table 4.18.

| Performance Variable | | Sum of Squares | df | Mean Square | F | Probability |
|---|---|---|---|---|---|---|
| Test Score | Between Groups | 132.46 | 2 | 66.23 | 3.421 | 0.040 * |
| | Within Groups | 1026.09 | 53 | 19.36 | | |
| | Total | 1158.55 | 55 | | | |
| Time on Test (seconds) | Between Groups | 211901 | 2 | 105951 | 3.633 | 0.033 * |
| | Within Groups | 1545659 | 53 | 29163.4 | | |
| | Total | 1757560 | 55 | | | |
| Total Clicks | Between Groups | 107388 | 2 | 53693.9 | 0.489 | 0.617 |
| | Within Groups | 4280761 | 39 | 109763 | | |
| | Total | 4388148 | 41 | | | |

* $p<0.05$

Table 4.18 : One-way ANOVA of Test Score, Time on Test, and Total Clicks.

Both time on test and test score analysis revealed significant differences ($F(2,53)=3.63$, $p<0.05$; and $F(2,53)=3.42$, $p<0.05$) and thus were analyzed with post hoc Scheffe tests.

Post hoc Scheffe tests (Table 4.19) showed significant differences between the Model Building and Comparative Feedback group means on the amount of time taken to complete the on-line test.

| Dependent Variable | Treatment | | Mean Difference (F) | Std. Error | Prob. | 95% Confidence Interval Lower Bound | Upper Bound |
|---|---|---|---|---|---|---|---|
| Test Score | 1 | 2 | 2.05 | 1.447 | .375 | -1.60 | 5.69 |
| | | 3 | -1.74 | 1.428 | .482 | -5.33 | 1.86 |
| | 2 | 1 | -2.05 | 1.447 | .375 | -5.69 | 1.60 |
| | | 3 | -3.78 | 1.447 | .040* | -7.43 | -.14 |
| | 3 | 1 | 1.74 | 1.428 | .482 | -1.86 | 5.33 |
| | | 2 | 3.78 | 1.447 | .040* | .14 | 7.43 |
| Time on Test (seconds) | 1 | 2 | 114.56 | 56.17 | .135 | -26.91 | 256.03 |
| | | 3 | 140.68 | 55.41 | .048* | 1.14 | 280.23 |
| | 2 | 1 | -114.56 | 56.17 | .135 | -256.03 | 26.91 |
| | | 3 | 26.13 | 56.17 | .898 | -115.34 | 167.60 |
| | 3 | 1 | -140.68 | 55.41 | .048* | -280.23 | -1.14 |
| | | 2 | -26.13 | 56.17 | .898 | -167.60 | 115.34 |
| Total Clicks | 1 | 2 | -118.48 | 126.52 | .648 | -440.45 | 203.49 |
| | | 3 | -33.17 | 130.34 | .968 | -364.85 | 298.51 |
| | 2 | 1 | 118.48 | 126.52 | .648 | -203.49 | 440.45 |
| | | 3 | 85.31 | 121.25 | .782 | -223.24 | 393.86 |
| | 3 | 1 | 33.17 | 130.34 | .968 | -298.51 | 364.85 |
| | | 2 | -85.31 | 121.25 | .782 | -393.86 | 223.24 |

* $p < 0.05$

Table 4.19 : Performance Variables Post-hoc Scheffe Multiple Comparison Results (1-Model Building, 2-Reference Learning, 3-Comparative Feedback).

The post-hoc analysis also showed that the Comparative Feedback group performed significantly better on the test than the Reference Learning group (see Figure 4.19). The Comparative Feedback group did not perform significantly better than the Model Building group. No significant differences were found for total number of mouse clicks.

Figure 4.4 provides an alternative representation of test scores by treatment group. In this line chart, participants are rank-ordered by test score within each treatment group.

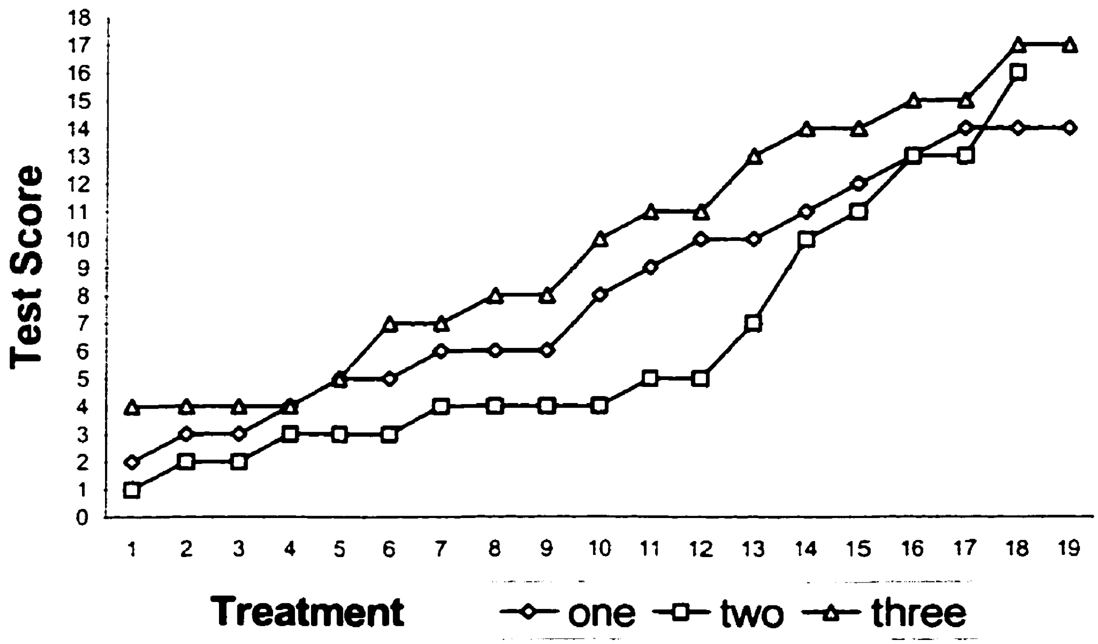


Figure 4.4 : Test Score by Treatment Group (1- Model Building, 2- Reference Learning, 3- Comparative Feedback). Participants are Rank-ordered by Test Score.

