

**Ground-Based Study of an Expert System for Human Assistance on the
STS-95 Sleep and Respiration Experiment**

by

Melinda Gallo Delaney

Bachelor of Science, Aerospace Engineering
University of Virginia, Charlottesville, VA, 1997

Submitted to the Department of Aeronautics and Astronautics
in partial fulfillment of the requirements for the degree of

Master of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

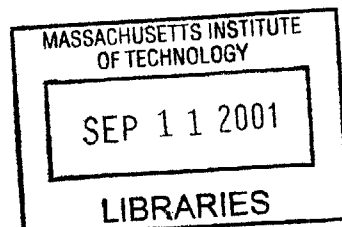
February 2001

© Massachusetts Institute of Technology, 2001. All Rights Reserved.

Author
Department of Aeronautics and Astronautics
December 7, 2000

Certified by
Laurence R. Young
Apollo Program Professor of Astronautics
Thesis Supervisor

Accepted by
Professor Wallace E. Vander Velde
Chair, Committee on Graduate Students
Department of Aeronautics and Astronautics



Aero

Ground-Based Study of an Expert System for Human Assistance on the STS-95 Sleep and Respiration Experiment

by

Melinda Gallo Delaney

Submitted to the Department of Aeronautics and Astronautics on November 15, 2000, in partial fulfillment of the requirements for the Degree of Master of Science in Aeronautics and Astronautics

Abstract

The expert system, *Principal Investigator-in-a-Box*, or [PI], was designed to assist astronauts or other operators in performing experiments outside their field of expertise. It helped the crew calibrate instruments for a Sleep and Respiration Experiment that flew on the Space Shuttle Missions STS-90 and STS-95. [PI] displays electrophysiological signals in real time, alerts astronauts when poor signal quality is detected, and advises astronauts how to restore good signal quality. A ground-based study sought to assess the utility of on-board expert systems, in general, for performing experiments and troubleshooting complex instrumentation systems.

Thirty subjects received training on the sleep instrumentation and the [PI] interface. Each subject was then tested on two separate sessions with the instrumentation, once with [PI] assistance and once without. Results indicate a beneficial effect of [PI] in reducing anomaly troubleshooting time. Further, questionnaires showed that most subjects preferred monitoring the [PI] display, together with monitoring waveforms, to monitoring of the waveforms alone.

Thesis Supervisor: Laurence R. Young, ScD.

Title: Apollo Program Professor of Astronautics

This research was supported by the National Space Biomedical Research Institute, NASA Cooperative Agreement NCC 9-58, and the NASA Ames Research Center, grant number NCC 2-570.

Acknowledgements

The completion of this thesis would not have been possible without the help, guidance, and inspiration of several people throughout my year and a half at MIT:

Professor Young, for allowing me to be a part of the final chapter of the [PI] saga. I've learned so much from my research here, and your help and support have been invaluable.

Allen Atamer, what can I say? It's been an unforgettable experience. I think the MVL will truly be a different place without the Allen-Mindy dynamic. =)

Dr. Alan Natapoff, for all of your guidance, support, and words of wisdom through the last year and a half. Thanks for everything!

Dennis Heher, who helped Allen and I get [PI] up and running early in the Fall of 1999. I hope that your new endeavour takes you far!

Marsha Warren, for answering all of my questions - I know you thought they'd never end! Thanks so much for your help with editing the [PI] reports, and for lending an ear to listen to my neverending commentary on life in the MVL.

Will Fournier, our dedicated UROP. I don't have to tell you, because you already know - without your mediating skills last January, I don't know how we would've finished our experiment on time (early, even!). Best of luck to you in the future.

To Luca Callini, my email pen-pal, and Susanne Essig - thanks for helping me out to learn the ropes, and filling me in on all the "old" MVL gossip. It's been great knowing you both! Best of luck in the future. (I'll see you in Houston, right?)

To the Brigham & Women's hospital sleep team: Dr. Charles Czeisler, Eymard Riel, Dr. Rod Hughes, and Dr. Ken Wright. Thanks for all of your help with learning the ins and outs of sleep signal analysis and finding new SleepNets.

To Jessica Marquez - for keeping me sane, being one of our "volunteer" subjects for our experiment at the last minute, and for the great company all of those lonely nights when my hubby was off at class.

To all of our guinea pigs: especially Dr. Heiko Hecht and Jason Richards. Thanks for agreeing to be "instrumented" countless times. What would we have done without you?

To the rest of the MVLers for always being around for a good laugh, a drink at the Muddy, or a night on the town.

To my husband, Rich, for agreeing to be dragged all the way to MIT from sunny Florida so that I could pursue this crazy dream of mine. Thanks for putting up with me!

Finally, to my family: Mom, Dad, Mike and Michele, thanks for believing in me, your "rocket scientist."

Table of Contents

Abstract.....	3
Acknowledgements	5
Table of Contents	7
List of Figures	9
List of Tables	11
1 Introduction.....	13
1.1 Background Information.....	13
1.2 Current Use of [PI]-in-a-Box.....	14
1.3 Past Conclusions on [PI] for the Sleep and Respiration Experiment	16
1.4 January 2000 Study.....	18
2 Overview of the STS-95 Sleep Experiment.....	19
2.1 Apparatus	19
2.2 Signal Overview	22
2.3 Signal Anomalies.....	26
3 Methods	29
3.1 Participants.....	29
3.2 Experimental Design.....	29
3.3 Procedures.....	31
4 Data Analysis and Results	37
4.1 Definitions	37
4.2 User Action Analysis	39
4.3 Analysis	40
5 Discussion and Conclusions	55
5.1 Discussion.....	55
5.2 Conclusions.....	56
5.3 Recommendations.....	57
6 References.....	59
Appendix A NASA Troubleshooting Guide.....	61
Appendix B Pre-experiment questionnaire.....	73
Appendix C Informed Consent Forms.....	75
C.1 INFORMED CONSENT FORM FOR TEST SUBJECTS	75
C.2 INFORMED CONSENT FORM FOR SLEEP SUBJECTS	77
Appendix D Subjective questionnaire	79
Appendix E Subject Comments.....	81

List of Figures

Figure 1.1: Electrophysiological Signals as Displayed on [PI] Interface	15
Figure 1.2: Senator John Glenn Donning Sleep Instrumentation Aboard STS-95	17
Figure 2.1: Electrode Net, E*Net, or SleepNet.....	20
Figure 2.2: Digital Sleep Recorder	21
Figure 2.3: Hardware Schematic	21
Figure 2.4: Electrode Locations.....	23
Figure 2.5: Example of Alpha Activity in EEG Signal	24
Figure 2.6: The Electro-oculogram Eye Movements - Left and Right	25
Figure 2.7: The Electro-oculogram Eye Movements - Up and Down.....	25
Figure 2.8: EMG Relaxed and Clenched Jaw	26
Figure 2.9: Example of a Popping Signal on the C4/A1 Electrode	27
Figure 2.10: Example of a Noisy Electro-oculogram Signal	27
Figure 2.11: Example of Saturated Signal.....	28
Figure 2.12: Example of a Flat EMG Signal	28
Figure 3.1: Experiment Setup	31
Figure 3.2: Close-up of Gray Checkbox and Red and Green LEDs	32
Figure 3.3: Interface Used to Input System State Description	33
Figure 3.4: Gray Checkbox with 'X' After Subject Entry	33
Figure 3.5: Interface Used to Input Problem Description.....	35
Figure 3.6: Interface Used to Input Solution Description.....	36
Figure 4.1: Timeline Showing Detection, Troubleshooting and "Fix" Times.....	37
Figure 4.2: Troubleshooting Time is Significantly Improved with [PI] Help.....	44
Figure 4.3: Detection Time by Fault Type	47
Figure 4.4: Detection Time Histogram for Group 1 Day 1 (With [PI]).....	50
Figure 4.5: Detection Time Histogram for Group 1 Day 2 (Without [PI])	51
Figure 4.6: Histogram of Detection Time for Group 2 Day 1 (Without [PI])	52
Figure 4.7: Histogram of Detection Time for Group 2 Day 2 (With [PI])	52
Figure 4.8: Subjective Questionnaire Results, Females	53
Figure 4.9: Subjective Questionnaire Results, Males	54

List of Tables

Table 3.1: Experimental Group Design	30
Table 3.2: Training Session Schedule.....	30
Table 3.3: Command or Questions to Ignore or Avoid	35
Table 4.1: Breakdown of User Actions	39
Table 4.2: Main Effects on T_{TS}	43
Table 4.3: Day x Group Table	43
Table 4.4: Main Effects on T_{TS}	45
Table 4.5: Median Test for T_D , Fault Types 0, 1, and 2 by Group.....	48
Table 4.6: Median Test for T_D , Fault Types 0, 1, and 2 by Group and Day	48
Table 4.7: Median Test for T_{TS} , Fault Types 1 and 2 by Group and Day	49
Table 4.8: Significance of Fault Type 1 and 2 for Detection Time	49

Chapter 1

Introduction

In a typical space flight, astronauts must learn a large number of scientific experiments as well as the flight operational procedures. Astronauts' lack of initial experience with any particular experiment is made more costly by the long delays between their training and the experiment's execution. This human factors problem will become increasingly critical in the era of the International Space Station (ISS), with its long flight intervals. Stress, fatigue, and high mental workload can also degrade astronaut performance. Principal Investigators can rarely accompany their experiments into space; and ground-to-air contact with the astronauts performing the experiment is not always possible. One way to approach this problem is to provide a computer decision aid, such as an expert system, to help guide the operator in real time. An expert system uses heuristic rules based on a real expert's knowledge to reason and assist untrained users. The study outlined in this thesis is unique. It is the first formal evaluation of an expert system designed for a space life sciences experiment. The problem of complex troubleshooting and extended time between training and execution is common to many human operator situations on earth.

1.1 Background Information

The original version of the expert system, *Principal Investigator-in-a-Box*, also known as the Astronaut Science Advisor (ASA), is the first documented attempt to use a biomedical diagnostic expert system on a space mission [4]. *Principal Investigator-in-a-Box*, or [PI], was designed to assist astronauts in performing experiments outside their expertise. Laurence R. Young, at the Massachusetts Institute of Technology (MIT), conceived the initial [PI] concept. The software was used to assist astronauts in conducting the "Rotating Dome" visual-vestibular interaction experiment on the STS-40 ground experiment and

STS-58 Space Life Sciences 2 (SLS-2) Space Shuttle mission in 1993. This version of [PI] provided data collection capabilities, as well as protocol assistance, scheduling, and protocol modification suggestions. An additional feature consisted of an "interesting data" filter, designed to perform quick-look data analysis and report any unexpected findings to the astronauts during the experiment. Although crew feedback on this demonstration was positive, no data was available concerning the performance of [PI] or the correctness of the advisories that it issued.

The Rotating Dome experiment was a successful implementation of the ASA during a space flight experiment. Extending upon this success, MIT and NASA Ames Research Center collaborated on the development of a new version of [PI]. Dennis Heher, at NASA Ames Research Center, was the core [PI] software developer. The second version of [PI] was used in conjunction with the "Sleep, Respiration and Melatonin in Microgravity" experiment (commonly referred to as the Sleep and Respiration Experiment), led by Dr. Charles Czeisler of Brigham and Women's Hospital (BWH, Boston, MA) and Dr. John West of the University of California, San Diego. The experiment flew aboard the STS-90 (Neurolab) Space Shuttle mission in April-May 1998 and on STS-95 in October 1998, with Senator/Astronaut John Glenn.

1.2 Current Use of [PI]-in-a-Box

The most recent version of [PI] displays sleep and cardiorespiratory electrophysiological signals in real time, alerts astronauts via indicator lights when a poor signal quality is detected, and advises astronauts how to restore good signal quality. The software interface consists of a waveform display for electrophysiological (EP) signals such as the electroencephalogram (EEG, commonly known as brain waves), electro-oculogram (EOG, or eye movements), and electromyogram (EMG, or jaw muscle activity). The interface has several components: a display window for EP signals, an array of indicator lights located

beside the waveforms, and a diagnostics window. The lights display [PI]'s assessment of signal quality. A red light denotes a poor quality signal; a green light indicates a good quality signal. The diagnostics messages window on the right part of the screen displays the procedures to follow to determine the cause of the problem. The [PI] interface is shown in Figure 1.1. In this case, the indicator light next to the C3_A2 EEG signal is red, indicating that the waveform is poor quality.

