6-1998

A Quasi-Experimental Approach for Assessing Air Traffic Controller Workload

Ana Theresa Borja

Embry-Riddle Aeronautical University - Daytona Beach

Follow this and additional works at: https://commons.erau.edu/db-theses

Part of the Aviation Commons

Scholarly Commons Citation

https://commons.erau.edu/db-theses/300

This thesis is brought to you for free and open access by Embry-Riddle Aeronautical University – Daytona Beach at ERAU Scholarly Commons. It has been accepted for inclusion in the Theses - Daytona Beach collection by an authorized administrator of ERAU Scholarly Commons. For more information, please contact commons@erau.edu.
A QUASI-EXPERIMENTAL APPROACH FOR ASSESSING AIR TRAFFIC CONTROLLER WORKLOAD

by

Ana Theresa Borja

A Thesis Submitted to the
Office of Graduate Programs
in Partial Fulfillment of the Requirements for the Degree of
Master of Aeronautical Science

Embry-Riddle Aeronautical University
Daytona Beach, Florida
June 1998
INFORMATION TO USERS

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

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.
A QUASI-EXPERIMENTAL APPROACH FOR
ASSESSING AIR TRAFFIC CONTROLLER WORKLOAD

by

Ana Theresa Borja

This thesis was prepared under the direction of the candidate’s thesis chair, Dr. Gerald Gibb, Department of Human Factors and Systems, and has been approved by the members of her thesis committee. It was submitted to the Office of Graduate Studies and was accepted in partial fulfillment of the requirements for the degree of Master of Aeronautical Science.

THESIS COMMITTEE:

Dr. Gerald Gibb
Chairman

Dr. John Wise
Member

Dr. Dan Garland
Member

Department Chair, Aeronautical Science

Dean of Graduate Studies, Daytona Beach campus
ACKNOWLEDGEMENTS

The author wishes to express special gratitude to the Thesis Chair, Dr. Gerald Gibb, whose constant encouragement and guidance were crucial to the successful outcome of this thesis. Appreciation is also due to Drs. Wise and Garland, Thesis Committee Members, for their assistance in preparing this manuscript.

This statement of acknowledgment would be incomplete without a formal expression of sincere appreciation and gratitude to the author's family and friends for providing the constant encouragement and love needed to complete the task.
The objective of this thesis was to evaluate and determine the operational impacts to the Oceanic Air Traffic Controller (controller) from deficiencies of an Oceanic Data Link system. These deficiencies in the Oceanic Data Link system are in regards to the Computer Human Interface (CHI) and its effect on the cognitive effort and physical task requirements imposed on the controller. The various workload methodologies and techniques were reviewed for specific workload techniques applicable to the operational environment when resources, such as time and funding, are lacking for a laboratory design. Data was collected from a live oceanic control facility where the Oceanic Data Link system is currently being utilized at a single sector on the control room floor. Qualitative measures were used to assess controller workload associated with performing Air Traffic Control (ATC) tasks. The data collection activities utilized the analysis of data from the NASA-Task Load Index (TLX), observation, and questionnaires. Subjective workload analysis was used and collected from eleven oceanic controllers. Analysis of the NASA-TLX revealed that the use of the Oceanic Data Link system received the highest rating in mental demand and temporal demand followed closely by frustration and effort. The Oceanic Data Link system imposes higher workload in cognitive demand rather than physical demand, but does not affect their performance.
# Table of Contents

ACKNOWLEDGEMENTS ........................................................................................................ III

ABSTRACT ........................................................................................................................... IV

ACRONYMS ......................................................................................................................... IX

INTRODUCTION .................................................................................................................. 1

BACKGROUND
  Oceanic Data Link System .......................................................................................... 2
AIR TRAFFIC CONTROLLER WORKLOAD ...................................................................... 5
PRODUCTIVITY, PERFORMANCE, AND AUTOMATION .................................................. 8
DIFFICULTIES IN MEASURING WORKLOAD .................................................................... 8

OVERVIEW OF EMPIRICAL WORKLOAD ASSESSMENT ................................................ 10

INFLUENCES ON ATC WORKLOAD .................................................................................. 10
  ATC Complexity ............................................................................................................. 12
  Hardware/software ......................................................................................................... 13
  Operator Behavior ......................................................................................................... 13
  Individual Differences .................................................................................................. 15

A SUMMARY OF EMPIRICAL TECHNIQUES .................................................................. 16

PERFORMANCE-BASED
  Primary-task Measurement ........................................................................................... 16
  Secondary Tasking .......................................................................................................... 17

SUBJECTIVE WORKLOAD .............................................................................................. 19

PHYSIOLOGICAL WORKLOAD ....................................................................................... 20

REVIEW OF THE RELATED LITERATURE ...................................................................... 22

METHODOLOGY ............................................................................................................... 26

OVERVIEW
  Preparation ....................................................................................................................... 26
  Assessment Methodology .............................................................................................. 26
  Participants ....................................................................................................................... 27
  Scenario ........................................................................................................................... 29
  Physical Setup ................................................................................................................ 30
  Data Collection Materials ............................................................................................. 30
  Assessment Conduct ....................................................................................................... 31

RESULTS ............................................................................................................................ 33

NASA-TLX .......................................................................................................................... 33
POST-RUN QUESTIONNAIRES ......................................................................................... 37

CONCLUSIONS ................................................................................................................... 42

REFERENCES ...................................................................................................................... 43

APPENDIX A ....................................................................................................................... 45

APPENDIX B ....................................................................................................................... 47

APPENDIX C ....................................................................................................................... 50
LIST OF TABLES

Table 1  Classification of Operator Behavior 14
Table 2  Summary Data Collection Techniques 27
Table 3  Responses to Statement of Opinion and Level of Agreement 29
Table 4  NASA-TLX Results 34
Table 5  Workload Factors 38
Table 6  Traffic Volume & Complexity 40
Table 7  Perceived Workload Characteristics 41
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Schematic Diagram of Controller Workspace Layout</td>
<td>5</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Relationship Between Workload and Performance</td>
<td>7</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Factors Affecting Controller Workload</td>
<td>11</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Mean Rating Per Subscale Across Participants</td>
<td>36</td>
</tr>
<tr>
<td>ACRONYMS</td>
<td>EXPLANATION</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>AAS</td>
<td>Advanced Automation System</td>
<td></td>
</tr>
<tr>
<td>AIDC</td>
<td>Air Traffic Services Interfacility Data Communications</td>
<td></td>
</tr>
<tr>
<td>ARINC</td>
<td>Aeronautical Radio, Inc.</td>
<td></td>
</tr>
<tr>
<td>ARTCC</td>
<td>Air Route Traffic Control Center</td>
<td></td>
</tr>
<tr>
<td>ATC</td>
<td>Air Traffic Control</td>
<td></td>
</tr>
<tr>
<td>ATWIT</td>
<td>Air Traffic Workload Input Technique</td>
<td></td>
</tr>
<tr>
<td>CHI</td>
<td>Computer Human Interface</td>
<td></td>
</tr>
<tr>
<td>CPDLC</td>
<td>Controller/Pilot Data Link Communications</td>
<td></td>
</tr>
<tr>
<td>DYSIM</td>
<td>Dynamic Simulation Laboratory</td>
<td></td>
</tr>
<tr>
<td>FPS</td>
<td>Flight Progress Strip</td>
<td></td>
</tr>
<tr>
<td>FAA</td>
<td>Federal Aviation Administration</td>
<td></td>
</tr>
<tr>
<td>FIR</td>
<td>Flight Information Region</td>
<td></td>
</tr>
<tr>
<td>FPL</td>
<td>Full Performance Level</td>
<td></td>
</tr>
<tr>
<td>HF</td>
<td>High Frequency</td>
<td></td>
</tr>
<tr>
<td>ISD</td>
<td>Interim Situation Display</td>
<td></td>
</tr>
<tr>
<td>MOPS</td>
<td>Minimum Operational Performance Standards</td>
<td></td>
</tr>
<tr>
<td>NADIN II</td>
<td>National Airspace Data Interchange Network</td>
<td></td>
</tr>
<tr>
<td>NASA-TLX</td>
<td>NASA-Task Load Index</td>
<td></td>
</tr>
<tr>
<td>NATCA</td>
<td>National Air Traffic Controller’s Association</td>
<td></td>
</tr>
<tr>
<td>ODAPS</td>
<td>Oceanic Display and Planning System</td>
<td></td>
</tr>
<tr>
<td>ODL</td>
<td>Oceanic Data Link</td>
<td></td>
</tr>
<tr>
<td>SIQR</td>
<td>Semi-interquartile Range</td>
<td></td>
</tr>
<tr>
<td>SWAT</td>
<td>Subjective Workload Assessment Technique</td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>Telecommunication Processor</td>
<td></td>
</tr>
<tr>
<td>Code</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>--------------------------------------</td>
<td></td>
</tr>
<tr>
<td>WJHTC</td>
<td>William J Hughes Technical Center</td>
<td></td>
</tr>
<tr>
<td>ZAN</td>
<td>Anchorage Air Route Traffic Control Center</td>
<td></td>
</tr>
<tr>
<td>ZNY</td>
<td>New York Air Route Traffic Control Center</td>
<td></td>
</tr>
<tr>
<td>ZOA</td>
<td>Oakland Air Route Traffic Control Center</td>
<td></td>
</tr>
</tbody>
</table>
INTRODUCTION

The objective of this thesis was to perform an analysis of the impact of the Oceanic Data Link system on controllers’ workload in the operational environment. Due to the implementation of the prototype Oceanic Data Link system in the operational environment, many concerns regarding the impact of the system on controller workload were raised. A preliminary study conducted by the Human Factors Branch (ACT-530) of the William J Hughes Technical Center (WJHTC) suggested that several issues regarding the Oceanic Data Link system does negatively impact controllers’ workload and that further analysis was necessary to better understand the identified issues. The purpose of this study was to qualitatively measure controller workload in the operational environment with the usage of the Oceanic Data Link system. The data were obtained through observations, a subjective rating scale, and questionnaires.