**Compliance**

Observations were made during testing and video review to assess compliance to instructions. It is important to consider that users could accidentally click on a forbidden object, which may indicate violation of protocol. The KYCJ interface has many click-able objects, of which some are close to each other on the display. While video data Inter-rater reliability was very highly correlated (0.928-1.000) it is difficult to interpret the purpose of a mouse click, even in the context in which it occurs.

## Model Building

None of the participants in the Model Building group clicked on the Challenge Mode button demonstrating compliance with the restriction protocol.

## Reference Learning

Five of the eighteen participants in the Reference Learning group did click on a plane after being asked a question in Challenge Mode, which contradicted instructions. A visual analysis showed that these users spent on average less than one second in the restricted mode. In fact, the users would frequently follow the violation of protocol with the appropriate response (i.e., clicked the answer button). No attempts were made to take advantage of the exposed features resulting from the violation of protocol, suggesting that these incidents were user errors rather than specific attempts at violating protocol.

## Comparative Feedback

The Comparative Feedback group did not have any restrictions and thus participant behavior was not analyzed for compliance.

## General Interface Usage

In addition to the issue of compliance, it was also useful to consider how well participants utilized the specific features available to their particular treatment group within the learning sequence.

### Label Use

Only 55% of participants used the labels feature, which would allow them to view salient recognition features by placing a text label on the image in the image window. To determine the relative use of this feature, the ratio of label clicks to total clicks was used. Of those who did use the labels feature, that use, on average, only represented 2% of the total clicks, or approximately 17 mouse clicks.

### Multi-aspect Use

The multi-aspect feature allowed the participants (in all treatment groups) to explore the features of all the aircraft in order to facilitate effective classification and categorization. This optional feature is very useful when aircraft share similar features in the learning subset (as is the case with the Bosnian Conflict Scenario subset).

To determine how much use was being made of this feature, a new variable was created by taking the total number of clicks by each participant upon any of the five multi-aspect buttons (fifteen in all when considering the left, right and middle possibilities). This variable was then transformed as a percentage of the total number of clicks by each participant. The average use of this feature by the participants in all groups was only 9% (see Table 4.20).

| Treatment | Statistic | Percent |
|---|---|---|
| Model Building | Mean | 9.08 |
| | N | 12 |
| | Standard Deviation | 7.92 |
| Reference Learning | Mean | 11.94 |
| | N | 16 |
| | Standard Deviation | 14.68 |
| Comparative Feedback | Mean | 6.81 |
| | N | 14 |
| | Standard Deviation | 7.52 |
| Totals | Mean | 9.41 |
| | N | 42 |
| | Standard Deviation | 10.88 |

Table 4.20 : Valid N, Mean, Standard Deviation, and Percentage of Multi-aspect use by Treatment.

Right and Left Interface Use

To determine the percentage of use of each side of the KYCJ interface, another set of variables were created to total all of the left and right clicked objects on the KYCJ main screen. The results are provided in Table 4.21.

| Treatment | Statistic | Left Clicks | Right Clicks |
|---|---|---|---|
| Model Building | Mean | 463.83 | 166.50 |
| | N | 12 | 12 |
| | Standard Deviation | 152.46 | 108.12 |
| Reference Learning | Mean | 464.75 | 104.75 |
| | N | 16 | 16 |
| | Standard Deviation | 321.37 | 94.05 |
| Comparative Feedback | Mean | 508.21 | 83.71 |
| | N | 14 | 14 |
| | Standard Deviation | 138.85 | 75.65 |
| Totals | Mean | 478.98 | 115.38 |
| | N | 42 | 42 |
| | Standard Deviation | 224.89 | 96.64 |

Table 4.21 : Left and Right Side utilization of interface - Number of cases, Mean, and Standard Deviation.

The majority of participants used the left side of the interface screen, even though both sides were visually identical ($t = 10.973$, $d.f. = 41$; $p<0.001$).

## Challenge Mode Use

Use of Challenge Mode represented the major difference between treatments. Model Building was not supposed to use this mode while the Reference Learning group was requested to limit their use of this mode. The Comparative Feedback group was given the opportunity to fully utilize this feature. A new variable was constructed to represent how much use was made of this feature. This variable consists of the sum of questions asked while in Challenge Mode as a ratio of total clicks (in percentage form). The results are presented in Table 4.22 below.

| Treatment | Statistic | Percent |
|---|---|---|
| Reference Learning | Mean | 12.10 |
| | N | 15 |
| | Standard Deviation | 11.18 |
| Comparative Feedback | Mean | 8.74 |
| | N | 13 |
| | Standard Deviation | 6.76 |
| Totals | Mean | 10.54 |
| | N | 28 |
| | Standard Deviation | 9.38 |

Table 4.22 : Percentage Use of Challenge Mode - Number of Cases. Mean. and Standard Deviation.

As can be seen from Table 4.22. very little use was made of Challenge Mode with respect to total clicks.

Additional Features

Participants did not make use of either subset reduction nor timed response features. These features allow the user to restrict the learning subset in order to focus upon a smaller group of planes. which may share very similar features and are more difficult to categorize and classify. The time response feature allows the participant to restrict the amount of time given to respond to computer generated questions.

# CHAPTER 5

# DISCUSSION

In this study, an attempt was made to determine the effectiveness of three different levels of Learner Control while using an aircraft recognition software package called *Know Your Combat Jets* (KYCJ). By limiting the functionality of the user interface it was found that user performance is significantly affected. The personality, questionnaire, and performance results are reviewed and discussed. Learner control issues, study limitations, and considerations for future research are then given. The section concludes with a summary and conclusions of the study.