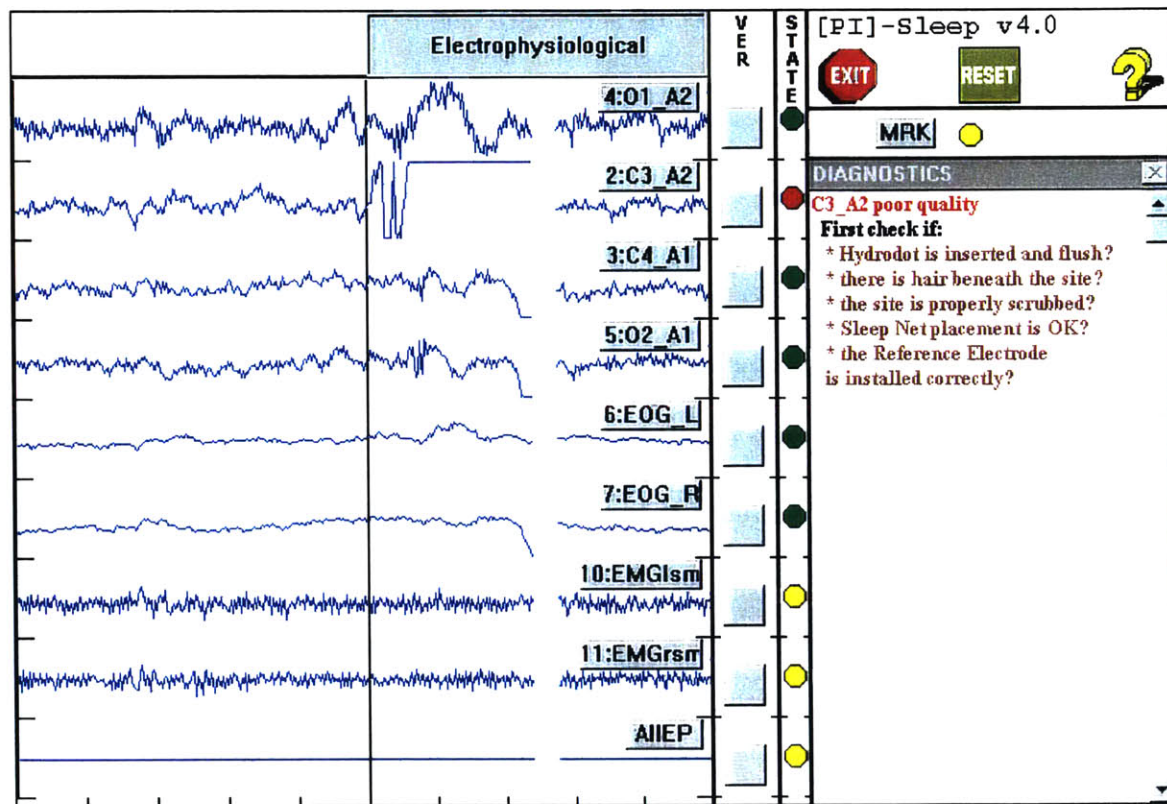


Figure 1.1: Electrophysiological Signals as Displayed on [PI] Interface

The [PI] software interface can also display cardiorespiratory signals in real time. However, this part of the software was not studied in the January 2000 experiment and is not discussed in this thesis. For more information regarding [PI] and the cardiorespiratory experiment that flew on STS-90 and STS-95, refer to Luca Callini's thesis [2].

1.3 Past Conclusions on [PI] for the Sleep and Respiration Experiment

Several earlier studies of [PI] for the space sleep and respiration study were conducted. A pilot study in January 1998 tested 12 subjects' ability to detect poor quality signals. The pilot study and observations of Neurolab and STS-95 data were used to evaluate [PI]'s ability to help with detection times for anomalous signals. The results of the pilot study showed that [PI] assistance reduced the detection time, though not by a statistically significant amount. Training, or the cross effect of [PI] and Day, was found to be significant. The study also found that the number of undetected anomalies was significantly lower when [PI] was available. Gender effects were also found to be significant for the detection task; female subjects outperformed male subjects by an average of 5 seconds.

The Neurolab and STS-95 data were comprised of signal recordings of the first few minutes of each instrumentation session. It was found that [PI] correctly detected 84% of the signal anomalies that were not saturated from the Neurolab data [3]. In the STS-95 data, [PI] correctly detected 86% of all signal anomalies. Overall, the cardiorespiratory indicator lights were the most reliable, while the electroencephalogram (EEG), and electro-oculogram (EOG) signals were the most prone to false alarms from [PI] indicator lights. One of the more famous astronauts, John Glenn, flew on the STS-95 mission and can be seen in Figure 1.2 donning the experiment hardware.



Figure 1.2: Senator John Glenn Donning Sleep Instrumentation Aboard STS-95

Phase I of the ground-based experiment tested the efficacy of [PI] for assisting "astronaut surrogates" in detecting realistic experiment artifacts. This phase of the ground-based study measured the time for thirty two test subjects to detect and identify the nature of a signal anomaly. Pre-recorded electrophysiological signals were played back for the subjects through a PC laptop computer, which either enabled or disabled the [PI] diagnostic routines. The subjects received 90 minutes of training on the electrophysiological sleep signals and [PI] software before monitoring sleep EP parameters. Half of the subjects received [PI] assistance in their first exposure to the tests, and half in their second.

Two distinct fault stimulus files were used for the analysis, File A and File B. Without [PI], subjects found it more difficult to correctly identify saturated and noise signals in File A than they did in File B. Consequently, [PI] was more beneficial when used with File A. Furthermore, [PI] had a significant effect on improving subjects' ability to cor-

rectly identify saturated signals in File A. Though subjects found File A more difficult to interpret than File B when [PI] was not active, the difficulty of the files became similar when [PI] was active. This result is encouraging; it suggests that any differences between files are essentially nullified when [PI] is active.

A statistically significant effect of day on response time was observed for popping signals. No other statistically significant effects on response time were found. This is consistent with the pilot study, which found only a significant effect of [PI] x Day on response time for saturated signals.

1.4 January 2000 Study

The January 2000 ground-based study sought to assess the utility of on-board expert systems, in general, for performing experiments and troubleshooting complex instrumentation systems. This study was labeled Phase 2 of the ground-based study. Specifically, we were interested in four major aims. First, to validate [PI]'s heuristic rules that are used to diagnose sleep signals. Second, to determine the required training time necessary to qualify astronaut surrogates in the use of [PI]. Third, to evaluate the effectiveness of [PI] as a decision aid. Fourth, to generalize from these results to the utility of on-board expert systems for experiment performance in space.

This study was completed via the collaboration of two full-time graduate Research Assistants, Allen Atamer and myself. Thirty two new subjects, or *astronaut surrogates*, were trained to read sleep EP signals and use the [PI] software. This experiment required the trained astronaut surrogates to not only monitor a sleep instrumentation setup and detect faults, but also to suggest a method to return the signal to good quality. The results presented in this thesis indicate the performance improvements noted with [PI]. Allen Atamer's thesis provides an in-depth analysis at the reliability of [PI]. [1]

Chapter 2

Overview of the STS-95 Sleep Experiment

The Sleep and Respiration Experiment flew onboard the STS-90 (NeuroLab) and STS-95 missions. The purpose of the Sleep Experiment was to record astronaut sleep EP signals during spaceflight. Investigators on the ground reviewed these recordings after the shuttle mission in order to score the sleep data. This information can be used to determine the quality and duration of sleep during spaceflight, which is of major concern because lack of sleep can lead to lower performance during a space mission. For long-duration spaceflight, as is necessary on the International Space Station, this cumulative lack of sleep can result in poor performance and affect an astronaut's abilities to carry out his/her duties. [PI] assisted the astronauts, after the hardware had been donned, in detecting and troubleshooting problems in the instrumentation setup before they retired to bed for the night.

The following section gives a brief overview of the hardware used in the spaceflight experiment and ground-based study. For the purposes of this study, only the sleep EP signals were used. The cardiorespiratory signal screen of [PI] and the corresponding hardware were not used, and therefore are not described.

2.1 Apparatus

The sleep experiment equipment includes the following:

- Electrode Net (e-Net), commonly referred to as simply the SleepNet: an elastic web-like cap containing 13 electrode sockets to record the electrophysiological signals: EEGs, EMGs and EOGs. (Physiometrix, Inc., North Billerica, Massachusetts, USA).
- Hydrodots: Ag/AgCl electrodes which fit into the Electrode Net's sockets and contact the skin. (Physiometrix, Inc., North Billerica, Massachusetts, USA).
- Digital Sleep Recorder (DSR): a device that converts the raw analog signals from the var-

ious electrodes and instrumentation to digital signals, which are then recorded onto a PCMCIA FlashRAM card. (Copyright 1996 Vitaport EDV System GmbH. Distributed by TEMEC instruments BV, The Netherlands).

- IBM Thinkpad laptop - Pentium class processor, with [PI] v4.04 software installed.

- Materials for preparing the electrode site (such as Q-Tip cotton swabs, Hydroprep solution [Physiometrix], adhesive pads, etc).

A diagram of the SleepNet is shown in Figure 2.1. It is made of an elastic material that allows it to fit properly on subjects with different size heads. It is designed to allow for easier and correct placement of the 13 electrodes. The SleepNet uses a 50-pin multi-connector to connect to the Digital Sleep Recorder.

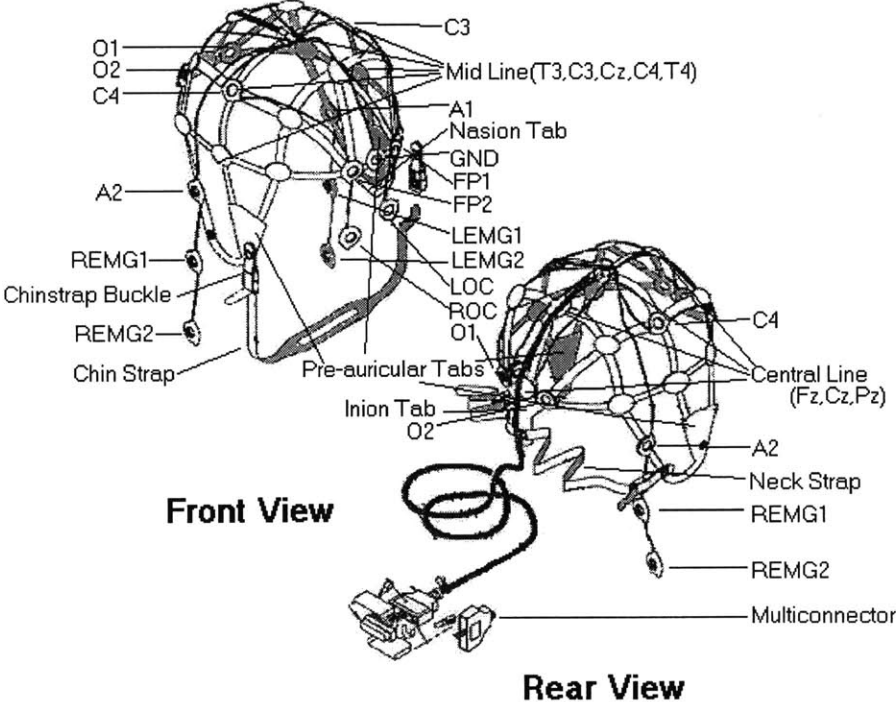


Figure 2.1: Electrode Net, E*Net, or SleepNet

The Digital Sleep Recorder (DSR) is shown in Figure 2.2, below. The top (blue) slice of the DSR is used for the SleepNet connection, the second (black) slice of the DSR is used for connection of the Respiratory Inductance Plethysmography (RIP) suit, which collects the Cardio-Respiratory signals. The bottom slice of the DSR is not used. For the January 2000 study, only the blue "sleep" slice of the DSR was in use. The DSR is powered by 4 AA batteries and records data on an 85 MB FlashRAM PCMCIA data card. The data can be directly downloaded from the card to the laptop, via insertion of the card in the laptop's PC Card Slot. The DSR is connected to the IBM Thinkpad via an optical RS-232, as shown in Figure 2.3, below.