There has been much research conducted on controller workload and workload studies in general. However, the implications of these studies and workload techniques developed are written in a manner towards workload assessments in the controlled environment. As with any ATC study conducted in the operational environment, there can be confounding variables which cannot be controlled for (e.g., weather, traffic volume, traffic type, etc.). However, these confounding variables, which may impact the assessment analysis, must be accounted and worked around by the human factors analyst.

Background

The airspace of the United States accommodates hundreds of thousands of aircraft movement everyday. Further significant increases in air traffic volume are projected over the next several years. The traffic is handled by ATC Centers and the responsibility for the separation of these flights are handled by individuals or small teams of controllers who are each assigned a
volume of the airspace It is postulated that the significant increase in air traffic volume will
demand higher performance from the air traffic controllers To offset these higher traffic loads
from the controllers, automated ATC systems are currently being developed and will be
implemented within the next several years

For years, human factors analysts have expressed concerns about the psychological
consequences of automation on controllers (Vortac, Edwards, & Fuller, 1994, Murphy, 1995) In
hopes to increase their overall performance without compromising safety, much of the new tools
have automated much of the manual and redundant tasks of the controllers Many of these
automated ATC tools have been in development for the last 10 years and are currently
undergoing field-testing Many of these tools, although products of years of development, have
not undergone an empirical human factors analysis in such important areas as workload and
situation awareness Due to lack of funding and time, the majority of the work in these areas will
be conducted in the operational field testing environment rather than the laboratory setting There
are many constraints, which limit the human factors analyst in performing an accurate empirical
assessment of controller workload in the operational environment These limitations are due to
the high cost of performing a laboratory assessment, the high cost of expensive laboratory
simulations, lack of personnel, and the lack of time to perform an assessment Furthermore, the
workload techniques that have been developed are most appropriate to the laboratory
environment

Oceanic Data Link System

The Federal Aviation Administration (FAA) is responsible for providing air traffic
control services to aircraft flying within the Flight Information Region (FIRs) in a portion of the
western half of the North Atlantic Ocean, a large portion of the Arctic Ocean, and a major portion
of the Pacific Ocean These areas are controlled by Air Route Traffic Control Centers (ARTCC)
in New York (ZNY), Oakland (ZOA), and Anchorage (ZAN) The Oceanic Data Link system, a
concept of automating the data exchange between the pilot and the ATC oceanic controller via data link, is currently in development by Raytheon Systems Company.

The current communication system used by the controller to communicate with the pilots are via the Telecommunications Processor (TP), Aeronautical Radio, Inc (ARINC), or through the High Frequency (HF) telephone. The recent integration of TP in the ocean was to reduce controller workload by automating ATC non-critical tasks and improve safety by reducing errors. The use of TP is to automate the message composition task of the controllers by presenting precomposed ATC phrases with aircraft-specific data included. The implementation of the Oceanic Data Link system is to further enhance this automated message composition process by expanding the number of messages included in the message set. The Oceanic Data Link system is to replace TP, and it is postulated that this change will further decrease controller workload and enhance controller efficiency. The Oceanic Data Link system significantly changes the ways in which oceanic controllers perform their tasks. Whereas with voice communications, controllers have hands and eyes free for scanning and marking flight strips, Oceanic Data Link system requires that controllers to monitor and input messages communication into the computer. Potential communication task efficiencies can only be attained with Oceanic Data Link system given an effective CHI for message inputs and display feedback, with proper training and documentation, and with an integrated cockpit and controller datalink system.

The Oceanic Data Link Control system is currently in operation at a single sector as a prototype system at an Oceanic ARTCC. Oceanic Data Link system will support the following functionality:

1. **Air-Ground Communication via controller/pilot data link communications (CPDLC) message processing,**
2. **Ground-Ground Data Communication - Support Air Traffic Services Interfacility Data Communications (AIDC) message set for Interfacility communications,**
3. **Existing TP functionality usage processing,**
4 Outgoing communications with ARINC radio operators for communications with HF equipped aircraft,

5 National Airspace Data Interchange Network (NADIN) II interface for CPDLC position report transmittal to the Oceanic Display and Planning System (ODAPS),

6 Precomposed Message Responses - Provide precomposed responses to incoming messages to reduce controller workload and allow rapid turnaround of air-ground and ground-ground messages,

7 and message archival

Figure 1 is a schematic representation of the controller workspace layout. The planned architectural design for the Oceanic Data Link system requires a workstation in an oceanic sector to contain the following

1 19-inch display monitor

2 CPU

3 Advanced Automation System (AAS) Keyboard

4 AAS Trackball

The Oceanic Data Link system will incorporate the use of windows, brightness, and color to provide a display which enables quick recognition of the different types of information that will be displayed to the controller. The graphical user interface of the system will contain the following windows

1 Aircraft List

2 Message History List

3 Flight Plan

4 Message Composition List

5 Command windows #1, #2, #3, and various corresponding windows
Throughout the years of assessing workload, there has not been a general consensus on the definition of workload (Murphy, 1995). In general, workload can be simply defined as the resources supplied by a system and the task demand imposed on the operator. These demands can be the amount of mental as well as physical effort expended in performing a given task. In defining ATC workload, much of the workload a controller expends for a given task is mainly cognitive or mental in nature rather than one that is physical. In general, ATC workload can be simply stated as the task demands or what the controller must perform in relation to each controlled aircraft. As such, much of the workload the controller experiences involve complex human information processing and decision making activities.

The goal of measuring ATC workload is to ensure that in developing a new system, it does not cause the operator to be underloaded or overloaded. One objective of assessing an automated ATC system is to ensure that the system does not overload the controllers’ natural human capabilities or underload the controllers so they become too complacent or bored with the
The relationship between workload and performance is rather complex and not well understood. Simply reducing workload does not guarantee improved performance or productivity. An objective of workload and performance measurement is to provide data that system designers can use as a basis for identifying and redesigning embedded sources of overload and underload (Murphy, 1995).

It is postulated that reducing workload improves performance. On the contrary, reducing workload does not always guarantee that performance is improved. The relationship between workload and performance is not as simple as it sounds; it is rather complex and not well understood. Performance and workload are affected by three major factors: 1) the operator tasks defined by the job, the environment, and the system, 2) the transitory state (i.e., initial states such as amount of rest, level of physical fitness, etc. which may or may not be appropriate for the task), and 3) the stable traits (i.e., goals/motivational state, knowledge/skills, and processing capabilities) of the operator. As demonstrated on Figure 2, reduction of workload does not increase performance (see Figure 2). In general, extreme levels of high workload degrade performance and extreme levels of low workload cause the controller to be too complacent and bored (Lysaght, Hill, Plamondon, Linton, Wierwille, Zaklad, Bittner and Wherry, 1989).
Figure 2. Relationship Between Workload and Performance. In Region 1, at extremely low levels of workload, the operator becomes bored. In Region 2, at a reasonable level of workload, there is an acceptable level of performance. In Region 3, as workload is increased, the level of performance degrades (Lysaght et al., 1989).

It is important to distinguish the differences between system workload and human workload. System workload is defined in terms of the number of inputs and outputs a computer system can handle in a unit of time. Human workload results from entering information into the computer system which is decided and acted upon by the following sources: ATC environment (e.g., traffic load), the hardware/software, and the individual behavior and individual differences of the operator. Human workload is broken down into those actions, which are observable (and quantified) and perceived (subjective). Observable workload considers the number of aircraft, complexity of the aircraft environment, communication, time on task, etc. In short, observable workload are those that are manual or verbal. Subjective/perceived workload, on the other hand, deals with the controller’s personal experience with or subjective perceptions of the system and their mental tasks, such as planning and problem solving. Because of this difference in measuring
ATC workload, both observable (objective) and perceived (subjective) aspects of operator workload should be considered (Murphy, 1995)

**Productivity, Performance, and Automation**

The objective of an automated ATC system is to remove or reduce the routine human functions and to reduce human workload so that the controller may handle increased amounts of air traffic while maintaining high standard of safety. It should also be designed to reduce or prevent human error and its consequences, and to achieve a successful and optimum matching of human and machine for the performance of air traffic control tasks (V D Hopkin, personal communication, 1995)

The addition of an automated system should not remove the controllers from their environment, but to keep the controllers “in-the-loop” of the system. In order for an automated system to be successful, it must support the controllers, natural information processing tasks and situation awareness of the environment. The goal of the automated system is to assist the air traffic controller, not to change his/her tasks and responsibilities. The automated system should gather, collate, summarize, and present the information to the air traffic controller without further processing or recording. It is expected that controllers will be more efficient, spend less time on controlling each aircraft, and provide better air traffic control service, while still maintaining safety (V D Hopkin, personal communication, 1995). When evaluating an automated system, the common trends in the ways that the system aids or impairs the controllers in performing their task should be included in the evaluation. These evaluations should consider individual efficiency and effectiveness, and not only system efficiency and effectiveness (Murphy, 1995)

**Difficulties in Measuring Workload**

The fundamental reason for the difficulty in measuring workload is that workload is complex and multi-dimensional. Various behavioral and physiological measures have proven
unreliable, which suggests that the majority of the workload the controller experiences is mental or cognitive. Because of this psychological aspect of controller workload, it is difficult to equate these subjective measures to an objective assessment (Stein & Garland, 1993).