## Eysenck Personality Inventory

For military personnel, the EPI Administration Guide reported a mean score of 13.2 on the extraversion scale with a standard deviation of 4, and a neuroticism scale mean of 9.1 with a standard deviation of 3.59 (Eysenck & Eysenck, 1968, p.10). These scores are similar to the results for the military participants in this study. The extraversion scale mean was 13.1 with a standard deviation of 3.63 and the neuroticism scale mean was 7.04 with a standard deviation of 3.59. These results suggest that the military personnel in this study have personality scores which match the expected profile, providing support for generalizing the results.

No significant differences were found relating either extraversion or neuroticism to performance. The absence of well defined neurotic introverts prevented any comparison with stable extraverts, and thus the hypothesis of neurotic introverts rating the effectiveness of this training method higher than stable extraverts is not testable with this sample.

Participant Questionnaire

The majority of participants rated the KYCJ program as having above average effectiveness and enjoyment. No significant differences were found between learning methods with regard to effectiveness and enjoyment. These results were surprising with respect to effectiveness considering the low online test scores with a mean score overall of eight correct out of twenty questions or less than fifty percent correct. However, online test results were not provided to participants prior to completing the questionnaire. Had participants received performance feedback from the recognition test results may have differed. Enjoyment was expected to be above average as this method employs multimedia and immediacy of feedback as compared to the slide show method of recognition training. Motivation levels were also expected to be high as the study involved material that was current and relevant.

Performance

The original hypotheses stated that the Comparative Feedback group would outperform both the Reference Learning and Model Building groups on the online test. It was also hypothesized that the Comparative Feedback group would take less time to complete the online test and use fewer mouse clicks during the instructional phase. Finally, it was hypothesized that the Reference Learning group, which had restricted learner control with advisement capability, would outperform the Model Building group requiring less time to complete the test and would use fewer mouse clicks in the training session.

After analysis by ANOVA it was found that the only significant differences on test scores existed between the Comparative Feedback and Reference Learning groups ($F(2,53)=3.42$, $p<0.05$), showing that the Comparative Feedback group performed significantly better than the Reference Learning group. Although it was expected that the Reference Learning group would outperform the Model Building group, there were no significant differences between the two.

A possible explanation for the Reference Learning group's poor performance relative to the other two groups may be due to the restrictions imposed which limited the functionality of Challenge Mode. Steinberg, Baskin, and Hofer (1986) found that when subjected to the use of a limiting software tool, participants became unwilling to make extensive use of it.

In Reference mode, participants were given a question and then were required to make a mental guess and then could choose to look at the correct answer. Although each participant had access to the correct answer, feedback was not provided regarding the nature of the error. Failure feedback can also increase anxiety, which Eysenck (1979) found reduced the effectiveness of memory encoding. Useful information for correctly categorizing and classifying aircraft is not easily identifiable using the Reference Learning mode. The use of unhindered Challenge Mode promotes this capability, which may explain the differences between the Comparative Feedback group and the Reference Learning group.

Test performance scores were better on average for the Comparative Feedback group than the Model Building group, but not significantly. A possible explanation is that the necessary tools for exploring misidentifications still existed and could be used by the Model Building group even though they were restricted from using Challenge Mode. In other words, the functionality of examining the misidentification space was available, but required extra initiative on the part of the learner.

It was hypothesized that the Reference Learning group would score significantly higher on the aircraft recognition test than the Model Building group. This was not the case; group means were higher on the test for the Model Building group than the Reference Learning group. One possible reason for this may be that the Model Building group was not required to receive simple verification (right/wrong) feedback when using the KYCJ software. Schloss, Wisniewski, & Cartwright (1988) stated that "Providing only 'Correct' or 'Wrong' feedback messages fails to provide sufficient informational content to be effective." (p.142).

Overall test performance results indicated an average correct response rate of approximately 41%. A previous study on combat vehicle recognition using traditional training techniques found that as the learning subset increased from five to thirty items, success rates dropped significantly (Smith (1989); as reported in Chang, Katz, & Wylie, 1997). This current study involves a subset of forty items. Success rates may have improved if the subset were reduced as recommended by Harnad (1987). Additional time to learn how to effectively use the KYCJ software could also improve scores; candidates may have used the subset reduction feature if they were more familiar with it's use.

Time on test can be a strong indicator of content mastery and is often inversely related to test score (within acceptable ranges). Although results follow the original hypothesis, significant differences were only found between the Comparative Feedback and Model Building groups ($F(2,53)=3.63$, $p<0.05$). The Comparative Feedback group finished the online test significantly faster than the Model Building group.

The original hypothesis regarding total mouse clicks during the learning phase stated that the Comparative Feedback group would use less clicks than either the Reference Learning or Model Building groups. However, there were no significant differences between groups on total number of mouse clicks. One possible explanation for the greater number of clicks than anticipated for both the Comparative Feedback and Model Building groups could be the use of Challenge Mode. Challenge mode encourages the user to explore the misidentification space, which could involve a series of clicks on the plane lists, multi-aspect buttons, and the question and answer buttons. Although this may be the case, it was shown previously that very little use was made of Challenge mode in relation to total clicks.

Interface Usage

It was quite clear from results that users tended to neglect the right hand side of the interface, tending to click items on the left side of the display. This reduces the effectiveness of the dual window display, which allows users to simultaneously view salient information. As a result, the user will probably require more memory use and is also possibly susceptible to repetition blindness (Bavelier, 1994) while exploring misidentifications.

Prior research has also shown that feedback on cumulative performance is useful in increasing overall learner performance (Schloss, Wisniewski, & Cartwright, 1988). This feature was provided to learners in the restricted subset or incremental learning feature of the KYCJ software, but participants did not use it.