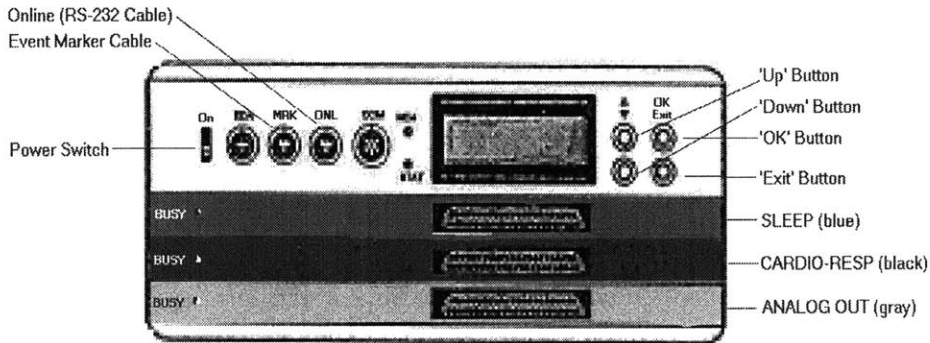


Figure 2.2: Digital Sleep Recorder

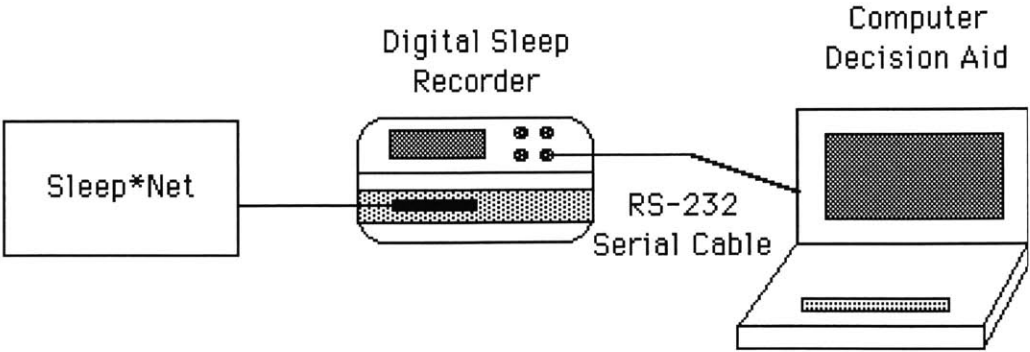


Figure 2.3: Hardware Schematic

2.2 Signal Overview

The SleepNet records three different types of electrophysiological (EP) signals: the electroencephalogram (EEG, or brain waves), the electro-oculogram (EOG, or eye movements) and the electromyogram (EMG, or jaw muscle activity). These signals are recorded via the 13 electrode locations on the head.

The international 10 20 system of electrode placement has been adopted in the field of sleep research as the standard for determining the location of each electrode [5]. This standard minimizes variation between different recording sessions. A schematic of this system is shown in Figure 2.4.

The nasion is located at the top of the nose, in the center of the nasofrontal structure. The inion is the protruding bone located on the back of the head, several inches above the neck. Marks are made at distances of ten and twenty percent between the subject's two preauricular points, located behind the ears, as well as around the circumference of the head. Electrodes are then placed at the intersections of the lines connecting these points. The ground electrode is located above the nasion on the forehead; it serves as a reference for all electrodes. Notice in Figure 2.4, below, that the left side of the head is labeled with even numbers and that the right side of the head is labeled with odd numbers.

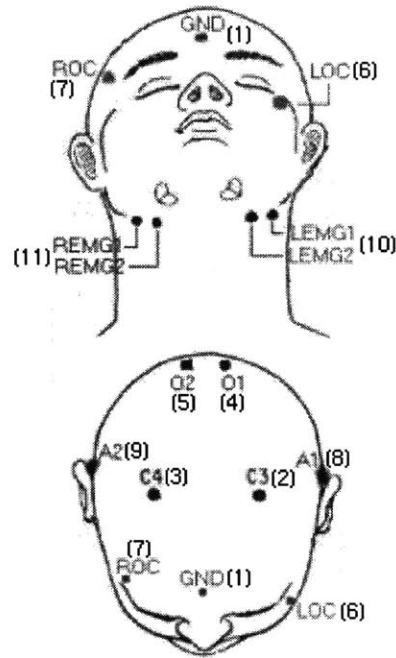


Figure 2.4: Electrode Locations

2.2.1 Electroencephalogram

The electroencephalogram (EEG) is the primary polysomnographic measure used in evaluating and scoring sleep data. The EEG signal uses 2 central electrodes (C4, C3) and 2 occipital electrodes (O2, O1). Only one central and one occipital electrode are necessary to score sleep data, therefore redundancy is built into the SleepNet system. The right side of the head (C3 and O1) are referenced to the opposite side's mastoid reference, the A2 electrode, and vice versa.

A pattern of a higher frequency signal (10-200 μV), known as alpha activity, becomes apparent when the subject's eyes are closed. While the subject is awake and relaxed with eyes open, the EEG is a mixed-frequency, low-amplitude signal. A standard amplitude range for the EEG signal is 10-200 mV. When the subject's eyes are closed, a rhythmic, higher frequency pattern becomes apparent in the EEG signal. This activity, known as alpha activity, is characterized by a frequency in the 8-12 Hz range. Alpha activity is gen-

erally most prominent in the occipital EEG, but is also discernible in the central EEG. An example of EEG activity is shown in Figure 2.5.

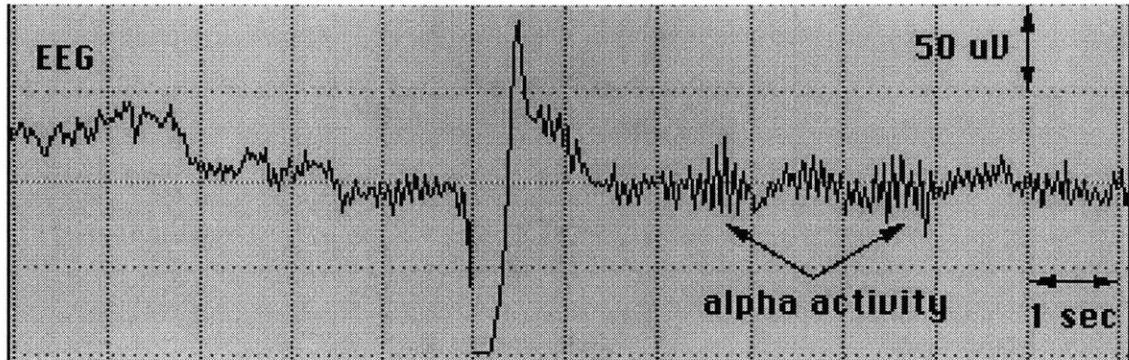


Figure 2.5: Example of Alpha Activity in EEG Signal

2.2.2 Electro-oculogram

The electro-oculogram (EOG) is recorded in sleep studies primarily to distinguish rapid eye movement sleep (REM) from non-rapid eye movement (NREM) sleep. The adopted convention for EP signals is to display a negative voltage signal as an upward signal deflection on the display, or "negative up." The human eyeball has a natural potential difference. The cornea, at the front of the eye, is positively charged with respect to the retina, at the back of the eye.

The placement of the electrode at the appropriate location above or below either eye must be correct. The right EOG is placed above the right eye, and the left EOG is placed below the left eye. Each EOG electrode is referenced to the mastoid reference electrode located behind the opposite ear (left EOG referenced to A2, and right EOG referenced to A1). If the electrode is placed below the right eye rather than above, voltage fluctuations will be opposite of those expected from those shown in Figure 2.6.

Two characteristics in EOG analysis are "up-and-down" eye movements and "left-and-right" eye movements. The movement of both eyes in either horizontal direction produces a negative voltage in one eye and a positive voltage in the other. Moving the eyes to the

left produce a downward deflection of the left EOG signal and an upward deflection of the right EOG signal. The reverse is true for the eyes moving to the right. These cases are shown in Figure 2.6, below.

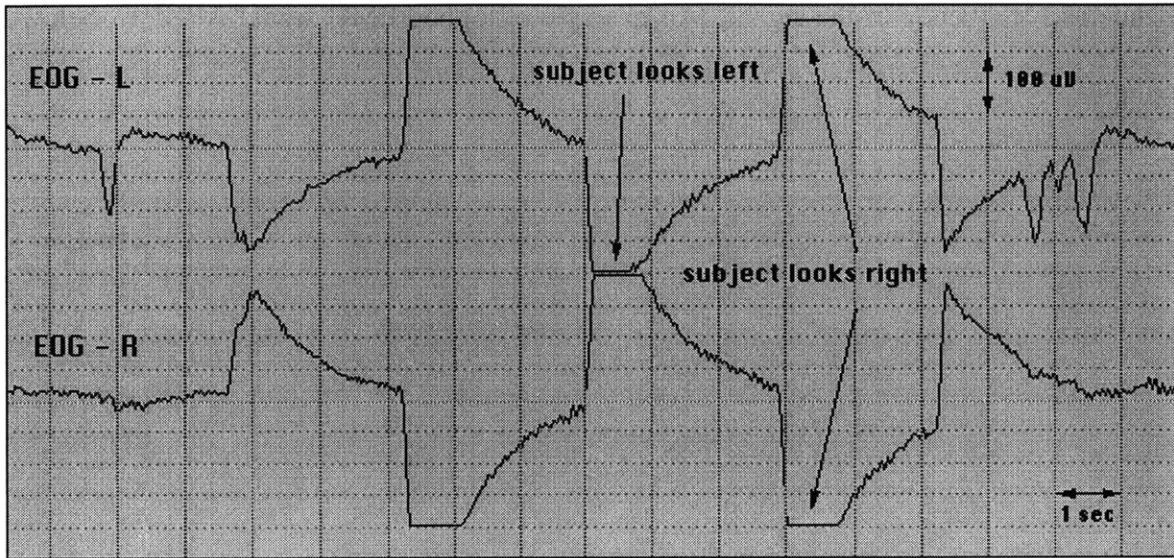


Figure 2.6: The Electro-oculogram Eye Movements - Left and Right

A similar characteristic is seen for up and down eye movements, but the deflections are less pronounced. Looking down causes an upward deflection in the right EOG. This is more difficult to see in the left EOG. Refer to Figure 2.7, below.

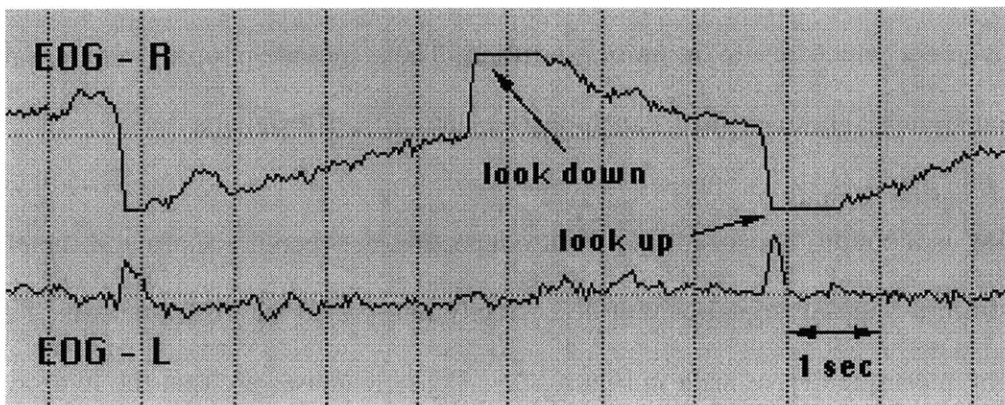


Figure 2.7: The Electro-oculogram Eye Movements - Up and Down

2.2.3 Electromyogram

The electromyogram (EMG) records muscle activity of the jaw, providing information on loss of muscle tone, which characterizes REM sleep. The EMG is a higher frequency signal than the EEG and EOG. Movements of the chin and jaw muscles increase the amplitude and frequency of the signal. An example of a good EMG signal is shown below in Figure 2.8.