Many workload studies have failed to measure workload because these studies often equate workload to task demands, although they are not one in the same. There are other variables, which need to be accounted for, such as experience, age, fatigue, etc (V D Hopkin, personal communication, 1995). Furthermore, controllers have the tendency to adopt different strategies for various situations (e.g., dealing with a number of different aircraft) which makes it difficult to assess workload in varying levels of traffic complexity, especially in low and high levels of traffic complexity (Stein, 1991).
OVERVIEW OF EMPIRICAL WORKLOAD ASSESSMENT

There are three classes of ATC assessments that an analyst can apply to measure workload. These three classes used by human factors analysts are performance-based (observable), perceived (subjective), and physiological assessments. Performance-based assessments measure some aspect of the operator's ability to perform tasks or system functions. Perceived workload assessments derive estimates of workload from controllers feedback concerning the workload or effort expenditures they experienced during task performance (Eggemeir & Wilson, 1991). Physiological workload assessments measure the body's response, such as heart rate, blood pressure, core temperature, and skin conductivity. These three classes will be further described in the following section.

There are several factors which influence workload, such as airspace characteristics, the ATC environment, and the controllers' behavior. These factors are observable, however, the workload perceived by the controllers may vary. This variation in perceived workload is often affected by other factors that are not observable. These unobservable factors include training, experience, skill, fatigue, etc. As previously stated, workload is a multidimensional construct which requires measures that tap into workload from different perspectives (i.e., those that are observable and those that are not). Therefore, although workload can be measured in an observational format, it is emphasized that workload assessments should investigate both observable and unobservable factors, and be measured via a combination of performance-based, perceived, and/or physiological assessments (Murphy, 1995, Eggemeir & Wilson, 1991).

Influences on ATC workload

There are four factors that influence controller workload. These four factors are:

- ATC complexity factors
- Hardware/software
- Operator behavior
- Individual differences of the operators

Based on a compilation of the available research, Figure 3 demonstrates the relationship among the above factors and the relationship to controller workload (Lysaght et al., 1989; Hart & Wickens, 1990, *The Complexity Construct in Air Traffic Control: A Review and Synthesis of the Literature*, 1995). ATC complexity consists of the air traffic pattern and sector characteristics. The mediating factors include the hardware/software (quality of equipment), operator behavior, and individual differences of each controller. The four factors that influence controller workload are explained in the following paragraphs.

![Diagram of ATC Complexity, Mediating Factors, and Controller Workload](image)

*Figure 3. Factors Affecting Controller Workload.*
**ATC Complexity**

ATC complexity is one factor that contributes to controller workload. As previously discussed, although measuring controller workload is mainly a subjective task, the traffic environment also directly influences the controller’s workload. ATC complexity has a measurable influence and accounts for a large proportion of the controller’s workload (*The Complexity Construct in Air Traffic Control*, 1995).

ATC complexity is a multidimensional construct which includes static sector characteristics (sector complexity) and dynamic traffic patterns (traffic complexity). The report *The Complexity Construct in Air Traffic Control* (1995) cites Grossberg (1989) who makes a “distinction between the attributes of a sector and their effects on the controller.” Complexity is defined as a construct that has both dynamic and static characteristics that affect the rate at which the controller workload increases. Controller workload is the activities, both mental and physical, which result from handling air traffic.”

ATC complexity is a construct that is composed of a number of sector and traffic complexity dimensions or factors. These functions can be physical aspects of the sector, or factors relating to the movement of air traffic through the airspace. Some factors cover both sector and traffic issue, e.g., required procedures and functions. Theoretically, the structure of a sector is separate from the characteristics of the air traffic. A given level of traffic density and aircraft characteristics may create more or less complexity depending on the structure of the sector. The following are examples of general ATC and sector complexity that a human factors analyst can apply as metrics when measuring air traffic controller workload:

- Traffic volume (number of aircraft controlled in the sector)
- Mixture of aircraft types
- Total number of flights handled
• Traffic mixture (arriving/departing vs overflying aircraft)

• Transfer of control

Hardware/software

The information display and the manner in which the human processes the information identify the hardware/software portion illustrated in Figure 3. When designing the CHI, designs must be chosen in relation to human capabilities of vision, information processing and understanding. Much research has been conducted on the information display and processing factors which affect ATC task difficulty and complexity. In general, studies have shown that the quality of the system transmitting the information about the sector and the aircraft within it affects the adequacy of the information reaching the controllers’ senses (Complexity Construct in Air Traffic Control, 1995, Hopkin, 1995).

Operator Behavior

Taxonomies are developed as aids in scientific classification. These classifications serve the useful purpose of grouping similar criteria together as well as being helpful in explaining their structure. The task taxonomy is one taxonomy that can be useful in helping to determine the appropriate workload techniques for a specific application. There are two main purposes for a task taxonomy: 1) classifying the nature of the operator tasks, and 2) classifying workload assessment techniques (Lysaght et al., 1989).

Before the operational assessment, for the purpose of building scenarios of predicting workload, a task analysis should be completed and information requirements identified. An attempt to develop such a scenario is taken from Berliner, Angell and Shearer (1964). Table 1 lists the four processes: perceptual, mediation, communication and motor. The attempt is to apply the four processes that generalize the ATC behaviors the controllers perform. Each general process is broken down further to the specific behavior the controller performs. Identifying
specific behaviors can be applied as a basis for task performance measures and analyzed to how much control and effort is exerted by the controllers when they interact with the automated ATC system (Berliner, Angell & Shearer, 1964, Casah & Wierwille, 1984)

Table 1

*Classification of Operator Behavior*

<table>
<thead>
<tr>
<th>Processes</th>
<th>Activities</th>
<th>Specific behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Perceptual processes</td>
<td>1 1 Searching for and receiving information</td>
<td>1 1 1 Detects</td>
</tr>
<tr>
<td></td>
<td>1 2 Identifying objects, actions, events</td>
<td>1 2 1 Discriminates</td>
</tr>
<tr>
<td>(2) Mediational processes</td>
<td>2 1 Information processing</td>
<td>2 1 1 Categorizes</td>
</tr>
<tr>
<td></td>
<td>2 2 Problem solving and decision-making</td>
<td>2 2 1 Analyses</td>
</tr>
<tr>
<td>(3) Communication processes</td>
<td>3 1 Advices</td>
<td>3 2 Answers</td>
</tr>
<tr>
<td></td>
<td>3 2 Answers</td>
<td>3 3 Communicates</td>
</tr>
<tr>
<td></td>
<td>3 3 Communicates</td>
<td>3 4 Directs</td>
</tr>
<tr>
<td></td>
<td>3 4 Directs</td>
<td>3 5 Indicates</td>
</tr>
<tr>
<td></td>
<td>3 5 Indicates</td>
<td>3 6 Informs</td>
</tr>
<tr>
<td></td>
<td>3 6 Informs</td>
<td>3 7 Instructs</td>
</tr>
</tbody>
</table>
### Individual Differences

Air traffic controllers experience workload differently because they differ in terms of individual traits or capabilities. They can differ in many ways that may make the same task harder or simpler to perform (Lysaght et al., 1989). Factors that must be considered include experience, training, knowledge/skills, etc. Gathering data regarding individual differences can only be obtained by questioning the controllers and cannot be obtained via an observational format.
A SUMMARY OF EMPIRICAL TECHNIQUES

Three types of ATC assessments which an analyst can identify to adopt a particular technique to measure workload are performance-based, perceived (subjective), and physiological assessments. Performance-based measures utilize some aspect of the operator's capability to perform tasks or system functions in order to provide an assessment of workload. Subjective measures derive estimates of workload from operator reports concerning the workload or effort expenditures that were experienced during task performance (Eggemeir & Wilson, 1991). Physiological workload measures the body's response, such as heart rate, blood pressure, core temperature, and skin conductivity. The following is a brief discussion of the types of workload measurement techniques that are available for assessing controller workload and productivity.

Performance-based

The two major categories of performance-based techniques are (1) primary-task measurement, and (2) secondary-task methodology. Primary-task measures assess some aspect of the operator's capability to perform the task or system function of interest. The air traffic controller's primary task is to maintain separation of aircraft. Secondary tasking derives an index of workload from the operator's capability to perform two concurrent functions or tasks along with the primary task (Eggemeir & Wilson, 1991). Secondary tasks of a controller include those actions that are secondary to maintaining aircraft separation such as flight progress strip (FPS) maintenance, etc.

Primary-task Measurement

Primary-task measurement is applied with the expectation that the speed and/or accuracy of performance will decrease as workload increases beyond a critical value or threshold for
unimpaired performance. Automation or decision support system capabilities may increase this threshold. It can also be applied in evaluating the efficiency of the operator’s information-processing. This approach utilizes modifications in the way in which a controller performs a task in order to gain an index of changes in workload (Eggemeir & Wilson, 1991).

The primary task is critical to an operator’s performance within a system and it should always be included as part of the workload evaluation. There are many studies which have proven primary-task measurement is sensitive to variations in workload in a wide variety of system-related applications. Instances where primary task measures prove to be insensitive is due to the operator’s ability to apply extra processing resources to meet increased demand, thereby maintaining adequate levels of primary task performance over some levels of increased workload (Wierwille & Eggemeir, 1993). Air traffic controllers compensate for increases in demand by varying their strategies in low and high levels of workload. That is, as workload increases, the controller’s primary task performance will remain adequate and will not vary between the changes of workload, particularly between low to moderate levels of workload.

In selecting the metrics for a primary task technique, latency and error scores are excellent candidates and have been reported as sensitive across a half dozen studies reported by O’Donnel and Eggemeir, 1986. The task taxonomy, like the Universal Operator Behaviors developed by Berliner, Angell, & Shearer, (1964) also serves as an excellent place to assist in deciding what aspects of controller behavior are to be measured. Spreading the tasks across the categories increases the opportunities for identifying performance measures that are sensitive to workload and it is also highly useful to measure two tasks which fall in different categories (Lysaght, et al., 1989).