**Learner Control**

In accordance with suggestions by Goforth (1994), Reeves (1993) and Williams (1992), this study examined the effects of using different levels of LC, rather than LC verses PC. The Comparative Feedback group consisted of a form of LC that incorporated LC with advisement (LCWA) so that, while in Challenge Mode, the participant was given the questions, and also provided with misidentification information, as well as an opportunity to explore the misidentification space. Performance results are consistent with previous studies involving LCWA (Arnone & Grabowski, 1992; Williams, 1988) where LCWA was viewed as a more effective form of LC. Unrestricted use of Challenge Mode also provides users with the structure and tools to make more efficient use of their own intelligence and knowledge; a necessary product of effective CBI (Scardamalia et. al., 1989).

Even though the Comparative Feedback group did not perform significantly better than the Model Building group on the aircraft recognition test, they did take significantly less time to complete the recognition test. As there were no significant differences between groups regarding perceived enjoyment and effectiveness levels, the Comparative Feedback form of LC can be considered the most effective in this study.

The overall low levels of achievement on the aircraft recognition test suggests that participants had low initial knowledge levels concerning the content provided in the Bosnian Conflict Scenario. Gay (1986) found that learners with low levels of domain knowledge coupled with a high degree of leaner control often made inappropriate decisions regarding content mastery and sequencing. In the present study, the Model Building group had the highest degree of LC, while the Comparative Feedback group, when in challenge mode, had access to LCWA and program control, which had the potential to assist with content mastery via the DLE

Many previous studies on CBI found that learners tended to not make effective use of available software features that would assist with the learning process (Tennyson & Park, 1985; Arnone & Grabowski, 1991; Poncelet & Proctor, 1993). Consistent with these studies, participants in this study did not take advantage of the more effective features of the KYCJ software. Both the Reference Learning and Comparative Feedback groups made little use of Challenge Mode. Results also found that all participants demonstrated limited use of labels and Multi-aspect views. None of the participants used the subset reduction (incremental learning) or timed response features. One possible cause for this occurrence is the complexity of the interface. Both the subset reduction and time response features involved using menus and required the user to perform more complex actions to enable the particular feature. Participants also had very little time to learn the interface and would probably have benefited greatly from a training session with a set of planes not included in the Bosnian Conflict Scenario, where they could practice using the various KYCJ software features. Time was also a restricting factor, as the number of planes in the subset was rather large in relation to the amount of time given to learn their features. Cronbach and Snow (1977), Reeves (1993), and Relan (1995) suggest sufficient time be allocated to acquaint students with innovative instructional methods and content when using CBI.

**Limitations**

Using cameras to record screen activity created logistical problems. Video required reliance on many machines (8 at a time), and also limited the number of participants at any one time to the number of working video cameras. Participants needed to ensure that they were not blocking the camera while using the software, which could distract them from the learning task. Finally, the loss of video data interfered with the statistical analysis, and also limits the usefulness of the video data. The loss of video data did not invalidate the performance measures for questions 1 – 20 on the online test, but the loss of questions 21 – 30 affects the ability to discuss participant performance for recognizing silhouettes.

An additional problem was the limited time participants had to become familiar with the KYCJ interface, and to learn a subset of 40 aircraft. Participants neglected to use the subset reduction feature, and this probably significantly increased the difficulty of the recognition task.

Treatment restrictions required participants to actively avoid using software features. While there was compliance, the artificial nature of the protocol may have interfered with the learning process for both the Reference Learning and Model Building groups.

While the participants of this study seem to be representative of the military personnel profile based upon the EPI results, there was no random selection of participants from the military population, only random assignment to groups, so generalizability of the results is limited.

Frederico (1989) found that no significant differences existed between paper and computer based recognition training performance. Without some form of direct comparison with traditional methods, it is not possible to compare the effectiveness of the KYCJ software with traditional recognition training methods.

**Future Considerations**

*Know Your Combat Jets* appears to be a useful learning tool and it's DLE could be applied to other knowledge domains. Future training sessions with this software should utilize a smaller learning subset than 40 aircraft. More training with the KYCJ software should also be given, providing the users with detailed explanations of the more powerful features of the software. Participants made little or no use of the learning tools that were built into the KYCJ software. Even with such limited use, significant performance differences were found between treatment groups. Increased levels of feature use have the potential to increase performance effectiveness.

Further investigation regarding the type of learning that is taking place would also possibly provide more generalizable results. As LC attempts to allow the user to utilize their inner cognitive structure, investigations regarding the type of learning task and the resulting effects upon LC could be beneficial. Although beyond the scope of this study, the comparison of task and process learning in relation to learner control in CBI could possibly provide useful information for future CBI design and LC use.

The effects of restricting Challenge Mode and requiring participants to receive correct answer only feedback after being asked a question appears to restrict performance and should be researched further, especially since this is often the type of feedback provided when using CBI (Allessi & Trollip, 1997).

When designing software for the recognition test, care must be taken to validate the test results. An expert in the content domain should take the test, and comment upon the results, along with a careful analysis of the algorithm used to store and score results.

The video data collected in this study has potential for future analysis beyond the scope of this thesis. The video data may provide researchers with an opportunity to analyze path and sequencing decisions and investigate how LC treatments effect navigation methods. An investigation of how participants behaved within the learning phase could produce significant LC and CBI design recommendations.

**Summary and Conclusions**

It was found that the Comparative Feedback group performed the best overall with respect to scores on the aircraft recognition test and time taken to compete the recognition test. Restricting the use of the Challenge Mode function (Reference Learning) appears to interfere with learning. Overall performance on the aircraft recognition test for all treatment groups was lower than expected. This was probably due to three main factors, 1] the learning subset was too large for the amount of time given, 2] insufficient time was provided to learn how to effectively utilize the KYCJ software, and 3] little use was made of the KYCJ's built-in learning tools. Since learning gains were still evident, providing more time, reducing the learning subset, and encouraging greater use of the built in learning tools, have the potential to improve learner performance, especially when comparisons are made with traditional methods. Self reports from the user questionnaire found that 64% of participants found the experience to be effective, and 68% rated the experience as an enjoyable one. A strong positive correlation between effectiveness and enjoyment indicated that the majority of participants who rated the experience as effective also rated it as being enjoyable.