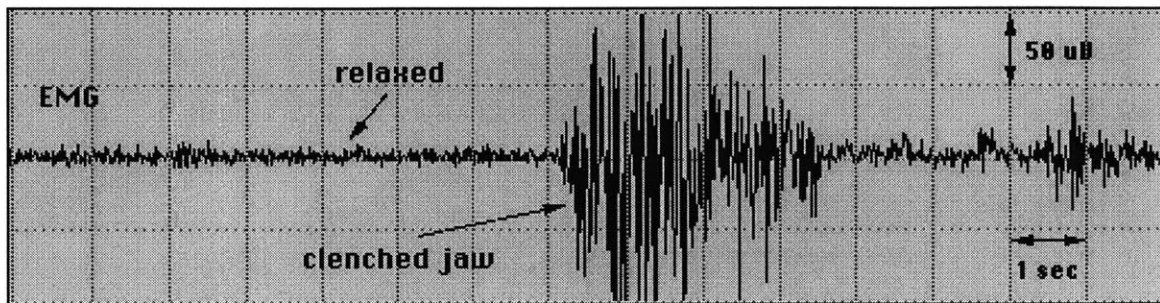


Figure 2.8: EMG Relaxed and Clenched Jaw

2.3 Signal Anomalies

Anomalous signals result from problems in the instrumentation. A bad signal can be caused by a poorly placed electrode, insufficient scrubbing of the electrode site, hair beneath the Hydrodot or hardware malfunction, or even the sleep subject moving, or yawning. Because the electrodes are all connected by elastic material, they are sensitive to movement which results in motion artifacts. These include popping, noise, saturation, and flat signals. These artifacts can appear on any of the EP signals.

2.3.1 Popping Signals

Popping is a condition caused by a poor connection between the scalp and the electrode. Data from a popping signal cannot be correctly interpreted by sleep analysts. The signal appears intermittently, as seen in Figure 2.9. Possible causes are that the hydrodot is not flush with the SleepNet, hair is beneath the hydrodot or incorrect SleepNet placement.

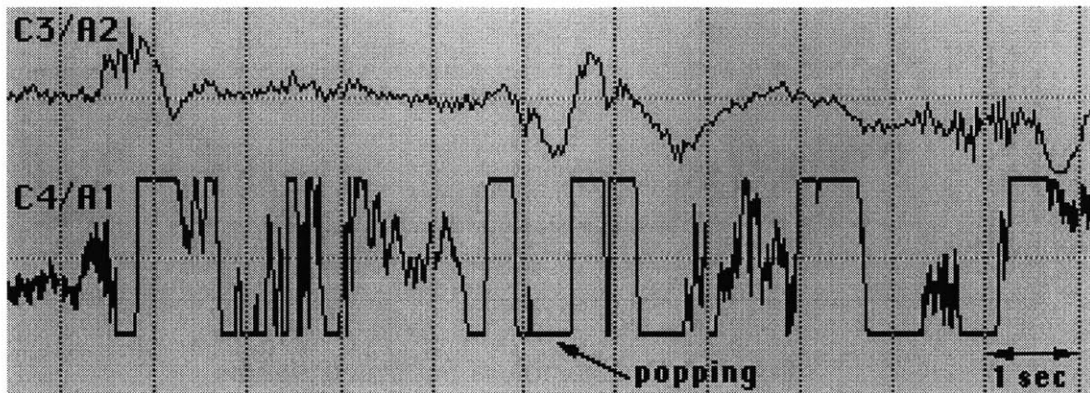


Figure 2.9: Example of a Popping Signal on the C4/A1 Electrode

2.3.2 Noise

Noise can be introduced into a signal as the result of a high resistance connection between the hydrodot and the skin, or by electromagnetic interference. Refer to Figure 2.10 for an example of a noisy signal. Noise is caused by hair beneath the hydrodot, an improperly scrubbed site, or 60 Hz interference from lights.

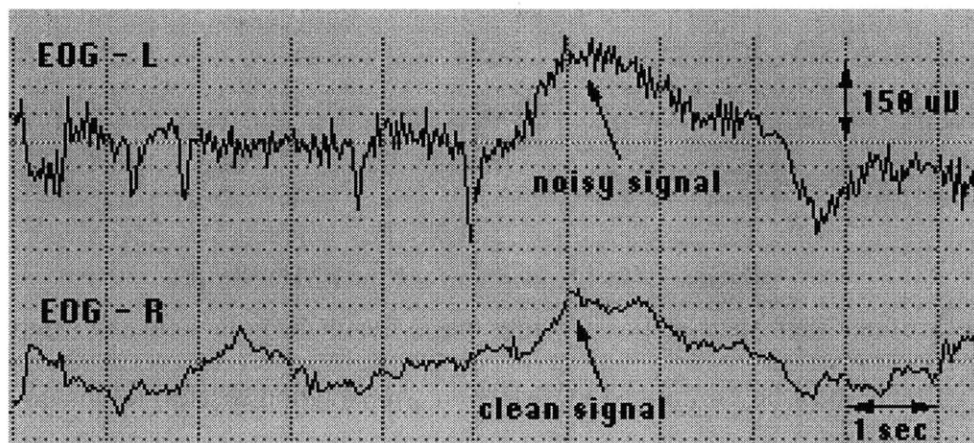


Figure 2.10: Example of a Noisy Electro-oculogram Signal

2.3.3 Saturation

A saturated signal occurs when the signal amplitude is out-of-bounds, driving the amplifier into saturation. This can occur when the electrode and hydrodot experience an excessive pressure force. Possible causes include a DC offset error or poor electrode location. Refer to Figure 2.11 for an example of a saturated signal.

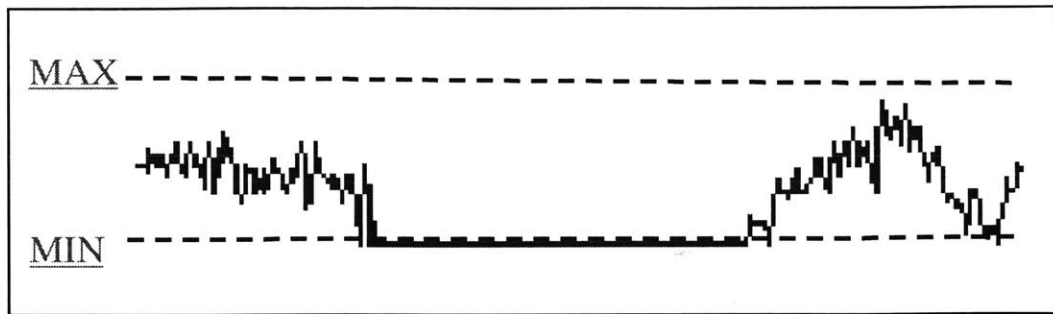


Figure 2.11: Example of Saturated Signal

2.3.4 Flat Signal

Improper electrode placement can result in a flat signal. Flat signals are caused by poor skin contact or the absence of a hydrodot. Refer to Figure 2.12, below.

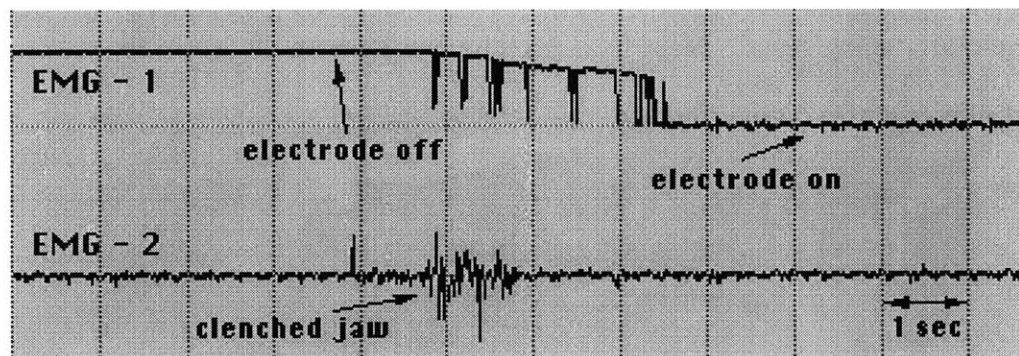


Figure 2.12: Example of a Flat EMG Signal

The artifacts described in the previous pages result in poor data for sleep analysis. These artifacts must be minimized in a sleep instrumentation session to ensure collection of good quality data. [PI] was provided to astronauts on STS-90 and STS-95 to help them troubleshoot these artifacts in the data they were recording before going to sleep. Subjects in our ground-based study were asked to look for and diagnose problematic signals in the sleep instrumentation, such as those seen in the previous pages.

Chapter 3

Methods

The ground-based experiment required trained astronaut surrogates to monitor a sleep instrumentation setup, detect faults, and suggest a method to return the signal to good quality. Execution of this experiment required the participation of several "sleep" subjects to wear the SleepNet and several test subjects to be trained as astronaut surrogates and monitor sleep signals.

3.1 Participants

MIT undergraduate and graduate students were recruited to act as sleep and test subjects and were paid for their participation. Sleep subjects, who donned the SleepNet for the duration of an experiment, were not allowed to later act as test subjects because they helped the research assistants create signal anomalies by tapping an electrode or holding an electrode away from the skin. This interaction might have given sleep subjects a biased ability to detect anomalies if they later acted as test subjects. Test subjects who had completed their testing were, however, allowed to act as sleep subjects. Of 32 test subjects (16 female, 16 male), 30 completed all of the testing sessions.

3.2 Experimental Design

The experiment used a balanced crossover design, where the 30 subjects were divided into two groups. Each subject was tested in two separate thirty minute sessions, on two days, with the instrumentation. Group 1 had [PI] assistance only on their day 1 and Group 2 only on their day 2 of testing. The groups were approximately gender balanced, because the January 1998 pilot study found a significant gender difference. (In the pilot study, the female subjects outperformed the male subjects by approximately 4.5 seconds). The (female, male) composition of Group 1 was (8,6); of Group 2 (6,10). Refer to Table 3.1

for the experiment design.

Table 3.1: Experimental Group Design

	Group1 N=14	Group 2 N=16
Day 1	[PI] on	[PI] off
Day 2	[PI] off	[PI] on

Each group received the same 3.5 hours of training on the sleep instrumentation and the [PI] interface. The schedule of the training session was based on information used for training sessions for the Pilot and Phase I ground-based studies. The training session schedule is shown in Table 3.2.

Table 3.2: Training Session Schedule

Time	Activities
9:00 - 10:00 AM	Administrative Activities: <ul style="list-style-type: none"> •Fill out consent forms •Read Training Manual Juice and Donuts available.
10:00 - 11:30 AM	Lecture: <ul style="list-style-type: none"> •Introduction to PI-in-a-Box •Show Instrumentation: SleepNet, DSR, Thinkpad, [PI] •Overview of Experiment
11:30 AM to 12:30 PM	Lunch on own
12:30 to 1:30 PM	Live Demonstration of [PI] Software: <ul style="list-style-type: none"> •How to use interface •How to input responses •How to interact with assistant and sleep subject

3.3 Procedures

The experiment involved the interaction of three people: an MIT research assistant trained by Brigham & Women's Hospital staff, a sleep subject who donned the instrumentation, and the test subject. The test subject sat opposite the sleep subject and the assistant, with a barrier in between. The test subject had an IBM ThinkPad (TP) laptop computer loaded with the [PI] software and a copy of Appendix A, the NASA Troubleshooting Guide. Figure 3.1 shows a diagram of the experimental setup.

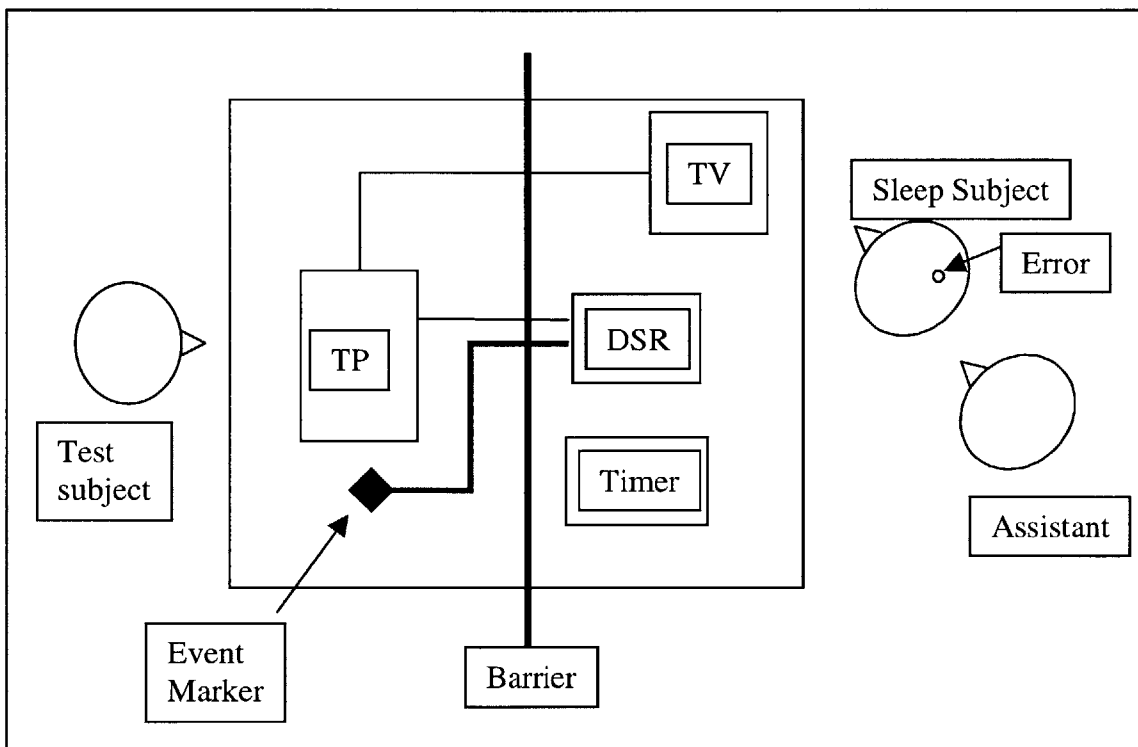


Figure 3.1: Experiment Setup

Each problem-solving cycle began with the assistant creating a problem. The test subject wore headphones playing music so that he/she could not hear any auditory cues, but was able to monitor the signals while the assistant created the problem. The assistant then tapped the test subject on the shoulder to indicate that he should begin the detection/troubleshooting process. The test subject removed his headphones and clicked on the event

marker, a device used for timekeeping purposes to mark the beginning of each event. Three minutes were allotted to detect and troubleshoot the problem, after which a timer sounded. At that time, the test subject stopped troubleshooting, turned the NASA Troubleshooting Guide back to Page 1 and replaced the headphones on his head. The assistant then fixed the previous fault and set the system up for the next problem.