*Secondary Tasking*

External secondary tasking is not part of the normal function of the controller and is intended to investigate different levels of human information processing and response functions.
Each controller has some inherent overall level of "workload capacity", and that each task or function performed draws upon or "uses up" a certain amount of that capacity. Further, it is generally assumed that if each type of task requires or uses a certain amount of this overall capacity, that whatever capacity level is left over, it represents a kind of "excess" or "reserve" capacity which might be available for other purposes. Finally, it is also frequently assumed that as numbers of tasks become more closely spaced in time, and/or the relative "demand value" of various of those tasks becomes greater, then more and more of this "capacity" is required to support those tasks — and correspondingly, less of the "excess" or "reserve" capacity remains.

The process involves the controller to perform the primary task (separating aircraft) and while also performing an additional or secondary task (e.g., FPS maintenance). This will measure how much cognitive resources are left over (e.g., residual memory). The relative workload associated with the primary task is reflected by the performance on the secondary task (Lysaght, et al., 1989). The most common secondary tasks include memory, mental mathematics, interval production, reaction time, time estimation and tracking (Wierwille & Eggermeir, 1993). The goal is to shift the controller's workload from low to moderate, where operator performance is expected to reflect variations in the workload associated with the performance of a task.

A disadvantage faced by analysts with the selection of secondary tasking is the risk of intrusion. Intrusion refers to the degradation of an ongoing primary-task performance associated with the application of a workload measurement technique. Furthermore, imposing an external secondary task on a controller may affect safety. Secondly, controllers may find this task to be artificial and bothersome and may cause the operator to fail on such a task (Eggemeir & Wilson, 1991). Care must be taken when exercising this type of measurement, so as to not cause unwarranted intrusion. Due to the problem of intrusion and the consequences in jeopardizing safety, secondary task measurement is not recommended to be performed in the field environment under operational conditions. However, this can be performed in the dynamic simulation laboratory (DYSIM) of the ATC Center.
Interpreting results of a secondary task measure must be proceeded with caution. A decline in secondary task performance does not necessitate high workload. Therefore, when secondary tasks are used to measure workload, several different secondary tasks should be used. According to Hart and Wickens, “each secondary task should demand different combinations of resources so that the origins and levels of primary task can be determined more accurately” (Murphy, 1995).

**Subjective Workload**

When assessing workload much of the work performed by an air traffic controller is mental or cognitive. The primary purpose of using subjective methods is to gain access to the experiences of the controller. Mental workload cannot be directly measured because it is internal and can only be inferred by observers. Subjective workload assessments are used to obtain and quantify the opinions and judgments of the controller and thus, researchers have suggested it is the most appropriate method to assess workload.

Subjective measures are designed to reflect variations in the subjective feelings or effort expenditure that are assumed to be associated with increases across low, moderate, and high workload (Eggemeir & Wilson, 1991). Subjective workload measurement techniques have been used for a number of applications in the past several years. Their use has been widely accepted and considered an integral part of the workload assessment (Hart & Wickens, 1990).

Speed and accuracy are two types of task performance indicators used as measurements of perceived workload. The assumption is that changes in task performance reflect changes in workload. However, other factors confound these indicators, such as motivation, time available, and operational strategies. That is, highly motivated air traffic controllers will put a much greater effort in producing faster and accurate results in separating aircraft (Murphy, 1995).

The two broad classes of subjective methods are (a) rating scales and (b) questionnaires and interviews. Over the years, there have been several subjective measurements.
rating scales that have been developed which have demonstrated a capability to reflect variation in demand across a variety of different tasks. Three rating scales have been widely used and include variants of the Cooper-Harper, NASA-Task Load Index (TLX) and the Subjective Workload Assessment Technique (SWAT).

Two of the most commonly used rating scales are NASA-TLX and SWAT. These involve a procedure whereby ratings on several scales are combined to produce a summary score. Choice of one technique or another for applications to such environments is dependent upon direct comparisons of techniques with respect to properties such as ease of use and sensitivity. For example, the use of SWAT procedure has facilitated verbal report of ratings by pilots in the flight environment (Schick & Hahn, 1987). Until more extensive data is developed, the choice technique for a particular measurement application should be influenced by the capability of the procedure to meet both the objectives of the evaluation and the constraints of the individual environment (Eggemeir & Wilson, 1991, Casali & Wierwille, 1983, Wierwille & Eggemeir, 1993, Hart & Wickens, 1990).

A pitfall of subjective measurement is its reliance on the perception of the operator and may not be an indicator of controller workload. Moreover, if workload is close to being the same between two comparative systems, the subjective measures could not measure the very small changes in low workload. Another constraint is that many controllers are reluctant to rate any level of workload as too high and as something they cannot handle. Therefore, it is recommended to use an objective measure, such as primary task along with a subjective test (Murphy, 1995).

Physiological Workload

Physiological measures include heart rate, biochemical changes, galvanic skin response, blood pressure, brain activity, and various optical measures, etc. These do not measure workload directly but, they measure the body's response to stress induced by workload. Assessing physiological workload relies on the individual. That is, how one person reacts to stress can be
different from another person. In general, the less intrusive the measurement technique, the better will the results reflect pure workload values (Murphy, 1995). Physiological measurements are difficult to obtain, not from the viewpoint of the human factors analysis, but rather from the controller’s workload environment which is approved and dictated by the National Air Traffic Controller’s Union (NATCA).
REVIEW OF THE RELATED LITERATURE

In examining the abundant literature related to air traffic control, an analyst can become inundated with the vast amount of information on the assessment of air traffic controller workload. These studies vary widely in the approaches used to assess workload, ATC complexity, and measurement techniques. The goal of this section is to disseminate the related literature of ATC workload and to apply any techniques learned from those studies to an assessment in the operational environment. In retrospect, although there are many studies conducted on ATC workload, the amount of information workload performed in the operational environment is not as vast.

Edwards, Fuller, and Vortac (1995), performed a study of automation and its replacement of flight paper strips. Although this study was performed not for the sake of a workload study, its content and approach can be applied to an operational approach. This observational study focused its approach by examining the control actions, communication events, and computer interactions of the controllers. Two observers sat behind the controllers with laptops and recorded their time-based behaviors. The goal of reviewing the paper was to assess the time-series analysis in which the behaviors of individuals and teams of controllers were categorized “on-line” while controlling simulated traffic. The behavioral categories included the range of activities that the controller performed on the strips. Communication events and computer entries were also recorded. Time-series models were developed to predict flight strip activity from communication events, and computer interactions.

The analysts time-stamping the controller’s behavior and comparing the time differences between the two systems in the study obtained their data. This type of assessment of time stamping
task performance behaviors could be used to compare the behavior of the controllers' and can be performed in the DYSIM and the operational environment without being intrusive.

Stein (1985) conducted a simulation that determined the relationship between the number of airspace factors and controller workload. Workload was measured by the Air Traffic Workload Input Technique (ATWIT) in which the controller was asked to press buttons from 1 to 10 to indicate the level of workload experienced. Importantly, controllers were able to provide real-time workload estimates using ATWIT without any noticeable decrement in performance.

Casali and Wierwille (1984) performed a workload study on pilots in a flight simulator. The experiment examined fourteen distinct mental workload estimation measures, including opinion, secondary task, physiological, and primary task measures. An assessment of the relative sensitivity of the measures to changes in mental workload and the differential intrusion of the changes on primary task were performed. This was part of a multiple workload assessment that systematically compared measures across a series of experiments.

The objective of the research, to establish the relative sensitivity and intrusion of a variety of workload estimation techniques with regard to a flight-related perceptual loading changes, was to a large degree met. The performance measures were obtained by using the Universal Operator Behavior (Berliner, Angell, & Shearer, 1964). This type of technique shows a successful use of the Universal Operator Behavior and can be applied to the operational environment approach.

Whitefield (1979) used off-line discussions with ten air traffic controllers who were in training to research the air traffic controller’s mental picture. The use of questionnaires and discussions generated opinions on sources of ATC workload, comments on differences between the comprehension abilities of experienced and less-experienced controllers and references to a sense of foreground and background in the picture. This technique is not intrusive and provides preliminary insights. However, it does not lend itself to statistical analysis (Mogford, Harwood, Murphy, & Roske-Hofstrand, 1994). Its application to the operational environment is therefore,
for the purpose of obtaining an analytical workload assessment of a workload study in the DYSIM or the field environment is not feasible.

The use of the DYSIM "provides a rich, high-fidelity environment for data collection. Although DYSIM lacks the ultimate reality of live operations, controllers generally find such problems compelling and highly motivating." Collecting data in the DYSIM is non-disruptive to the ATC operations in the operational environment. It also provides an opportunity to video and audio tape the air traffic controllers, a technique that is difficult to be approved in the operational environment. A study conducted by Zachary et al. (1989), used structured DYSIM problem solving "to capture performance strategy." Results from this were used in an analysis to derive a glossary of display strategies, control strategies, and workload-reduction strategies (Mogford et al., 1994).

According to Mogford et al. (1994) the use of DYSIM provides a degree of control not available in a live environment. They provided the following example, different participants can be given identical problems, thus, providing a valid basis for within-group and between-group comparisons. Variables such as time pressure, problem difficulty/complexity can be held constant or manipulated. Therefore, the use of the DYSIM is preferred over the operational environment for assessing ATC workload. Furthermore, many of the techniques covered in the previous sections can be utilized in the DYSIM environment.

Rehman, Stein and Rosenberg (1983), performed two studies on subjective pilot workload. The method employed the use of a switch box device in the cockpit simulator. The first study used pilots and non-pilots who input workload evaluations each minute during a critical tracking test. The second test was used to determine if pilots could differentiate between three flights in which the level of difficulty was varied. A postflight questionnaire was also provided. The workload rating scale used a switch box device containing an array of 10 push buttons that was used to obtain minute-by-minute workload responses during task execution.
In determining its use in the ATC environment, it should be noted that this type of assessment is more functional in the laboratory setting. However, it is postulated that with careful planning, this type of assessment can be used successfully in the operational environment. The method should allow the controllers to provide their on-line assessment of the perceived workload. The workload data can be correlated with other variables which influence workload such as traffic loads, the number of type of communication made, and number of motor movements, etc.
METHODOLOGY

Overview

This section discusses the assessment methodology of controllers’ workload operating the Oceanic Data Link system in the operational environment. It includes the types of data collected, a description of the data collection procedures, the results, and the protocols employed.