Perhaps the introduction of an adaptive LCWA module could allow specific performance monitoring which would regulate the amount of LC and also the amount and type of advisement that the learner would receive. This could compensate for those learners with low prior knowledge of the learning domain, who require more PC than LC. This may also assist participants to utilize the significant learning aids that are built into the software. If a participant avoids using a particular tool such as subset reduction, and yet continues to demonstrate low recognition scores, the software could provide advice related to the use of the learner subset reduction feature. Consideration of these factors, and further research within the KYCJ model, may result in an extremely effective aircraft recognition training program.

## REFERENCES

A Hybrid Framework for Categorization. Retrieved January 28, 1998 from the World Wide Web: http://www.soton.ac.uk/~coglab/coglab/Thesis/

Alessi, S. M., & Trollip, S. R. (1985). Computer-based instruction. New Jersey: Prentice-Hall.

Allen, G. W. (1990). Learner control of review in computer assisted instruction within a military training environment. (ERIC Document Reproduction Service No. ED 323 913)

Arnone, M. P., & Grabowski, B. L. (1991). Effects of variations in learner control on children's curiosity and learning from interactive video. (ERIC Document Reproduction Service No. ED 334 972)

Arnone, M. P., & Grabowski, B. L. (1992). Effects on children's achievement and curiosity of variations in learner control over an interactive video lesson. Educational Technology, Research and Development, 40(1), 15-27.

Axia International Inc. (1994). Know Your Combat Jets [Computer Program].

Bavelier, D. (1994). Repetition blindness between visually different items: the case of pictures and words. Cognition, 51, 199-236.

Beer, J. M., Gallaway, R. A., & Previc, F. H. (1996). Do individual's visual recognition thresholds predict performance on concurrent attitude control flight tasks? International Journal of Aviation Psychology, 6(3), 273-297.

Boretz, H. F., Brown, S. W., & Kahn, H. (1984). Reaction time, intelligence and learning curve: an exploratory investigation. Ellenville, NY. (ERIC Document Reproduction Service No. ED 255 582)

Bridgeman, B., & Post, T. (1985). Assessing M1 tank commanders with a computerized hand-held tutor. San Diego, CA. (ERIC Document Reproduction Service No. ED 338 627)

Carrier, C. A., & Williams, M. D. (1988). A test of one learner-control strategy with students of differing levels of task persistence. American Educational Research Journal, 25(2), 285-306.

Chang, E. (1996). Learning methodology templates. Unpublished Manuscript.

Chang, E. (1996). Psychological foundations for effective individual learning using multimedia. Unpublished Manuscript.

Chang, E. (1996). The computational simplicity of discrimination learning systems. Unpublished Manuscript.

Chang, E., Katz, L., & Wylie, M. (1997). Discrimination learning: coupling categorization and classification. Proceedings of Ed Media & Ed Telecom '97 Conference, Calgary, AB.

Cho, Y. (1995). Learner control, cognitive processes, and hypertext learning environments. NECC '95 Proceedings, 289-293.

Clark, R. E. (1984). Research on student thought process during computer-based instruction. Journal of Instructional Development, 7(3), 2-5.

Crooks, S. M., Klein, J. D., Jones, E.K., & Dwyer, H. (1996). Effects of cooperative learning and learner control modes in computer-based instruction. Journal of Research on Computing in Education, 29(2), 109-123.

Dempsey, J. V., & Driscoll, M. P. (1989). The effects of four methods of immediate corrective feedback on retention, discrimination error, and feedback study time in computer-based instruction. San Francisco, CA. (ERIC Document Reproduction Service No. ED 309 745)

Duchastel, P. (1990). Assimilatory tools for informal learning: prospects in ICAI. Instructional Science, 19, 3-9.

Ellermann, H. H., & Free, E. L. (1990). A subject-controlled environment for paired associate learning. Journal of Computer-Based Instruction, 17(3), 97-102.

Eysenck, H. J., & Eysenck, B. G. (1968). Eysenck Personality Inventory. San Diego: Educational and Industrial Testing Service.

Eysenck, M. W. (1979). Anxiety, learning, and memory: a reconceptualization. Journal of Research in Personality, 13, 363-385.

Fitzgerald, G. (1995). The effects of an interactive videodisc training program in classroom observation skills used as a teaching tool and as a learning tool. Computers in Human Behavior, 11(3-4), 467-479.

Francis, L., Katz, Y., & Evans, T. (1996). The relationship between personality and attitudes towards computers: an investigation among female undergraduate students in Israel. British Journal of Educational Technology, 27(3), 164-170.

Frederico, P. (1989). Computer-based and paper-based measurement of recognition performance. San Diego, CA. (ERIC Document Reproduction Service No. ED 306 308)

Gay, G. (1986). Interaction of learner control and prior understanding in computer-assisted video instruction. Journal of Educational Psychology, 78(3), 225-227.

Gay, G. (1985). Interaction of learner control and prior conceptual understanding in computer-assisted video instruction. Chicago, IL. (ERIC Document Reproduction Service No. ED 265 845).

Goetzfried, L., & Hannafin, M. J. (1985). The effect of the locus of CAI control strategies on the learning of mathematics rules. American Educational Research Journal, 22(2), 273-278.

Goforth, D. (1994). Learner control = decision making + information: a model and meta-analysis. Journal of Educational Computing Research, 11(1), 1-26.

Goh, D. S., & Moore, C. M. (1977). Personality and academic achievement in three educational levels. New York, NY. (ERIC Document Reproduction Service No. ED 143 907)

Goldstone, R. L. (1994). The role of similarity in categorization: providing a groundwork. Cognition, 52, 125-157.

Gray, S. (1989). The effect of locus of control and sequence control on computerized information retrieval and retention. Journal of Educational Computing Research, 5(4), 459-471.