Detecting the problem

The first step for the test subject was to detect the presence of a faulty signal, or a set of them. When diagnostics are enabled, [PI] assesses the quality of each channel, and displays a rating: a green light represents a good quality signal, a yellow light represents an unknown quality signal, and a red light represents a poor quality signal. The test subject turned diagnostics on for the problematic channel by clicking on the indicator light and reading the messages displayed in the diagnostics window. When [PI] diagnostics were disabled, the indicator lights remain yellow, and the diagnostics window remains blank, regardless of whether or not the test subject clicks on an indicator light. Once an anomaly was detected, the test subject clicked the gray rectangular checkbox beside the waveform, as seen in Figure 3.2.

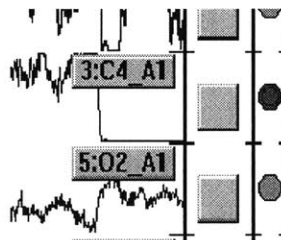


Figure 3.2: Close-up of Gray Checkbox and Red and Green LEDs

Clicking the checkbox brought up a dialog window, as seen in Figure 3.3. In this window, the test subject selected from a list of states the one that best describes the current status of the system.

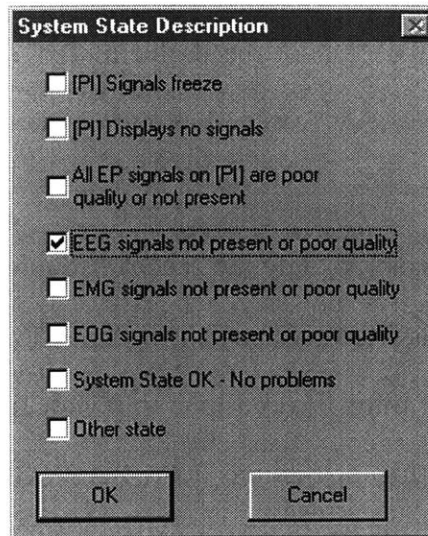


Figure 3.3: Interface Used to Input System State Description

If none of the descriptions are adequate, "Other state" was selected. If a mistake was made in selecting the system state from the dialog window, subjects could fix it by clicking on the checkbox with the "X" in it, (Figure 3.4), and selecting the "Go Back" button. **Note that two null faults were included in the experiment where NO problem was introduced into the system.** The subject indicated this by clicking "System State OK - No Problems."

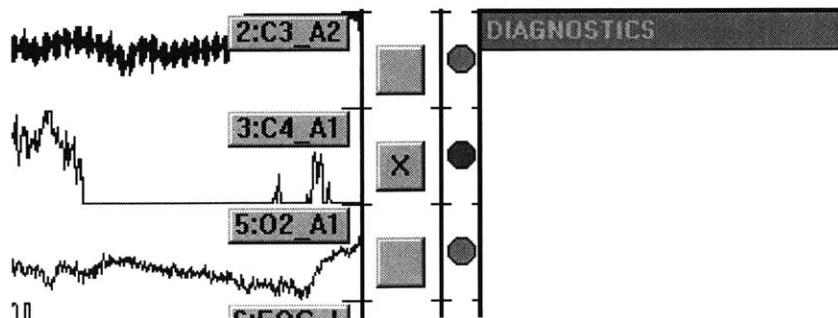


Figure 3.4: Gray Checkbox with 'X' After Subject Entry

Troubleshooting the Problem

The next step was to troubleshoot the problem. Test subjects followed the troubleshooting flowcharts in the NASA Troubleshooting Guide or the diagnostics suggestions listed by [PI].

With [PI] diagnostics disabled, only the NASA Troubleshooting Guide was available for troubleshooting. Subjects were given a copy of the NASA Troubleshooting Guide during the training session and were allowed time to review the procedures before their test sessions. Some troubleshooting instructions in the NASA Guide were not applicable to the ground-based study, such as “Replace DSR.” These instructions are shown in Table 3.3. Whenever these commands were encountered, the test subject was instructed to ignore them and move on to the next command, or step back in the procedures, as a mistake must have been made. The experiment was designed so that no catastrophic failure would be the cause of the problem - i.e., no DSR or IBM Thinkpad laptop failures. These problems were viewed as rare and beyond [PI]’s capability to repair. No serious hardware failures occurred during the experiment.

Table 3.3: Command or Questions to Ignore or Avoid

Commands and questions to Ignore or Avoid while using the NASA Guide
[PI] Icon Displayed?
Exit [PI]
ThinkPad Power Off
Unstow or Replace RS-232 Cable
Unstow or Replace DSR
Stop Recording
Turn DSR Off
Unstow or Disconnect or Replace Borg Harness

With [PI] diagnostics enabled, test subjects viewed troubleshooting steps in the diagnostics window by clicking on the appropriate indicator light. The subject entered the identification of the fault by clicking on the gray checkbox corresponding to the problematic channel. This action displayed the pop-up window seen in Figure 3.5.

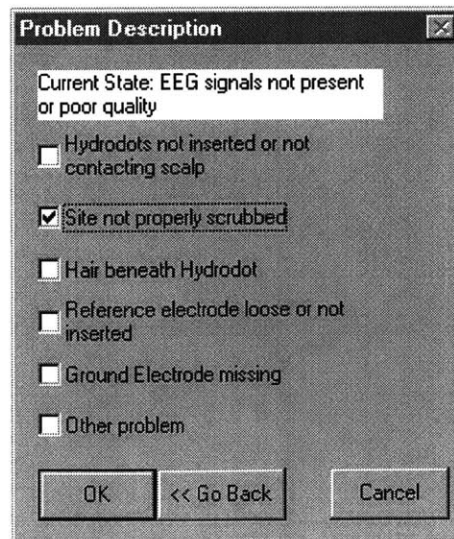


Figure 3.5: Interface Used to Input Problem Description

Subjects then selected the description of the fault that most closely matched the current problem and clicked OK. If a mistake was made in selecting the problem description, the "Go Back" button was used to change the input. The subject interacted with the assistant by asking questions about a specific electrode. Test subjects also gave commands to the sleep subject. If the left EOG channel (EOG_L) was showing anomalous behavior, then the subject could ask the assistant to check if the Left EOG Hydrodot was inserted into the socket and flush. They could also ask the sleep subject to move his or her eyes, to determine if the EOG waveforms track the eye movement. After successfully fixing the instrumentation, the assistant confirmed the fix command, and set up the next problem. Test subjects were restricted from physically interacting with the sleep subjects.

Fixing the Problem

After entering the problem description, the test subject instructed the research assistant to fix the problem, i.e. "Insert the O2 hydrodot." Once the problem was fixed, the test subject popped up the dialog window, seen in Figure 3.6, and entered the solution description. The "X" then disappeared from the checkbox. After successfully fixing the instrumentation, the assistant confirmed the fix command, and set up the next problem.

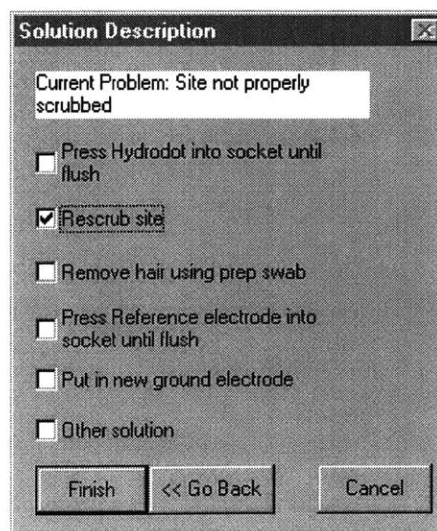


Figure 3.6: Interface Used to Input Solution Description

Chapter 4

Data Analysis and Results

Each test session produced a log file of times, events, channels and subject actions associated with [PI] firings, faulty channels, and signal quality. A subject action includes the event marker time (the beginning of each error cycle), time of detection, the name of the channel on which a fault was detected, the system's state, and the name of the problem identified. A Matlab script created an Excel file. A sample log file entry is shown below:

```
13:10:42 >> 0:00:00 01_A2 poor quality
13:10:43 >> 0:00:00 OP LED on: 2
13:10:45 >> 0:00:00 OP LED off: 2
13:11:05 >> 0:00:00 Menu Popup
13:11:09 >> 0:00:00 OP chose State 1
13:11:09 >> 0:00:00 Check X 2
```

The task was divided into detection, troubleshooting, and fixing the problem, as represented in Figure 4.1. Fault onset is the time at which the subject presses the event marker. Test subjects were allowed a total of 180 seconds to detect and troubleshoot the problem. If no fault was detected in the given time, he/she was said to have *timed-out* on the error. In some cases, subjects were able to detect that a fault exists, but were unable to diagnose it within the allotted time.

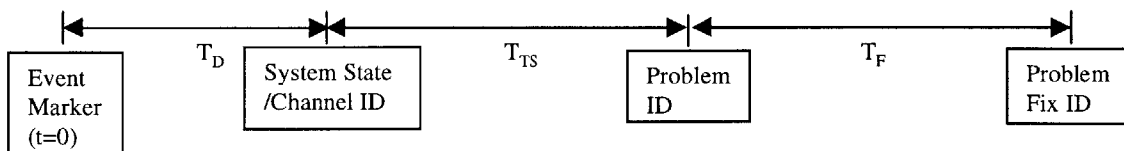


Figure 4.1: Timeline Showing Detection, Troubleshooting and “Fix” Times

4.1 Definitions

The time it takes to detect that the quality of the signal has fallen below a prescribed level is called the detection time, T_D . It is equal to the time at which the subject reports the sys-

tem state and the channel identification showing the fault minus the event marker time.

$$T_D = \text{Channel Time} - \text{Event Marker Time} \quad (4.1)$$

The troubleshooting time, or T_{TS} , is the time at which the subject reports the problem identification minus the system state/channel identification time.

$$T_{TS} = \text{Problem ID Time} - \text{System State/Channel ID Time} \quad (4.2)$$

The "fix" time is a measure of the time it takes the **research assistant** to fix the problem after the test subject has instructed the assistant to do so. Test subjects were asked to enter the fault they found into the pop-up window, and then to instruct the assistant to fix it, i.e. "Please insert the O2 hydrodot." After the research assistant had fixed the problem and the signal has returned to good quality, the test subject was asked to input the problem fix. T_F is the time the subject reports the problem fix minus the problem identification time.

$$T_F = \text{Problem Fix Time} - \text{Problem ID Time} \quad (4.3)$$

The fix time is a measure of the time it takes the research assistant to fix the fault. It is not subject dependent, and is, therefore, not analyzed, although it might be useful some day to know exactly how long it takes to fix a particular problem in an instrumentation setup. If a future version of [PI] includes a scheduler, for example, the time needed to fix each error would be coded into the software and that information could help determine if a particular compound problem could be fixed in a specified amount of time. If, for instance, the C3 electrode fails while the C4 electrode shows a good quality signal, it may not be necessary to fix the C3 channel since those electrodes are redundant. If an astronaut has only 10 minutes to apply electrodes and determine that the system is operating properly, it may not be efficient to spend 4 of those minutes fixing the C3 electrode.