Preparation

As part of the initial planning phase of the assessment, the following steps were taken and the following documents reviewed:

- Conducted technical discussions with ARU-100, ATS-240, NATCA, and various contractors and developers of the Oceanic Data Link system.
- Visited the ARTCC site and conducted preliminary discussion with the ARTCC’s Oceanic Manager and Oceanic Data Link representatives to enhance understanding of oceanic operations and procedures. This helped to identify any additional data collection risks and issues.
- Thoroughly reviewed documentation relating to the Oceanic tower controller tasks to establish a solid background regarding the controllers’ duties. Oceanic Data Link documentation, prepared by the FAA, contractors, and developers of the system were reviewed in order to gain a thorough understanding of the system functions, features, and characteristics.
- Reviewed the information obtained in a past human factors assessment performed by the Human Factors Branch (ACT-230) of the WIHTC.

Assessment Methodology

The assessment was conducted to assess the effects of the Oceanic Data Link system on controller workload under high traffic load conditions. The current single sector in which the
Oceanic Data Link system is operational was evaluated to examine the expected envelope of usage under high traffic load conditions. The reason for conducting the assessment under high traffic load conditions is that the preliminary WJHTC study and controller reports have suggested that the usage of the Oceanic Data Link system creates more workload for the controller than the current ATC system during peak traffic conditions. Over the course of the assessment period, the human factors analyst observed and collected data from a total of 11 different controller participants. Qualitative measures were used to assess workload and identify issues associated with the controllers' oceanic ATC tasks. Qualitative information included the NASA-TLX, questionnaires, and interviews regarding the workload associated with the Oceanic Data Link system. Provided below is a summary list of the data collection techniques.

Table 2

<table>
<thead>
<tr>
<th>Subjective Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rating Scales</strong></td>
</tr>
<tr>
<td>Post hoc method using questionnaire rating scales to quantify perceived opinion and judgments</td>
</tr>
<tr>
<td><strong>Questionnaires and Interviews</strong></td>
</tr>
<tr>
<td>Post hoc method to obtain contextual information and estimates, judgments, evaluations, comparisons, attitudes, and opinions</td>
</tr>
</tbody>
</table>

Participants

A total of 11 controller participants (over 50% of the user population was sampled) from the oceanic data link sector participated in the study. Selection was a convenience sample. These participants were volunteers who had agreed to participate in the study. They are all oceanic full-performance level (FPL) controllers who are qualified to work the oceanic data link sector traffic. During a preliminary briefing, each controller was asked to complete a participant entry.
questionnaire in which they were asked to rate their current controller skill, the amount of training and experience received on the Oceanic Data Link system, their agreement with their status as "volunteers," and whether or not they were looking forward to the experiment (see Appendix A). They were asked to complete the Participant Entry Questionnaire to obtain some information regarding their experience and current attitudes (see Appendix B). These controllers had received previous training on the Oceanic Data Link system, and were experienced in the use of the Oceanic Data Link system to control the sector's traffic. They ranged in oceanic air traffic control experience from 4 - 16 years, with 8.5 average years of experience at the ARTCC under study. All controllers had worked all oceanic sectors at the ARTCC. All controllers received training in the form of classroom and on-the-job training on the Oceanic Data Link system within the past 2 years.

Controllers indicated the amount of experience using the Oceanic Data Link system was a mean of 1.25 years, with a range of 4 months to 2.5 years. It would have been most appropriate to statistically analyze the controller workload against experience with the Oceanic Data Link system, however, it would not be feasible because of the various system software upgrades since its inception at the Center. They rated themselves on their level of usage of the Oceanic Data Link system as 1 beginner, 6 intermediates, 4 advanced. A beginner was defined as having no knowledge or little knowledge of using the system. An intermediate user was defined as having enough knowledge to use the system with little assistance. An advanced user was defined as having knowledge to use the system using several of the shortcut keys. An advanced user was defined as having knowledge to efficiently use all shortcut keys. Since the Oceanic Data Link system is a computerized system and which differs from their present manual way of controlling traffic, the controllers were asked several questions regarding their computer experience level. In regards to their computer experience level, 1 controller rated himself as a beginner and the remaining 10 rated themselves as intermediate. In regards to entering numeric data, 2 controllers rated themselves as beginner, 3 intermediate, and 5 advanced.
Table 3 shows the controllers' responses regarding their opinion and level of agreement for each statement. The table shows the median and semi-interquartile range (SIQR). The statistics were computed with Microsoft Excel Version 5 0e software. The median is the score or point below which 50% of the raw score lies. The SIQR measures the variability of a distribution that describes the spread or range of scores in the distribution. The SIQR is one-half the difference between the score values at the 75th and 25th percentiles. Overall, all controllers responded that they freely participated in the study.

### Table 3

**Responses to Statement of Opinion and Level of Agreement**

<table>
<thead>
<tr>
<th>Question</th>
<th>Median</th>
<th>SIQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;I know a great deal about ODL&quot;</td>
<td>5 0</td>
<td>1 3</td>
</tr>
<tr>
<td>&quot;ODL is important for airspace system development&quot;</td>
<td>4 5</td>
<td>1 2</td>
</tr>
<tr>
<td>&quot;I need additional training for ODL&quot;</td>
<td>5 0</td>
<td>1 2</td>
</tr>
<tr>
<td>&quot;The level of training I receive on new ODL features is adequate&quot;</td>
<td>5 0</td>
<td>1 0</td>
</tr>
</tbody>
</table>

*Note: Questions were on a scale from “1” strongly agree to “10” strongly disagree.*

Each controller was briefed concerning his right to informed consent and privacy. Data collection was accomplished using numbers so that specific data could not be traced back to an individual participant.

**Scenario**

The operational conditions for data collection were to be held during the sector’s daily peak traffic conditions which normally occurs during the time of 10 PM to 5 00 AM, PST (07 00 - 15 00, UTC). The data collection period was scheduled during the peak conditions to analyze the distribution of workload under peak traffic loads, during which the controllers' experience the greatest workload level of operating the Oceanic Data Link. That is, greater the traffic volume,
greater the workload the controller experiences. As with any analysis performed in the operational environment in lieu of a controlled laboratory environment, there are many factors (such as weather and traffic) which cannot be controlled. These factors are measures of controller workload and when these factors deviate from the expected, the analyst must account and adjust for these unexpected factors. In this study, due to the weather conditions over the Pacific Ocean, the traffic was routed to traverse other adjacent sectors. Thus, the normal high peak traffic period over the oceanic data link sector did not occur during the data collection activity period. This weather condition minimized the expected high traffic volume and thus, as per the Oceanic supervisor, equated to an average day's level of traffic volume for the single sector.

Physical Setup

During the period of on-line ATC observations, the analyst observed and annotated the activities performed by each of the individual participants. The analyst sat directly behind the controller and encouraged him to provide a continuing verbal protocol (“think aloud commentary”), describing and explaining any issues of the Oceanic Data Link system which directly influenced his workload (see Appendix C). Other than this relatively unobtrusive side-task, the controllers were instructed to perform their duties in a normal manner.

Data Collection Materials

The data collection materials included two types of instruments. The first material consisted of a subjective questionnaire that asked them to rate those factors (on a scale from 1 very low to 10 very high) which contributed to their workload (see Appendix D). The second questionnaire, focused on questions on the traffic volume and complexity of the traffic period (on a scale from 1 low to 10 high), and general questions relating to their perceived workload with the usage of the Oceanic Data Link system (see Appendix E). The second type of subjective instrument, employed the use of a self-rating instrument, the NASA-TLX (see Appendix F). The
NASA-TLX workload assessment technique approaches workload as a combined construct made of several dimensions such as mental load, physical effort, temporal (time-based), performance, effort, and frustration. The technique allows for individual workload experiences, and that each controller assigns different levels of importance to each dimension. It also gathers both relative dimensional importance weightings, as well as experienced workload level ratings from each participant. To normalize the rating scale, during the preliminary data collection activities, the controllers were gathered to compare and agree upon the ratings on each dimension. During this briefing activity, the human factors analyst lead the discussion and asked each controller to describe and provide examples of what they consider very high, medium, and very low workload for each dimension. This discussion lead to the controllers to discuss, compare, and agree amongst themselves what they consider very high, medium, and very low workload of each workload dimension.

Assessment Conduct

The following scenario was conducted to observe and collect the data from each of the controller participants during the 3 day assessment.

- Each day of the data collection activity, 3-4 controllers participated as participants. Controllers participating in the evening’s activity were given a 25-30 minute orientation briefing prior to the start of their shift. The orientation provided the controllers with the purpose and objective of the study. They were also explained the data collection activities and asked to complete several forms (i.e., consent form and participant entry questionnaire). They were asked to sign a consent form which informed them that their participation was voluntary, anonymous, and that we were evaluating the Oceanic Data Link system and not their personal performance. Finally, the human factors analyst described the purpose of the computerized NASA-TLX and given some practice time to ensure that they developed the correct technique for inputting their responses. This activity also provided a short exercise in
which the controllers were asked to provide their views of the type of workload they
experience for each NASA-TLX factor. This activity was to normalize the rating scale for
each six subscales.

- Each controller were given the opportunity to control traffic for approximately 1 hour at the
  sector in study. During this time period, the human factors analyst observed and noted any
  workload issue and those described by the controller. While working at the sector, the
  controller was encouraged to “talk aloud”, describing and explaining any issues regarding the
  workload they experience using the Oceanic Data Link systems. Although this activity may
  seem obtrusive in nature, controllers commented that the “talk aloud” commentary did not
  disrupt their duties, and were able to control traffic in their normal manner.