Harnad, S. (1987). Categorical perception: the groundwork of cognition. New York: Cambridge University Press.

Henry, M. J. (1995). Remedial math students' navigation patterns through hypermedia software. Computers in Human Behavior, 11(3-4), 481-493.

Higginbotham-Wheat, N. (1990). Learner control: when does it work? (ERIC Document Reproduction Service No. ED 323 930)

Hodges, D. L. (1986). How the use of cognitive psychology findings can raise the productivity of computer-assisted instruction. Miami, FL. (ERIC Document Reproduction Service No. ED 295 714)

Hoffman, J. L., & Waters, K. (1982). Some effects of student personality on success with computer-assisted instruction. Educational Technology, 22(3), 20-21.

Katz, Y. J., & Francis, L. J. (1995). Personality, religiosity and computer oriented attitudes among trainee teachers in Israel. Computers in human behavior, 11(1), 1-8.

Kinzie, M. B., Sullivan, H. J., & Berdel, R. L. (1992). Motivational and achievement effects of learner control over content review within CAI. Educational Computing Research, 8(1), 101-114.

Kluger, A. N., & Adler, S. (1993). Person-versus computer-mediated feedback. Computers in Human Behavior, 9, 1-16.

Lee, S., & Lee, Y. H. K. (1991). Effects of learner-control versus program-control strategies on computer-aided learning of chemistry problems: for acquisition or review? Journal of Educational Psychology, 83(4), 491-498.

Mattoon, J. S., & Klein, J. D. (1993). Controlling challenge in instructional simulation. Educational Computing Research, 9(2), 219-235.

Mattoon, J. S., Thurman, R. A., Klein, J. D., & DeVilbiss, C. A. (1991). Effects of control mechanisms in instructional simulation. Chicago, IL. (ERIC Document Reproduction Service No. ED 336 077)

Medin, D. L. (1973). A context-sensitive theory of discrimination learning. New York, NY. (ERIC Document Reproduction Service No. ED 096 573)

Merril, M. D. (1988). Don't bother me with instructional design – I'm busy programming!: suggestions for more effective educational software. Computers in Human Behavior, 4, 37-52.

Microsoft Corporation (1996). Microsoft Visual Basic Programming System for Windows 95 [Computer Program].

Milheim, M. D. (1989). Perceived attitudinal effects of various types of learner control in an interactive video lesson. Dallas, TX. (ERIC Document Reproduction Service No. ED 308 828)

Milheim, W. D., & Azbell, J. W. (1988). How past research on learner control can aid in the design of interactive video materials. New Orleans, LA. (ERIC Document Reproduction Service No. ED 295 652)

Milheim, W. D., & Martin, B. L. (1991). Theoretical bases for the use of learner control: three different perspectives. Journal of Computer-Based Instruction, 18(3), 99-105.

Morrison, G. R., Ross, S. M., & Baldwin, W. (1992). Learner control of context and instructional support in learning elementary school mathematics. Educational Technology, Research and Development, 40(1), 5-13.

Nicholson, P. A. (1992). An investigation of learner characteristics and instructional control on grade five students. Canadian Journal of Educational Communication, 21(3), 223-241.

Niemiec, R. P., Sikorski, C., & Walberg, H. J. (1996). Learner-control effects: a review of reviews and a meta-analysis. Journal of Educational Computing Research, 15(2), 157-174.

Nosofsky, R. M., Gluck, M. A., Palmeri, T. J., McKinley, S. C., & Glauthier, P. (1994). Comparing models of rule-based classification learning: a replication and extension of Shepard, Hovland, and Jenkins (1961). Memory & Cognition, 22(3), 352-369.

Park, O. (1984). Empirically based procedures for designing a response-sensitive sequence in computer-based instruction: an example from concept-teaching strategies. Journal of Computer-Based Instruction, 11(1), 14-18.

Pcius, K. E. (1991). Personality factor sin human-computer interaction: a review of the literature. Computers in Human Behavior, 7, 103-135.

Poncelet, G. M., & Proctor, L. F. (1993). Design and development factors in the production of hypermedia-based courseware. Canadian Journal of Educational Communication, 22(2), 91-111.

Pridemore, D. R., & Klein, J. D. (1992). Effects of learner control over feedback in computer-based instruction. ECT '92 Proceedings, 604-611.

Pridemore, D. R., & Klein, J. D. (1993). Learner control of feedback in a computer lesson. New Orleans, LA. (ERIC Document Reproduction Service No. ED 362 194)

Pridemore, D. R., & Klein, J.D. (1991). Control of feedback in computer-assisted instruction. Educational Technology Research and Development, 39(4), 27-32.

Reeves, T. C. (1993). Pseudoscience in computer-based instruction: the case of learner control research. Computer-Based Instruction, 20(2), 39-46.

Relan, A. (1995). Promoting better choices: effects of strategy training on achievement and choice behavior in learner-controlled CBI. Educational Computing Research, 13(2), 129-149.

Rieber, L. P., Boyce, M. J., & Assah, C. (1989). Effects of computer animated lesson presentations and cognitive practice on adult learning in physical science. Dallas, TX. (ERIC Document Reproduction Service No. ED 308 834)

Ross, S. M., & Morrison, G. R. (1989). In search of a happy medium in instructional technology research: issues concerning external validity, media replications, and learner control. Educational Technology Research and Development, 37(1), 19-33.

Rubincam, I., & Olivier, W. P. (1985). An investigation of limited learner-control options in a CAI mathematics course. AEDS Journal, 18(4), 211-226.

Sales, G. C., & Williams, M. D. (1988). The effect of adaptive control of feedback in computer-based instruction. Journal of Research in Computing in Education, 21(1), 97-111.

Scardamalia, M., Bereiter, C., McLean, R. S., Swallow, J., & Woodruff, E. (1989). Computer-supported intentional learning environments. Journal of Educational Computing Research, 5(1), 51-68.