4.2 User Action Analysis

Subjects were trained to click on the appropriate channel and input the state of the system when they detected an error. The subject had to choose from a list of seven states (shown previously in Figure 3.3) the one that best fit the current state of the EP signals displayed. These states included options such as "EP Signals Not Present or Poor Quality" and "EEG Signal Not Present, Poor Quality" and correspond to the troubleshooting procedures laid out in the NASA Troubleshooting Guide, flown on STS-95.

A subject must correctly diagnose both the channel and the state. The time it takes to present a correct diagnosis is the detection time, T_D . There is an ambiguity in situations in which the subject did not correctly diagnose either the channel or the state on the first try. Table 4.1 below gives the distribution of situations (User Actions) observed in the data.

Table 4.1: Breakdown of User Actions

User Action	Description	# of Errors	% of Errors
1	Correct Diagnosis on 1st try	618	78%
2	Correct State, "Go Back," Correct State	12	1.5%
3	Wrong State, "Go Back," Correct State	44	5.5%
4	Correct State, Wrong Channel	27	3.4%
5	Wrong State, Correct Channel	55	6.9%
6	Correct State, "Go Back," Time Out	20	2.5%
7	Other	17	2.2%

We struggled over the definitions of the detection time, T_D , and the troubleshooting time, T_{TS} . Sometimes it was unclear whether the subject was still detecting that there is a problem or deciding which problem was detected. The answers to two questions helped us to define T_D and T_{TS} . First, when does the subject detect the correct problem? Second, when does the subject begin troubleshooting the correct problem?

We chose to define T_D as the first time the subject clicks on a channel and assesses the state. In most situations, this seems appropriate. In User Action 4, however, the subject has chosen the wrong channel. This may indicate that he/she has not detected the correct problem, or that the subject has accidentally checked the wrong box.

In User Action 2, the subject seems to be second-guessing his own decision. In User Action 3, the subject has chosen the wrong state, then changes to the right one. The determination of the correct state is the first step in deciding how to troubleshoot the problem.

We have defined the troubleshooting time, T_{TS} , as the time from T_D to the time when the correct problem is entered. An alternate approach might be to define a third time, say a "gray region" time that cannot clearly be called T_D or T_{TS} .

4.3 Analysis

The experiment was set up according to a repeated-measures design which allows each subject to serve as his/her own control. The same 13 faults were presented on each test session, but they were presented in reverse order on day 2. Subjects did not recognize the reversed order on the second day, according to their later report, and were therefore not expecting the particular fault they would have to diagnose next on day 2 any more than on day 1.

A repeated-measures ANOVA was performed on the values of T_D and T_{TS} by SYSTAT9 (SPSS Corporation), a statistical software package. Note that if a subject is missing even one of the 26 times recorded (each error tested twice) for the two days, his remaining data is omitted from the analysis. (A repeated measures analysis requires **every** measure from each subject). When a fault went undetected or incorrectly created, the subject had no corresponding value of T_D ; those losses and consequent deletions reduced the number of subjects in the T_D analysis from 30 to 25.

The repeated-measures ANOVA showed a significant main effect of Day on T_D , ($p=0.006$). Subjects detected a problem faster on their second day of testing than on their first: they had learned how to read the EP signal screen better and how to determine more swiftly if a signal quality had degraded. [PI] assistance did not show a significant main effect ($p=0.749$) on T_D , which agrees with the finding of the January 1998 pilot study. (The January 1998 pilot study found only a significant main effect of [PI] x Day, $p = 0.002$). [2]

The agreement between studies is interesting because it occurs despite differences in the experimental set-up, (January 1998 pilot vs. January 2000 Phase II). First, the experiment setup of January 1998 used a script of 59 errors that had been recorded during an earlier spaceflight and was, therefore, easily and precisely reproducible for each of the 12 subjects. In particular, the experimenters knew the exact onset time for each of the 59 faults. In the January 2000 study, by contrast, only 13 errors were created live, anew, and by hand in real time. This made it impossible to present each fault at exactly the same time to each subject. Instead, on a signal from the assistant indicating that they could begin, subjects pressed the event marker to set an explicit initial reference time-mark. If a subject forgot to press the event marker, an estimated event marker time was inserted. This consideration added uncertainty to T_D . Further, it was difficult to introduce a **noise fault** into the instrumentation. At times, noise in the EP signal appeared only intermittently during the 180 seconds while the test subject was looking for an error. In other situations, especially in creating the **ground electrode** fault, the error could not be reproduced correctly.

The time to troubleshoot, T_{TS} , is at least as important to this study as T_D because its main purpose was to determine how well [PI] assistance helps astronaut surrogates not only to detect but also to rectify faults. Past ground studies, by contrast, concentrated on

detection alone. The present study examines the [PI] system as a whole. Moreover, because the detection time is shorter (5 to 10 seconds), it contributes less to the total task time than troubleshooting (20-30 seconds on average). Assistance at this step is, therefore, potentially important.

A repeated-measures ANOVA was performed on T_{TS} for all 13 errors, two of which were “null” faults (the absence of a fault in the system). These null faults were included in the experiment to ensure that a subject could not expect to find a fault in each setup, [i.e. they did not have preconceived knowledge that a fault **must** exist]. (This simulates the problem facing astronauts who do not have this advantage during spaceflight.) While the null faults have an associated detection time, they do not, of course, have a troubleshooting time and only the 11 non-null errors can be analyzed for T_{TS} .

One of the faults, removing the ground electrode, had an especially high fraction of time-outs among the subjects. On Day 1, 9 of the 30 subjects were unable to troubleshoot the problem. More, this fault was particularly hard to create and the EP signals could unpredictably continue to be of good quality even after the ground electrode was disconnected. Sometimes all of the signals would fail, as they should when the ground electrode is removed. Also, the NASA Troubleshooting Guide was difficult to use in troubleshooting this fault, because subjects found its flowchart hard to follow. Overall, the ground electrode fault is not typical of the population of faults analyzed and is therefore not included in the repeated-measures analysis for T_{TS} .

Any subject who is missing data (e.g., T_D or T_{TS}) for even one of the 10 analyzed errors is omitted from the repeated-measures analysis, and 15 of the 30 subjects were lost in that way, leaving 150 total data points. Seven subjects were from Group 1 and eight from Group 2; seven subjects were male, and eight were female. In this analysis, the gender of the subject gives a marginally significant, ($p=0.049$), main effect on T_{TS} . That is,

male subjects outperformed female subjects in troubleshooting faults by 5 seconds, on average. Note that this result is true only for this analysis. Other repeated measures analyses (see the following paragraphs), considering fewer errors and therefore more of the subjects, did **not** show gender as a significant main effect on T_{TS} . The identity of the research assistant who ran the experiment and group factors did not show a significant main effect. Refer to Table 4.2 for the significance of main effects on T_{TS} .

Table 4.2: Main Effects on T_{TS}

Factor	P Value
Gender	0.049
Assistant	0.816
Group	0.146

The within-subjects analysis shows a significant Day x Group effect with $p=0.001$. This cross factor represents the [PI] assistance factor, which can be seen in Table 4.3.

Table 4.3: Day x Group Table

	Group 1	Group 2
Day 1	[PI]	No [PI]
Day 2	No [PI]	[PI]

The significant effect of [PI] for reducing troubleshooting time can also be seen in Figure 4.2. The top graph shows Group 1 subjects who received [PI] help on their first day and no help on their second day. All but one of these subjects performed better on their first day with [PI]. This result is important because it means that subjects took **longer** to rectify faults in the instrumentation on their second day, despite the fact that they had one day of experience in performing the experiment. In other words, [PI] provided more assis-

tance for troubleshooting than did one day of experience with the system. The bottom graph shows Group 2 subjects who received [PI] help on their second day only. All of these subjects performed better on their second day with [PI] help. Perhaps more noteworthy is the fact that Group 2 on their day 2, with [PI] assistance and one day of experience, outperformed Group 1 subjects on both their day 1 and day 2. Considering the results of Group 1 and Group 2 together indicates that in future spaceflights, especially for the International Space Station, using a [PI] system to assist astronauts in troubleshooting a life sciences experiment would be more beneficial than merely having astronauts practice without an expert system available for assistance.

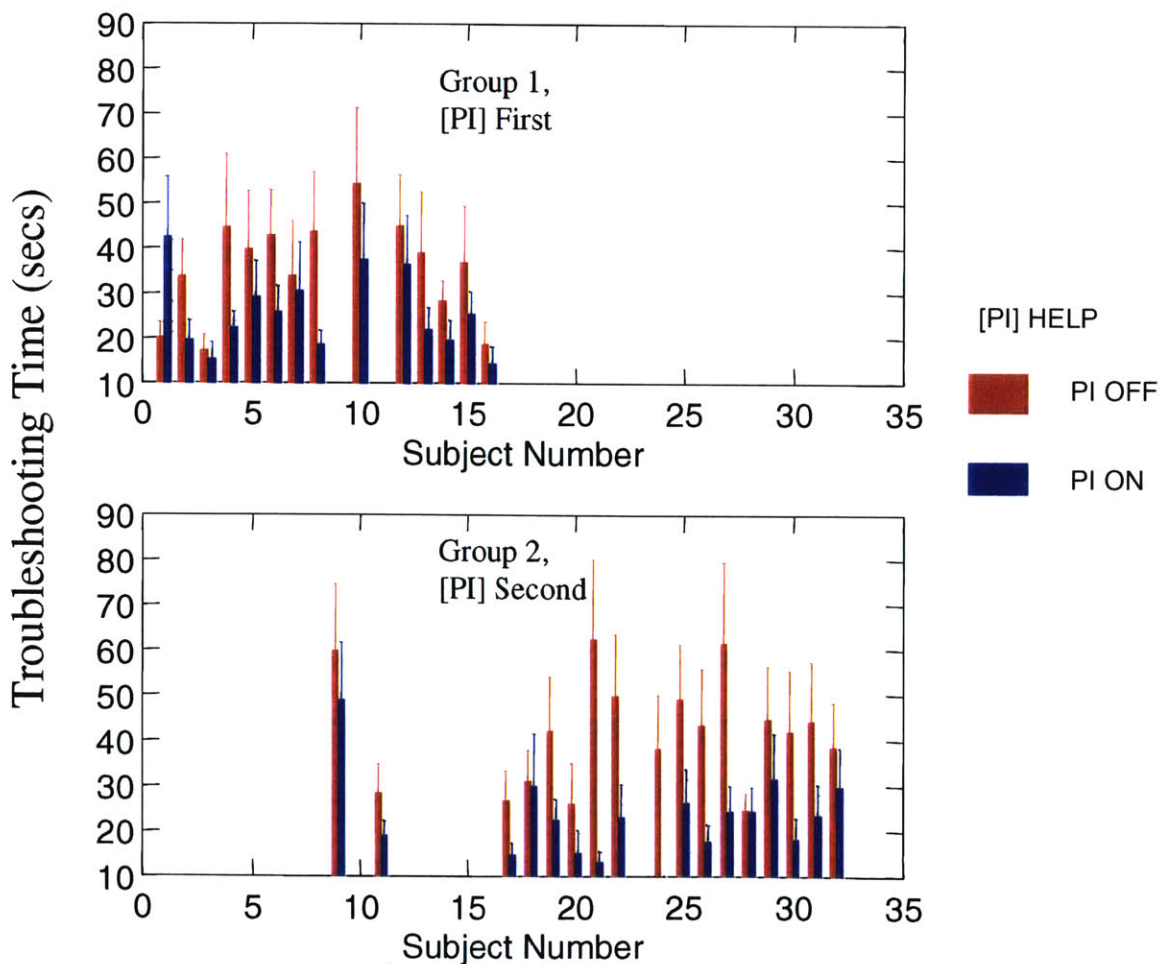


Figure 4.2: Troubleshooting Time is Significantly Improved with [PI] Help

There was no significant main effect of day on T_{TS} , $p=0.078$. The results taken together suggest that [PI] improves subjects' troubleshooting time, regardless of the day on which they received [PI] assistance. Table 4.4, below, shows the Main Effects and their significance for T_{TS} .