- After the session, each participant would go to a separate room and debriefed. During this
  debriefing, the participant would be asked questions regarding his experiences of the Oceanic
  Data Link system from the prior activity. They were also given the two workload
  questionnaires which asked them to rate factors which contributed to their workload, the
  traffic volume and complexity of the traffic period, and rate several questions of the
  perceived workload characteristics.

- During the above debriefing session, the controllers were instructed to input their responses
  on the NASA-TLX.
RESULTS

NASA-TLX

The NASA-TLX is a multi-dimensional rating procedure that provides an overall workload score based on a weighted average of ratings on six subscales: Mental Demands, Physical Demands, Temporal Demands, Performance, Effort, and Frustration. As described previously, there is no clear definition of workload, but all human factors analysts agree that workload is complex and multi-dimensional. The NASA-TLX takes into consideration that the determinant of workload is specific sources of loading imposed by different tasks. Although the NASA-TLX and SWAT have appeared to be valid assessors of subjective workload, the NASA-TLX was chosen over SWAT because of the following reasons. The first reason was due to its relative convenience of obtaining the computerized version of the NASA-TLX. The second reason is its practical application in the operational environment and the relative ease of quickly obtaining the weights and ratings (Hart & Staveland, 1988). Third, the NASA-TLX workload parameters evaluation is specifically intended to reduce the between-subject variability in the workload ratings. Fourth, the technique weights subjects' ratings by their own individual biases about the importance of each scale to workload (Vidulich & Tsang, 1985). The NASA-TLX combines the six subscale ratings that are weighted according to their subjective importance to participants in a specific task. Therefore, the ratings of factors deemed most important in creating the workload of a task is given more weight in computing the overall workload score (Task Load Index).

The NASA-TLX is a two-part evaluation procedure consisting of both weights and ratings. According to the Task Load Index Manual, the administration of the NASA-TLX is possible to be administered in the operational setting since the participants can provide their ratings quickly. The first part of the administration process required the participants to rate their
contribution of each factor (its weight) to the workload of a given task. Secondly, the participants rate the magnitude of each factor (its rating) in a given task. The overall workload score is computed for each participant by multiplying each rating by the weight given to that factor by the participant. The sum of these weighted ratings is divided by 15 (the sum of the weights) to get the overall workload score. The ratings are assumed to be interval properties (Lysaght et al., 1989). The computerized version of the NASA-TLX was used to gather the ratings and weights, and compute the weighted workload scores. Table 4 summarizes the NASA-TLX computed data for all participants. Columns two through seven are the computed version of the weights and ratings, and the last column shows the computed Weighted Workload Score (WWL). The WWL can range from 0 to 100. The last row of the table shows the mean computer ratings and weights, and the mean score of WWL across all participants.

Table 4

<table>
<thead>
<tr>
<th>Participant</th>
<th>Mental Demand</th>
<th>Physical Demand</th>
<th>Temporal</th>
<th>Effort</th>
<th>Performance</th>
<th>Frustration</th>
<th>WWL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>5</td>
<td>5</td>
<td>20</td>
<td>5</td>
<td>45</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>70</td>
<td>20</td>
<td>85</td>
<td>65</td>
<td>80</td>
<td>45</td>
<td>55</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>15</td>
<td>50</td>
<td>15</td>
<td>5</td>
<td>55</td>
<td>31</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
<td>35</td>
<td>25</td>
<td>30</td>
<td>20</td>
<td>5</td>
<td>21</td>
</tr>
<tr>
<td>7</td>
<td>20</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>5</td>
<td>70</td>
<td>32</td>
</tr>
<tr>
<td>8</td>
<td>80</td>
<td>5</td>
<td>50</td>
<td>70</td>
<td>20</td>
<td>65</td>
<td>56</td>
</tr>
<tr>
<td>9</td>
<td>80</td>
<td>25</td>
<td>85</td>
<td>80</td>
<td>15</td>
<td>65</td>
<td>64</td>
</tr>
<tr>
<td>10</td>
<td>70</td>
<td>15</td>
<td>85</td>
<td>80</td>
<td>20</td>
<td>75</td>
<td>65</td>
</tr>
<tr>
<td>11</td>
<td>55</td>
<td>5</td>
<td>55</td>
<td>55</td>
<td>5</td>
<td>7</td>
<td>28</td>
</tr>
<tr>
<td>Mean</td>
<td>4182</td>
<td>1273</td>
<td>4182</td>
<td>400</td>
<td>1682</td>
<td>4064</td>
<td>3527</td>
</tr>
</tbody>
</table>


As shown on table 4, the overall workload rated for each participant Participant 4, 8, 9, and 10 rated their workload experienced in the assessment higher than the rest of the participants.

The following statistical analysis was computed with Allyn & Bacon Stats Demo for Shavelson by David W Abbott, Ph D, Copyright 1989. The mean ratings per subscale mental demand, physical demand, temporal, effort, performance and frustration are 41.82, 12.73, 41.82, 40.00, 16.82, and 40.64 (respectively). These means do differ significantly with a one-way within subjects ANOVA, $F(5, 60) = 2.85, p = 0.022$. Thus, the use of the Oceanic Data Link system influences the scores of each subscale. An Omega Square of 12 indicates that 12% of the variability in the controller workload scores of each subscale is related to the Oceanic Data Link system.

Figure 4 shows the mean workload rated across each subscale by all participants. As the figure shows, subscales Mental Demand and Temporal Demand (41.82 and 41.82, respectively) received higher rating contribution than the subscales, physical, effort, performance, and frustration. Subscales effort and frustration (40.00 and 40.64, respectively), although close in rating to Mental Demand and Temporal Demand, was rated a higher contribution of workload than subscales physical demand and performance (12.73 and 16.82, respectively).

These apparent differences were evaluated with a Tukey HSD post hoc procedure, HSD = 33.65, p = 0.05. It appears that there are no significant difference in the usage of the Oceanic Data Link and scores on each subscale.

In summary, scores of each workload subscale of the NASA-TLX are significantly related to the Oceanic Data Link system. However, there are no significant differences in the means between subscale mental demand, physical demand, temporal, effort, performance and frustration. However, we are unable to conclude that the Oceanic Data Link system produces these differences since this is a post facto study.
In comparison to the other subscales, it is not surprising that Mental Demand and Temporal Demand received the highest workload rating since controllers’ work is mainly cognitive in nature. Also, the level of Effort and Frustration the controllers’ exerted both were rated high by the controllers. These findings are interpreted to mean that the Oceanic Data Link system imposes higher workload in cognitive demand rather than physical demand, but it does not affect their performance. In the analyst opinion, these NASA-TLX scores somewhat correspond with the human factors analyst recorded analysis and the controllers statements. During the assessment, the controllers stated that they were sufficiently performing their ATC duties with the Oceanic Data Link system. However, the controllers referenced several issues regarding the physical usage of the Oceanic Data Link System which makes several tasks cumbersome, imposes more time, and somewhat difficult to perform. These physical issues relate to the integration of the workstations and input of communication messages into the Oceanic Data Link system.
Link system These issues, although physical demands, do impede on controller mental workload that does support that the workload the controller experience is mainly cognitive in nature

Post-Run Questionnaires

At the end of each period of the air traffic control assessment, the participant proceeded into an adjacent room and completed the two questionnaires regarding the workload they experienced. The first questionnaire asked the participants to rate several various factors which contributed to their workload - such as traffic volume, weather, strip usage. The second questionnaire asked questions to rate the traffic volume and complexity of the previous traffic period. The questionnaire also asked 6 questions regarding perceived workload characteristics.

The responses to the first questionnaire are listed in table 5. The table shows the various workload factors identified, the median scores, and the SIQR for each factor. The rating scale for the workload factors were on a scale from “low” 1 through to “high” 10. As shown on the table 5, the highest factors rated by the controllers fell within a medium range of 5.5 to 5.0. These factors are: number of aircraft (Median 5.5), coordination with other sectors/facilities (Median 5.0), and strip usage (Median 5.0).

The following factors were rated between the scale of 4.5 to 3.0 indicating the controllers’ responses were moderate workload to more than very low workload. These factors are: weather (Median 4.5), sector geometry/complexity (Median 4.5), housekeeping (Median 4.5), accepting transfer of control (Median 4.0), giving transfer of control (Median 4.0), monitoring and resolving conflicts (Median 4.0), and console layout (Median 4.0), number of altitude changes (Median 4.0), traffic mixture (Median 4.0), pilot route/altitude deviation (Median 3.0), use of ODL keyboard (Median 3.0), using strips and ODL concurrently (Median 3.0), and use of ODL trackball (Median 3.0).