Schloss, P. J., Sindelar, P. T. Cartwright, G. P., & Smith, M. A. (1988). Learner control over feedback as a variable in computer assisted instruction. Journal of Research in Computing in Education, 20(4), 310-320.

Schloss, P. J., Wisniewski, L. A., & Cartwright, G. P. (1988). The differential effect of learner control and feedback in college students' performance on CAI modules. Journal of Educational Computing Research, 4(2), 141-149.

Schnackenberg, H. L. (1997). Learner control over full and lean computer-based instruction under differing ability levels. Ph.D. Dissertation. (ERIC Document Reproduction Service No. ED 408 935)

Schnackenberg, H. L., & Sullivan, H. J. (1997). Learner ability and learner control in computer assisted instructional programs. Albuquerque, NM. (ERIC Document Reproduction Service No. ED 409 868)

Schyns, P. G., & Rodet, L. (1997). Categorization creates functional features. Journal of Experimental Psychology: Learning, Memory, and Cognition, 23(3), 681-696.

Secrist, G. E., & Hartman, B. O. (1993). Situational awareness: the trainability of the near-threshold information acquisition dimension. Aviation, Space & Environmental Medicine, 64(10), 885-892.

Shin, E. C., Schallert, D. L., & Savenye, W. C. (1994). Effects of learner control, advisement, and prior knowledge on young students' learning in a hypertext environment. Educational Technology, Research and Development, 42(1), 33-46.

Siegel, M. A., & DiBello, L. V. (1980). Optimization of computerized drills: an instructional approach. (ERIC Document Reproduction Service No. ED 207 555)

Siegel, M. A., & Misselt, A. L. (1984). Adaptive feedback and review paradigm for computer-based drills. Journal of Educational Psychology, 76(2), 310-317.

Sims, R. (1997). Interactivity: a forgotten art? Computers in Human Behavior, 13(2), 157-180.

Smith, N. (1989). Summary of research on combat vehicle identification completed at Fort Hood Field Unit. US Army Research Institute for the Behavioral and Social Sciences Research Report 154C, Fort Hood, TX.

Spielberger, C. D. (1970). The effects of anxiety on computer-assisted learning. Salishan, OR. (ERIC Document Reproduction Service No. ED 045 720)

SPSS Inc. (1998). SPSS for Windows v7.0 [Computer Program].

Steinberg, E. R., Baskin, A. B., & Hofer, E. (1986). Organizational/memory tools: a technique for improving problem solving skills. Journal of Educational Computing Research, 2(2), 169-187.

Tallmadge, G. K., Shearer, J. W., & Greenberg, A. M. (1968). Study of training equipment and individual differences: the effects of subject matter variables. Palo Alto, CA. (ERIC Document Reproduction Service No. ED 043 194)

Tarr, M. J. (1995). Rotating objects to recognize them: a case study on the role of viewpoint dependency in the recognition of three-dimensional objects. Psychonomic Bulletin and Review, 2(1), 55-82.

Tennyson, R. D., & Park, S. I. (1985). Interactive effect of process learning time and ability level in concept learning using computer-based instruction. Journal of Structured Learning, 8, 241-260.

Thackray, R. I., Jones, K. N., & Touchstone, R. M. (1974). Personality and physiological correlates of performance decrement on a monotonous task requiring sustained attention. British Journal of Psychology, 65(3), 351-358.

Williams, M. D. (1993). A comprehensive review of learner-control: the role of learner characteristics. New Orleans, LA. (ERIC Document Reproduction Service No. ED 362 211).

Young, J. D. (1996). The effects of self-regulated learning strategies on performance in learner controlled computer-based instruction. Educational Technology Research and Development, 44(2), 17-27.

**Appendix A**

Research Cover Letter

Dear Research Participant

My name is Mark Wylie. I am a graduate student in the Department of Educational Psychology at the University of Calgary, conducting a research project under the supervision of Dr. Larry Katz, as part of the requirements for the course EDPS 692.82. I am writing to provide information regarding my research project on Discrimination Learning with Multimedia so that you can make an informed decision regarding your participation.

The purpose of the study is to examine the effectiveness of a multimedia discrimination learning software program. As part of the study you will be asked to:

- listen to a introductory information session concerning the study
- use a multimedia software program to learn basic facts about aircraft
- take a short test concerning the material in the software
- complete a personal questionnaire

These procedures will take approximately 2 hours. You should be aware that even if you give your permission you are free to withdraw at any time for any reason and without penalty.

Participation in this study will involve no greater risks than those ordinarily experienced in daily life.

Only your computer screen will be video taped to record all procedures used when interacting with the software. These tapes will be secured within a locked filing cabinet and will be erased two years after study completion.

You will complete a questionnaire and short test individually with paper and pencil. Once collected, responses will be kept in strictest confidence, reviewed only by those individuals directly involved with the study. Data will be gathered in such a way to ensure anonymity: each participant will be assigned a number, and all results will be encrypted and stored on a computer for a period of two years. Only group results will be reported in any published studies. The raw data will be kept in a locked file cabinet and destroyed two years after completion of the study.

If you have any questions, please feel free to contact me at (403)220-4269 or mowylie@acs.ucalgary.ca. My supervisor, Dr. Larry Katz, can be reached at (403)220-3418 or katz@acs.ucalgary.ca. You may also contact the Chair, Faculty of Education Joint Ethics Committee at 220-5664, or the Office of the Vice-President (Research) at 220-3881.

Two copies of the consent form are provided. Please return one signed copy to me and retain the other copy for your records.

## CONSENT FOR RESEARCH PARTICIPATION

I, the undersigned, hereby give my consent to participate in a research project entitled Discrimination Learning with Multimedia.

I understand that such consent means that I will take part in a lecture, use a software program, write a short test, and complete a questionnaire. I also understand that my computer screen will be video taped to record all procedures used when interacting with the software.