Table 4.4: Main Effects on T_{TS}

Factor	P Value
Day	0.078
Day x Sex	0.265
Day x Group	0.001
Day x Asst	0.627

Subsets of the 13 faults are considered for repeated-measures analysis in order to recover those subjects who were lost because they missed some of the original T_{TS} measurements. [If, however, that fault with missing data is deleted, the analysis of the remaining faults can be strengthened by adding that subject's data]. We are forced to balance the inclusion of as many measured variables as possible against the inclusion of as many of the subjects as possible.

For the above reasons, considering only 7 of the 11 errors for a repeated-measures ANOVA on T_{TS} allows analysis of 21 of the 30 subjects. Ten subjects were from Group 1 and eleven from Group 2; nine subjects were male and twelve subjects were female. This analysis reveals that no between-group factors, (sex, assistant, or group), are significant (Sex $p=0.160$, Assistant $p=0.882$, Group $p=0.275$). **For within-subject results, both Day and Day x Group are significant with $p=0.023$ and $p=0.003$. These subjects perform the troubleshooting task better on Day 2, than Day 1, and also better with [PI] assis-**

tance than without it. It follows from this that Group 2 on Day 2 with PI assistance had the best overall troubleshooting times.

A sensitivity analysis was performed to determine the robustness of the results against different choices of faults retained in the analysis. When 8 and 9 errors are analyzed, data for 20 and 16 subjects can be analyzed, and the results from these analyses are similar to those discussed previously.

4.3.1 Detection Time by Fault Type

A repeated measures analysis showed no significant main effect of [PI] on T_D . Faults were categorized into null-faults (2 per day), single-channel faults (6 per day) and multiple-channel faults (5 per day). A single-channel fault is one in which only one of the 8 EP signals is of poor quality; a multiple-channel fault has two or more signals of poor quality. (One type of multiple-channel fault produces three poor quality EP signals while the other four types affect the quality of all eight EP signals). In the following tables and results, null-, single-, and multiple-channel faults are noted as 0, 1, and 2, respectively.

Detection time was highest for the null-faults, as indicated by the bar chart in Figure 4.3 on the following page. Subjects took longer to evaluate each signal before deciding that no error existed in the system than they did when a fault was present. It is possible that including the null error T_D with the T_D for the two other fault types may decrease the significance of [PI] as a main effect. For this reason, a repeated measures ANOVA was done on T_D for single- and multiple-channel faults only (excluding the null-faults). This analysis again indicated that [PI] was **not** a significant main effect for detection of any fault type.

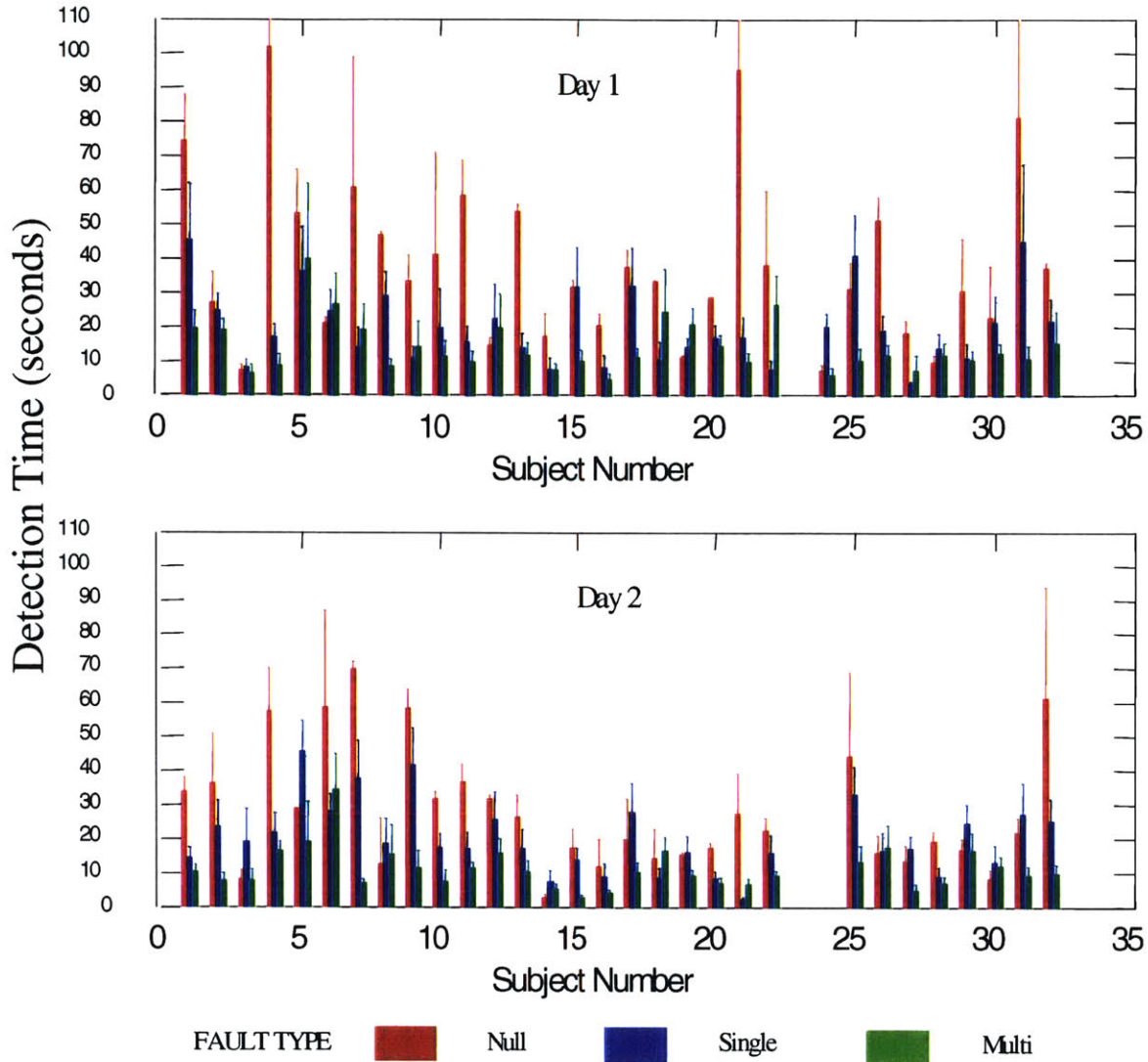


Figure 4.3: Detection Time by Fault Type

A significant median test (StatExact, Cytel, Inc.) would indicate that at least two of the three fault types have significantly different values of the median T_D . Such a result might suggest that the fault types should be analyzed separately. Separate repeated-measures analyses (which, unlike the median test, presumes a normal distribution) for each of the three fault types does not, however, show a significant main effect of [PI] on T_D .

A preliminary median test on T_D and T_{TS} on all subjects (ungrouped by Group or Day) showed a significant difference among the three fault types for T_D ($p=0.0000$) but not for

T_{TS} ($p=0.4689$). This suggests that subjects were able to detect certain fault types faster than others, but troubleshooting time was not affected by the type of fault. [This intuitively makes sense because the methods used to troubleshoot a single- or multiple-channel fault (reading questions from the [PI] software or from the NASA Troubleshooting Guide) are similar.]

A second median test on T_D and T_{TS} was performed for each group separately. Results are given in Table 4.5 below: the fault type is a significant factor for all four scenarios tested (T_D for Group 1 and 2 and T_{TS} for Group 1 and 2).

Table 4.5: Median Test for T_D , Fault Types 0, 1, and 2 by Group

	Group 1	Group 2
Detection Time (T_D)	0.0000	0.0000
Troubleshooting Time (T_{TS})	0.0099	0.0202

A median test on T_D across fault types by Group and Day showed a significant difference among them for all four Group-Day combinations. That is, the detection time is significantly different for at least one of the three fault types compared to the other two. There were, however, no significant differences between the fault types (single- and multiple-channel) for troubleshooting time (see Table 4.6 and Table 4.7).

Table 4.6: Median Test for T_D , Fault Types 0, 1, and 2 by Group and Day

	Group 1	Group 2
Day 1	0.0004	0.0001
Day 2	0.0000	0.0000

Table 4.7: Median Test for T_{TS} , Fault Types 1 and 2 by Group and Day

	Group 1	Group 2
Day 1	0.1007	0.4693
Day 2	0.8671	0.5392

The lowest null-fault detection time was higher than the highest detection time, T_D , for the single- and multiple-channel faults. Another median test was, therefore, performed on T_D for only the two non-null fault types: single- and multiple-channel faults, leaving out the null-faults. This analysis shows that T_D for single- and multiple-channel faults are significantly different for both groups on Day 2 of testing (refer to Table 4.8, below), but not on Day 1. This indicates a learning effect for detection of multiple-channel faults, not an effect of [PI], and supports the repeated measures findings reported earlier in this chapter.

Table 4.8: Significance of Fault Type 1 and 2 for Detection Time

	Group 1	Group 2
Day 1	0.1049 (with [PI])	0.1686 (no [PI])
Day 2	0.0001 (no [PI])	0.0040 (with [PI])

4.3.2 Detection Time by Fault Type - Histograms

Reviewing subjects' detection time by histogram, for each group and day, yields several interesting results. On average, Group 1 subjects detected null faults 10 seconds faster on their first day, with [PI] assistance, than on their second day without assistance. On their first day, subjects with [PI] assistance and no experience took between 10 and 30 seconds to determine that there was no fault in the system. On their second day of troubleshooting, when they no longer had the assistance of [PI], subjects took between 20 and 40 seconds. Refer to Figure 4.4 and Figure 4.5.

For the single-channel faults, by contrast, there is relatively little difference between Group 1's day 1 and day 2 mean. This implies that the value of [PI] assistance in detecting a single-channel fault is minimal. This is expected because [PI] false alarms on day 1 distract subjects from detecting which single-channel has a fault. Refer to Atamer's thesis for [PI]'s false alarm rate [1].

For multiple-channel faults, there is also relatively little difference between Group 1's day 1 and day 2. This is because the multiple-channel faults are easy to detect; it is a simple task to detect that the majority of the EP signal waveforms are of poor quality. Because the task is simple, [PI] would not be expected to give much help.

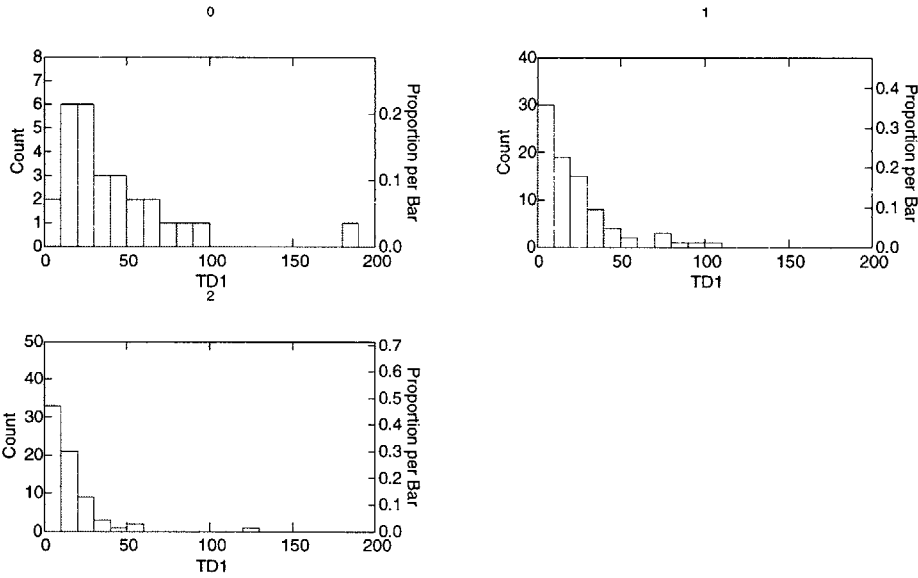


Figure 4.4: Detection Time Histogram for Group 1 Day 1 (With [PI])

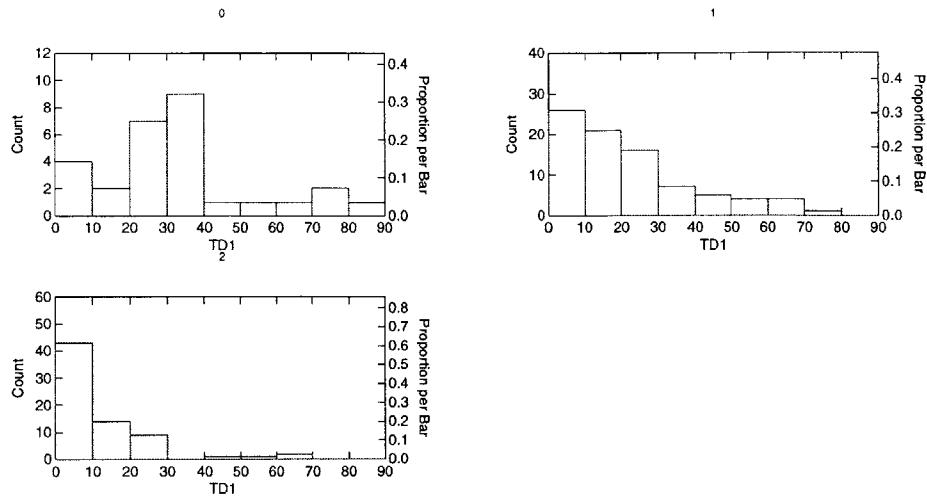


Figure 4.5: Detection Time Histogram for Group 1 Day 2 (Without [PI])

A comparison of Group 2 on day 1 and 2 mirrors the results for Group 1. The majority of Group 2 subjects detected null faults within 30 seconds on their day 2 with [PI] assistance. On their first day, without [PI], subjects detected null-faults within 50 seconds. The Day 2 result, 20 seconds faster than the Day 1 result, indicates a beneficial effect of both [PI] assistance and experience. Refer to Figure 4.6 and Figure 4.7.