The lowest factors rated by the controllers and, thus, not major contributors to their workload fell within the scale of 1.0 to 2.0. These factors are number of route changes (Median 3.0),...
2 0), pilot verbal response errors/delay (Median 2 0), a/c flight characteristics (Median 2 0),
monitor size and position (Median 2 0), number of airspeed changes (Median 1 0), giving
information transfer (Median 1 5), accepting information transfer (Median 1 0), and area
restrictions (Median 1 0)

Table 5  Workload Factors

<table>
<thead>
<tr>
<th>Workload Factor</th>
<th>Median</th>
<th>SIQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of aircraft</td>
<td>5 5</td>
<td>1 5</td>
</tr>
<tr>
<td>Number of route changes</td>
<td>2 0</td>
<td>2 5</td>
</tr>
<tr>
<td>Number of altitude changes</td>
<td>4 0</td>
<td>1 8</td>
</tr>
<tr>
<td>Number of airspeed changes</td>
<td>1 0</td>
<td>5</td>
</tr>
<tr>
<td>Weather</td>
<td>4 5</td>
<td>1 0</td>
</tr>
<tr>
<td>Pilot verbal response errors/delay</td>
<td>2 0</td>
<td>2 0</td>
</tr>
<tr>
<td>Pilot route/altitude deviations</td>
<td>3 0</td>
<td>1 8</td>
</tr>
<tr>
<td>Accepting transfer of control</td>
<td>4 0</td>
<td>75</td>
</tr>
<tr>
<td>Giving transfer of control</td>
<td>4 0</td>
<td>1 25</td>
</tr>
<tr>
<td>Accepting information transfer</td>
<td>1 0</td>
<td>5</td>
</tr>
<tr>
<td>Giving information transfer</td>
<td>1 5</td>
<td>1 0</td>
</tr>
<tr>
<td>Housekeeping (moving data blocks, removing strips)</td>
<td>4 5</td>
<td>1 25</td>
</tr>
<tr>
<td>Traffic mixture, patterns or unique a/c types</td>
<td>4 0</td>
<td>1 5</td>
</tr>
<tr>
<td>Sector geometry/complexity</td>
<td>4 5</td>
<td>1 8</td>
</tr>
<tr>
<td>Area restrictions</td>
<td>1 0</td>
<td>5</td>
</tr>
<tr>
<td>A/c flight characteristics (climb, descend, airspeed)</td>
<td>2 0</td>
<td>1 0</td>
</tr>
<tr>
<td>Coordination with other sectors/facilities</td>
<td>5 0</td>
<td>1 6</td>
</tr>
<tr>
<td>Monitoring and resolving conflicts</td>
<td>4 0</td>
<td>1 8</td>
</tr>
<tr>
<td>Console layout</td>
<td>4 0</td>
<td>1 5</td>
</tr>
<tr>
<td>Monitor size and position</td>
<td>2 0</td>
<td>5</td>
</tr>
<tr>
<td>Strip usage</td>
<td>5 0</td>
<td>1 0</td>
</tr>
<tr>
<td>Use of ODL keyboard</td>
<td>3 0</td>
<td>1 5</td>
</tr>
<tr>
<td>Using strips and ODL concurrently</td>
<td>3 0</td>
<td>1 5</td>
</tr>
<tr>
<td>Use of ODL trackball</td>
<td>3 0</td>
<td>1 5</td>
</tr>
</tbody>
</table>
In the analyst opinion, these findings correspond with the human factors analyst recorded analysis and the controllers' statements during the assessment. The problems that several controllers pointed out regarding the system included that it was not well-integrated within the oceanic workspace environment—neither physically or functionally. The main issue is that the ISD Workstation, flight strip bays, and the Oceanic Data Link system do not work well together since they were developed separately. That is, the controller must relocate his attention (physically and mentally) between the different systems to accomplish a task. Also, in the analyst's opinion, the strip usage median also correspond with the controllers' statements. That is, the controllers must divert their attention from the Oceanic Data Link system and enter redundant information onto the flight strips. Furthermore, the ways the controllers must enter certain Minimum Operational Performance Standards (MOPS) messages into the Oceanic Data Link system was redundant and cumbersome. As indicated on the median response above, these messages are due to coordination with other sectors/facilities. This task was not part of their previous ATC tasks, but now must be performed with the Oceanic Data Link system.

The second questionnaire was designed to elicit their perceived workload experienced during the evaluation period on such factors as the traffic volume, traffic complexity, sector complexity and their overall workload. The rating scale for these factors were on a scale from 1 “low” to 5 “moderate” to 10 “high.” Table 6 shows the median and SIQR for the traffic volume and complexity section of the questionnaire. Looking at the median scores on table 6, it appears that controllers rated traffic volume (Median 3.5), traffic complexity (Median 3.5), and sector complexity (Median 3.5) as factors which fall within a slightly less than moderate workload for the traffic period ended. For overall workload, the median was 3.0 which indicates that overall, controllers experienced less than moderate workload for the traffic period ended. As described previously, due to the weather pattern over the Pacific Ocean, the aircraft were traversed to other adjacent sectors causing the normal high traffic period to be one of an average traffic period.
Therefore, these ratings may reflect the somewhat less than moderate workload for the traffic period ended.

Table 6  Traffic Volume & Complexity

<table>
<thead>
<tr>
<th>Traffic Volume &amp; Complexity Factors</th>
<th>Median</th>
<th>SIQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic volume</td>
<td>3 5</td>
<td>1 6</td>
</tr>
<tr>
<td>Traffic complexity</td>
<td>3 5</td>
<td>1 3</td>
</tr>
<tr>
<td>Sector complexity</td>
<td>3 5</td>
<td>1 8</td>
</tr>
<tr>
<td>Your overall workload</td>
<td>3 0</td>
<td>1 6</td>
</tr>
</tbody>
</table>

Respondents indicated that on a scale of 0% to 50% to 100% busyness, participants indicated that the fraction of time that they were busy was slightly below 50% (Mean 4 4) of the time. Table 7 shows the median and SIQR for the perceived workload characteristic portion of the questionnaire. As shown on table 7, in response to the amount of effort they spent calculating, estimating, planning, and problem solving (on a scale of 1 “low” to 10 “high”), controllers median response was 4 0 indicating the amount of effort they experienced was moderate. However, in response to the question if they found this control period stressful, their median response was 2 0 (on a scale of 1 “relaxing” to 10 “stressful”). Again, these less than moderate score responses may be due to the less than average workload traffic period during the traffic period ended.

The last three questions were questions regarding the workload usage of the Oceanic Data Link system. Controllers’ median response to the question that the system lowers their workload, controllers median response was 5 0 (on a scale of 1 “agree” to 10 “disagree”). This indicates that the controllers were not in agreement, but also not in disagreement that the systems lowers their workload. Controllers median response was 4 5 to the question that the system makes traffic handling more efficient, indicating that they do not agree, but also do not disagree that the systems makes handling traffic more efficient. Finally, responses to the last question that the use
of the system impacts their non-messaging tasks (e.g., FPS), the median response was 5.5. This indicates that the controllers were slightly more in disagreement that the system negatively impacts their non-messaging tasks.

Table 7 Perceived Workload Characteristics

<table>
<thead>
<tr>
<th>Perceived Workload Characteristics</th>
<th>Median</th>
<th>SIQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>How much effort did you spend calculating, estimating, planning, and problem solving during this period?</td>
<td>4.0</td>
<td>1.5</td>
</tr>
<tr>
<td>Did you find this control period stressful?</td>
<td>2.0</td>
<td>1.0</td>
</tr>
<tr>
<td>I feel that the use of ODL lowers my workload?</td>
<td>5.0</td>
<td>1.0</td>
</tr>
<tr>
<td>I feel that the use of ODL makes handling traffic more efficient</td>
<td>4.5</td>
<td>1.0</td>
</tr>
<tr>
<td>I feel that the use of ODL negatively impacts my non-messaging tasks (e.g., FPS)</td>
<td>5.5</td>
<td>1.8</td>
</tr>
</tbody>
</table>
CONCLUSIONS

As reflected by the NASA-TLX ratings, the use of the Oceanic Data Link system on controller workload imposes higher Mental Demand and Temporal Demand followed closely by Effort and Frustration than physical demand and performance. The human factors analyst’s notes and controllers’ on-line statements also reflect these results. The analysis indicated that the issues relating to their increased cognitive workload are due to physical and functionality issues relating to the oceanic controller workload environment. Also noted are certain messages are difficult to perform because of their complexity to compose these messages and which were not performed in the TP environment. Lastly, the integration of the Oceanic Data Link system is not fully integrated with other sub-system components (i.e., ISD and FPS) which makes it difficult to physically and functionally perform some of their duties. Since the traffic complexity and traffic volume were not comparable to the normal high traffic period, further studies are recommended to study the impact of the Oceanic Data Link system on controller workload during high traffic period.
REFERENCES

workload a comparison of assessment techniques addressing sensitivity and intrusion issues
Ergonomics, 27(10), 1033-1050

Casali, J G & Wierwille, W W (1993) A Comparison of Rating Scale, Secondary-
Task, Physiological, and Primary-Tasks Workload Estimation Techniques in a Simulated Flight
Task Emphasizing Communications Load Human Factors, 25(6), 623-641

Edwards, M B, Fuller, D K & Vortac, O U (1995) The role of flight progress strip
in en route air traffic control a time-series analysis International Journal Human-Computer
Studies, 43, 1-13

Eggemeir, F T & Wilson, G A (1991) Performance-based and subjective assessment
of workload in multi-task environments In D L Damas (Ed ), Multiple task performance
London Taylor & Francis

Results of Empirical and Theoretical Research In P A Hancock and N Meshkati (Editors),
Human Mental Workload North-Holland Elsevier Science Publishers B V

Hart, S & Wickens C (1990) Workload Assessment and Prediction In H R Booher
(Ed ) MANPRINT An approach to system integration Van Nostrand Reinhold, New York

of Operator Workload Methodologies (Procurement Instrument Identification No MDA903-86-
C-0384) United States Army Research Institute for the Behavioral and Social Sciences

and Evaluation of Applied Research Techniques for Documenting Cognitive Processes in Air

Task Load Index (V 10) Moffett Field, CA NASA Ames Research Center


Appendix A
Understanding your Participation

in the

ODL Human Factors Evaluation Plan

Please read this agreement carefully

Purpose
This analysis effort has been requested by ARU-100 and being carried out by __________. The purpose of this study is to assess the Human Factors issues of ODL. We are NOT evaluating you or your capabilities.

Participant number
You will be randomly assigned a participant number. We will be using this number to help us in organizing our data. We would like you to write down and remember this number, because we will be asking you to enter that same number later during the data collection activities.

Information Collected
The observers will record information about how you use ODL, e.g., where it was easy or difficult to use. You will be asked to fill out a questionnaire, answer questions about the workload you experienced, and participate in a brief interview. The information you will give us, along with the information we collect from other participants, will be used in making recommendations for improving the ODL design.