I understand that participation in this study may be terminated at any time by my request or at the request of the investigator. Participation in this project and/or withdrawal from this project will not adversely affect me in anyway.

I understand that this study will not involve any greater risks than those ordinarily occurring in daily life.

I understand that the responses will be obtained anonymously and kept in the strictest confidence.

I understand that only group data will be reported in any published reports.

I have been given a copy of this consent form for my records. I understand that if I have any questions I can contact the researcher at 220-4269, or his supervisor at 220-3418. I may also contact the Chair, Faculty of Education Joint Ethics Committee at 220-5664, or the Office of the Vice-President (Research) at 220-3881.

______________________________________________

Date                                        Signature

______________________________________________

Participant's Printed Name

# Appendix B

Tutorial given to participants

All groups received the tutorial on using a graphical user interface use. This tutorial was comprised of the following items:

- How to move the mouse pointer
- How to click the mouse on objects
- Closing programs by clicking on the close icon on the title bar
- Dragging the mouse to move objects such as sliders
- Selecting menu items

All groups also received a basic tutorial on how to use the more prominent features of the *Know Your Combat Jets* software. These features included:

- The timing feature
- Subset reduction
- Choosing items from the plane lists
- Move through different media of the same item (look at other pictures of the same item)
- Changing aspects (different information screens for the same aircraft such as orthogonal views)
- Labels

The Model Building group was shown where the Challenge mode button was and was instructed to not click on it. The Reference Learning group was told that they could click on the Challenge mode button but had to follow up questions asked by pressing the answer button rather than choosing a plane from the list. The Comparative Feedback group was given no restrictions on how to use the software.

General Instructions given to all groups:

You will be tested on your ability to identify forty aircraft that could potentially be used in a Bosnian Conflict. You will be given a maximum of two hours with the instructional software and as much time as is needed to complete the online test. Upon completion of the test you will complete a questionnaire. If at any time you require assistance, please raise your hand I will assist you. If you have any questions before we begin please ask them now.

**Appendix C**

Testing program sample output file

```
<><> output begin <><>
"participant number"
"-----"
"Time Begin: ","7/30/96","11:24:24 AM"
"-----"
"Question# ",1
"Photo ","rafale"
"Answer: ","starfighter"
"Time: ",76
"Pressed Enter"
"-----"
<questions #2 through #29 removed for sake of brevity>
"-----"
"Question# ",30
"Photo ","mig23"
"Answer: ",""
"Time: ",12
"Pressed Enter"
"-----"
"Time complete: ","11:36:58 AM"
"-----"
```

**Appendix D**

Sample Data Collection Input File

```
Participant Code
"-----"
"-----"
"Time Begin: ","4/6/97","12:40:59 PM"
"-----"
"-----"
"Click# ",1
"Distant Button"
"Tape Counter: ","00.00"
"-----"
"Click# ",2
"Right Image Window"
"Tape Counter: ","00.08"
"-----"
```

<mouse clicks 3 through 845 removed for sake of brevity>

```
"-----"
"Click# ",846
"Left List Alpha Jet"
"Tape Counter: ","83.52"
"-----"
"Totals "
"Topic ",0
"Menu ",0
"left images ",5
"right images ",1
"images ",0
"left orthogonals ",12
"right orthogonals ",3
"orthogonals ",1
"left silhouette ",9
"silhouette ",0
"right silhouette ",5
"left performance ",0
"right performance ",0
"performance ",0
"left distant ",0
"distant ",1
```

```
"right distant ",0
"left video used ",0
"left video stopped ",1
"right video used ",0
"right video stopped ",0
"left image ",70
"right image ",66
"left left rotate ",0
"left right rotate ",0
"right left rotate ",0
"right right rotate ",0
"left text ",1
"right text ",0
"left list ",357
"right list ",121
"left solo ",1
"right solo ",0
"left zoom ",17
"right zoom ",0
"left labels ",46
"right labels ",0
"left scroll ",1
"right scroll ",1
"left scroll up ",35
"right scroll up ",29
"left scroll down ",40
"right scroll down ",22
"explore ",0
"challenge ",0
"A + V ",0
"Q",1
"A",
"break link",0
"main menu button",0
"Start ","00.00"
"End ","83.52"
comments
one crash was observed at time 28:43
```

## Appendix E

Bosnian conflict aircraft subset

All Jets - Central Europe

A-10
A-6E
A-7
Albatros
Alpha Jet
AMX
AV-8B
EA-6B
F-104
F-111F
F-117A
F-14
F-15A/C
F-15E
F-16A
F-16C/D
F-4E
F-4G
F/A-18
Harrier
Jaguar
MB-339
MiG-21
MiG-23
MiG-25
MiG-27
MiG-29
MiG-31
Mirage 2000C
Mirage F1
Mirage III
Rafale
Sea Harrier
Su-17
Su-24
Su-25
Su-27
Tornado ADV
Tornado IDS
EF-111A

# Appendix F

## Participant Questionnaire

STUDY QUESTIONNAIRE DATE

PLEASE COMPLETE THIS FORM BY PLACING THE REQUESTED INFORMATION IN THE SPACES PROVIDED

PARTICIPANT #

AGE

SEX

MALE FEMALE

HIGHEST LEVEL OF EDUCATION ACHIEVED

GRADE DEGREE(S)

TRADE (OCCUPATION)

PLEASE ANSWER THE FOLLOWING QUESTIONS BY PLACING A CHECK MARK IN THE APPROPRIATE BOX

AMOUNT OF PREVIOUS COMPUTER EXPERIENCE

NONE EXTENSIVE

AMOUNT OF PREVIOUS MILITARY VEHICLE RECOGNITION

NONE EXTENSIVE

HOW WOULD YOU RATE THE EFFECTIVENESS OF THIS TYPE OF LEARNING METHOD?

POOR EXCELLENT

HOW WOULD YOU RATE YOUR ENJOYMENT OF THIS METHOD OF LEARNING?

POOR EXCELLENT