Waiver
Your "on-line" work on the oceanic control room floor will be video and audio-taped. By signing this form, you give your consent to __________ to use your verbal statements, and your "on-line" work, but not your name, for evaluation and demonstration.

Comfort
You may take a break at any time you wish, just inform the analyst that you would like to do so.

Confidentiality
Please understand that your participation in this study is strictly voluntary and your right to privacy will be protected. Your participation in this study will be anonymous and strictly confidential.

Freedom to withdraw
You may withdraw from this study at any time.

If you agree with these terms, please indicate your acceptance by signing below.

Signed _________________________________ Date ____________
Participant Entry Questionnaire

Instructions  The purpose of this questionnaire is to obtain some information regarding your experience and current attitudes. The information will be used to describe participants as a group. All responses to these questions are anonymous. We ask you to be as accurate in your response as possible.

1. How many years have you been an oceanic air traffic controller? ______
   a) at ZOA? _____
   b) at other facilities? _____ (please list) ____________________________

2. How many years have you been a controller? _____

3. What oceanic sectors have you worked at ZOA? ____________________________

4. Have you received training on ODL? Yes _____ No _____
   If so, when? Date __________
   What type of training did you receive? Please check each that apply
   Classroom ____ OJT ____ Other ______________________

5. When did you most recently receive refresher training? Years _____ Months _____
   What type? Classroom ____ OJT ____ Other _____

6. How much experience do you have using ODL? Years _____ Months _____

7. What do you think is your level of using ODL? beginner ____ intermediate ____
   advanced ____ expert ____

8. What do you feel is your computer experience level? beginner ____ intermediate ____
   advanced ____ expert ____

9. What do you think is your level of training? beginner ____ intermediate ____
   advanced ____ expert ____

10. What do you think is your level of entering numeric data? beginner ____ intermediate ____
    advanced ____ expert ____

The next series of questions will ask you to examine statements of opinions and determine to what extent you agree or disagree with them. Circle the one number which best describes your level of agreement with each statement.

<table>
<thead>
<tr>
<th>Circle one</th>
<th>Strongly agree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>11. &quot;I freely volunteered to participate in this study&quot;</td>
<td>1   2   3   4   5   6   7   8   9   10</td>
<td></td>
</tr>
<tr>
<td>12. &quot;I know a great deal about ODL&quot;</td>
<td>1   2   3   4   5   6   7   8   9   10</td>
<td></td>
</tr>
<tr>
<td>13. &quot;ODL is important for airspace system development&quot;</td>
<td>1   2   3   4   5   6   7   8   9   10</td>
<td></td>
</tr>
<tr>
<td>14. &quot;I need additional training for ODL.&quot;</td>
<td>1   2   3   4   5   6   7   8   9   10</td>
<td></td>
</tr>
</tbody>
</table>
15. "The level of training I receive on new ODL features is adequate."

Is there anything else we (___________) should know regarding your participation in this assessment?
Appendix C
Today you will be participating in a Human Factors study of Oceanic Data Link (ODL). We ask you to work as you normally would and let us know what you are verbally thinking and doing as you go along. Your verbal feedback during this session is vital, it will help identify Human Factors issues associated with ODL. Please comment freely on anything you observe. Positive and negative comments are welcome!

**GUIDELINES**
While you are working

- **Think out loud**  Tell us what's going on  For example

  "I am now performing a traffic search  

  "I am now using ODL to compose a MOPS message  

- **If you have trouble with a task**

  1  Tell us what is happening and why it's different than you expected
  2  Tell us what you did to try to solve the problem

- **Work at your normal pacer**  Remember, we're evaluating ODL, NOT your performance!  If you have trouble, it probably means there's something that needs improvement

**Thank you for your participation.**
Appendix D
Workload Questionnaire

Regarding the previous assessment, please rate the factors which contributed to your workload.

<table>
<thead>
<tr>
<th>Workload Contributions was (circle 1)</th>
<th>Workload Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very low</td>
<td>Very high</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Number of aircraft</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Number of route changes</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Number of altitude changes</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Number of airspeed changes</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Weather</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Pilot verbal response errors/delay</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Pilot route/altitude deviations</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Accepting transfer of control</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Giving transfer of control</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Accepting information transfer</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Giving information transfer</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Housekeeping (moving data blocks, removing strips)</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Traffic mixture, patterns or unique a/c types</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Sector geometry/complexity</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Area Restrictions</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>A/C flight characteristics (climb, descend, airspeed)</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Coordination with other sectors / facilities</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Monitoring and resolving conflicts</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Console layout</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Monitor size and position</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Strip usage</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Use of ODL keyboard</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Using strips and ODL concurrently</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Use of ODL trackball</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td>Other (please specify):</td>
</tr>
</tbody>
</table>

Comments:
Appendix E
**ODL Workload Assessment**

**Workload Questionnaire**

**Traffic Volume & Complexity**
For the traffic period just ended, rate the following:

|------------------|---------------------|---------------------|--------------------------|

**Perceived workload characteristics**

What fraction of the time were you busy during the period you were controlling?

<table>
<thead>
<tr>
<th>1. 0%</th>
<th>2. 50%</th>
<th>3. 100%</th>
</tr>
</thead>
</table>

How much effort did you spend calculating, estimating, planning, and problem solving during this period?

<table>
<thead>
<tr>
<th>1. Low</th>
<th>2. High</th>
</tr>
</thead>
</table>

Did you find this control period stressful?

<table>
<thead>
<tr>
<th>1. Agree</th>
<th>2. Disagree</th>
</tr>
</thead>
</table>

I feel that the use of the ODL lowers my workload.

<table>
<thead>
<tr>
<th>1. Agree</th>
<th>2. Disagree</th>
</tr>
</thead>
</table>

I feel that the use of ODL makes handling traffic more efficient.

<table>
<thead>
<tr>
<th>1. Agree</th>
<th>2. Disagree</th>
</tr>
</thead>
</table>

I feel that the use of ODL negatively impacts my non-messaging tasks (e.g., FPS).

<table>
<thead>
<tr>
<th>1. Agree</th>
<th>2. Disagree</th>
</tr>
</thead>
</table>

Comments:

_________________________________________________________________
Appendix F
ODL Workload Assessment
Instructions for completing the rating scale

We are interested in the experiences you had during this assessment. Right now we are going to describe the technique that will be used to examine your experiences. In the most general sense, we are examining the “workload” you experienced. Workload is a difficult concept to define precisely, but a simple one to understand generally. The factors that influence your experience of workload may come from the task itself, your feelings about your own performance, how much effort you put in, or the stress and frustration you felt.

Rating Scale

One way to find out about workload is for you to describe the feeling you experienced. Because workload may be caused by many different factors, we would like you to evaluate several of them individually rather than lumping them into a single global evaluation of overall workload. This set of six rating scales was developed for you to use in evaluating your experiences during different tasks. Please read the descriptions of the scales carefully. If you have a question about any of the scales in the table, please ask me about it. It is extremely important that they be clear to you. You may keep the descriptions with you for reference during the experiment.

Procedure

Six rating scales will be presented to you on the screen. You will evaluate your task by marking each scale at the point that matches your experience. Each line has two endpoint descriptors that describe the scale. Note that “effort” goes from “good” on the left to “bad” on the right. This order has been confusing for some people. Move the arrow with the right and left arrow keys until it points at the desired location. Stop it by pressing the up arrow key. Press the down arrow key to enter your selection. Please consider your responses carefully. Consider each scale individually. Your ratings will play an important role in the assessment being conducted. Your active participation is essential to the success of this assessment, and is greatly appreciated.
ODL Workload Assessment
Instructions for assessing the importance of the rating scales

Throughout this assessment, the rating scales are used to assess your experience using the ATC system. Scales of this sort are extremely useful, but their utility suffers from the tendency people have to interpret them in individual ways.

The evaluation you are about to perform is a technique that has been developed by NASA to assess the relative importance of six factors in determining how much workload you experienced. Please read the definition for each of these factors (provided on the next page). If you need further classification, please ask me. The procedure is simple.

You will be presented with a series of pairs of rating scales titles (for example, Effort vs Mental Demands) and asked to choose which of the items was more important to your experience of workload during the evaluation period. Each pair of scale titles will appear separately on the screen. Select the Scale Title that represents the more important contributor to workload during the evaluation period.

Press “1” to select the top item in the pair, and “2” to select the bottom item. If you change your mind, press backspace to erase your choice. Press carriage return to enter it. After the carriage return, a new pair of scale titles will appear.

After you have finished the entire series, we will be able to use the pattern of your choices to create a weighted combination of the ratings from the task into a summary workload score. Please consider your choices carefully and make them consistent with how you used the previous rating scales. Don’t think that there is any correct pattern. We are only interested in your opinions.

If you have any questions, please ask them now.

Once again, thank you for your participation.
**ODL Workload Assessment**  
NASA-TLX rating scale description (NASA-Ames Research Center, 1986)

<table>
<thead>
<tr>
<th>Title</th>
<th>Endpoints</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental demand</td>
<td>Very Low/Very</td>
<td>How mentally demanding was it for you in controlling the traffic using the ATC system?</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>Physical demand</td>
<td>Very Low/Very</td>
<td>How physically demanding was it for you to control traffic using the ATC system?</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>Temporal demand</td>
<td>Very Low/Very</td>
<td>How hurried or rushed was the pace for you to control the traffic using the ATC system?</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td>Perfect/Failure</td>
<td>How successful were you in accomplishing what you were asked to do by using the ATC system?</td>
</tr>
<tr>
<td>Effort</td>
<td>Very Low/Very</td>
<td>How hard did you have to work to accomplish your level of performance?</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>Frustration</td>
<td>Very Low/Very</td>
<td>How insecure, discouraged, irritated, and annoyed were you using the ATC system?</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td></td>
</tr>
</tbody>
</table>