For the single-channel fault, however, there is relatively little difference between Group 2's day 1 and day 2. This suggests that [PI]'s false alarm rate distracts subjects from detecting single-channel faults, despite their day of experience in reading EP signal waveforms, in agreement with the results found for Group 1. For the multiple-channel faults, there is a steeper fall-off in Group 2's day 2 histogram that is not seen on day 1. This implies that both training and [PI] assistance help in faster detection of multiple-channel faults.

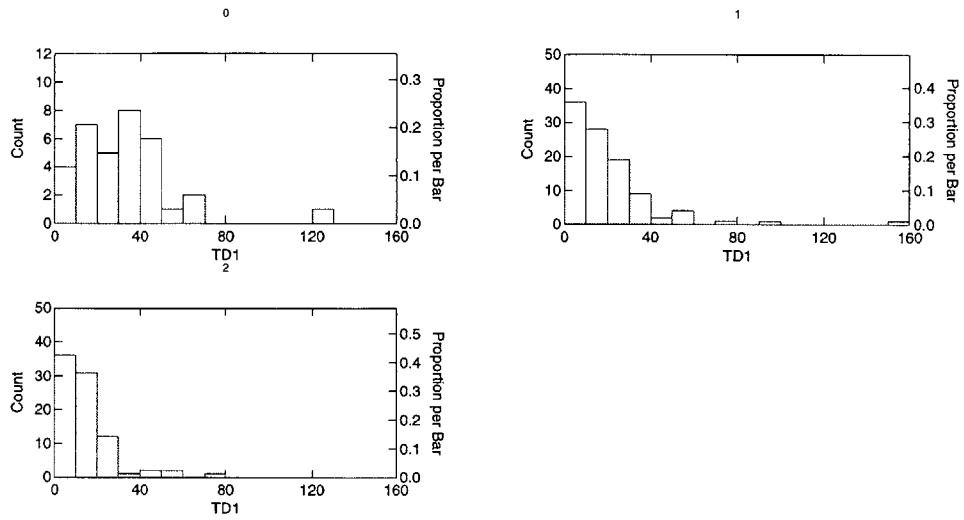


Figure 4.6: Histogram of Detection Time for Group 2 Day 1 (Without [PI])

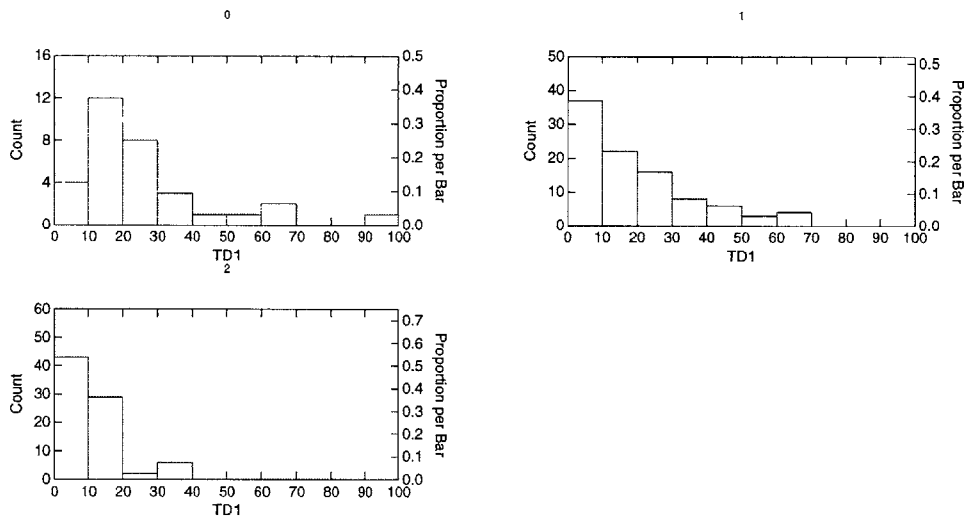


Figure 4.7: Histogram of Detection Time for Group 2 Day 2 (With [PI])

4.3.3 Subjective Responses

On completion of the experiment, subjects filled out questionnaires regarding the software and experimental setup. Refer to Appendix D for a listing of all questions. Questions were scored on a 1 (low) to 7 (high) scale. Figure 4.8 and Figure 4.9 show a boxplot of subject responses broken down by males and females. On average, subjects gave [PI] high scores. Male and female subjects showed few differences on subjective questionnaires. In addition to number scoring, subjects included constructive comments on the use of [PI] and the experiment setup as a whole. These comments can be found in Appendix E.

Differences were noted for Question 7, 8 and 11. Question 7 asked how well subjects understood the troubleshooting directions. Both males and females indicated that they had a fairly good understanding of the troubleshooting directions (mean=6). Males, however, show a larger variance about the mean. Question 8 asked how closely subjects followed the NASA Troubleshooting guide. Females mean response was a 6 while the males were a 6.5 with two outlying answers. Question 11 asked subjects to rate the usefulness of PI as a completely autonomous decision-making tool. Females mean response was a 5 while the males mean response was 5.5. Females show a larger variance with a low score of 1.

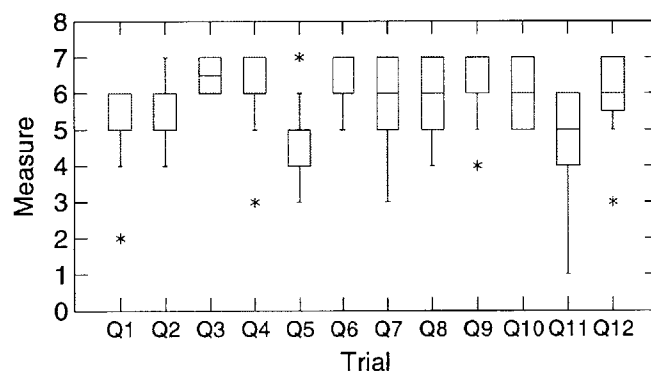


Figure 4.8: Subjective Questionnaire Results, Females

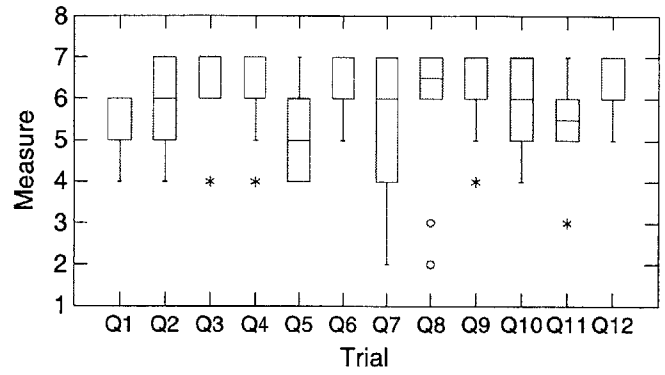


Figure 4.9: Subjective Questionnaire Results, Males

Chapter 5

Discussion and Conclusions

5.1 Discussion

A repeated-measures analysis was performed on the detection time, T_D . These results showed that [PI] assistance was not significantly helpful for detecting faults in a sleep instrumentation setup. Day, or training, however, was found to have a significant main effect on T_D . The implication of this result is that training is more important for astronaut surrogates in order to detect problems in instrumentation. This is true because astronauts are presumably intelligent and can use their natural abilities, with adequate training, to determine on their own when a fault appears. Further, detection of most of the faults in the sleep instrumentation setup were trivial tasks. That is, determining whether a poor quality signal is present is relatively simple. The significance of this result for spaceflight and experimentation on the International Space Station is that astronauts will benefit more from training than they will from use of expert systems to help in detecting faults. Principal Investigators must learn to adequately train astronauts and then rely on their intelligence to complete the task at hand.

A second repeated-measures analysis was performed on the troubleshooting time, T_{TS} . These results indicated that Day x Group, or [PI], did show a main effect for improving **troubleshooting**. In contrast to the detection task, day (or training) did **not** improve subjects' ability to rectify problems. (Recall that Group 1 performed the troubleshooting task on their Day 2 of testing, without [PI], slower than on their Day 1 with [PI]). One of the driving factors behind this result was that astronaut surrogates, in general, found the presentation of troubleshooting help easier to follow on the [PI] software than in the NASA Troubleshooting Guide. One subject commented, "The [PI] diagnostics were much easier

to use than the NASA Troubleshooting Guide, probably because the computer was already in front of me.” This result is most likely exasperated in the microgravity of space, when keeping a paper document in place is a difficult task in itself (items tend to float in microgravity). The significance of this result for spaceflight and experimentation on the International Space Station is that astronauts will benefit more from online computer assistance than they will from use of traditional NASA troubleshooting procedures.

As explained in Chapter 4, the faults introduced into the sleep electrophysiological setup were divided into three fault types: null-, single-, and multiple-channel faults. A median test on these fault types indicated that each of the faults types are significantly different from each other for the **detection** task, but not for the troubleshooting task. This is likely true because the process a subject uses to detect a null fault (i.e., that there is no fault) is different from the process a subject might use to detect a single- or multiple-channel fault. That is, it takes longer to closely inspect each EP channel and determine that all signals are of good quality than it does to determine that all are reading of poor quality (i.e., if all signals appear flat or noisy). This result is important for training purposes because it implies that different procedures are necessary to detect different kinds of faults in a system. It is also important because it implies that no special differentiation is necessary for **troubleshooting** different types of faults. Especially for ISS spaceflight experiments, it is important for astronauts to understand the different categories of faults they are looking for, and how best to do so.

5.2 Conclusions

The Phase 2 ground-based study run in January 2000 evaluated the troubleshooting assistance afforded by [PI] and sought to assess the utility of on-board expert systems, in general, for performing experiments and troubleshooting complex instrumentation systems. Thirty subjects, divided into two groups, received training on the sleep instrumentation

and the [PI] interface. Each subject was then tested on two separate days with a sleep subject instrumented in real time. Group 1 received [PI] help only on their Day 1 of testing; Group 2 received [PI] assistance only on their Day 2 of testing. During this phase of the ground-based study, the test subjects not only detected signal problems, but also interacted with the sleep subject and the trained operator to complete the diagnosis and repair. A balanced crossover design was used, and time and accuracy of the troubleshooting were measured.

Results indicate a beneficial effect of [PI] in reducing anomaly troubleshooting time. Only training, however, was found to significantly reduce detection time. In addition, no significant gender effects were found in this study. Post-experiment questionnaires showed that most subjects preferred monitoring the [PI] indicator lights while monitoring waveforms, rather than monitoring the waveforms alone. Results on the reliability of [PI]'s indicator lights can be found in Allen Atamer's thesis [1].

5.3 Recommendations

To best determine the effects of confounding Day (learning) and [PI] assistance, a 4 group design is recommended for further analysis in human-computer interaction studies. This should include two groups that are similar to the ones used in this ground-based study, as well as two additional control groups - Group 3 would receive no [PI] assistance on day 1 or day 2 and Group 4 would receive [PI] assistance on both days.

References

- [1] Atamer, A. "Reliability Evaluation of an Expert System Diagnostic Aid for a Sleep and Respiration Experiment." Unpublished Master's Thesis, Massachusetts Institute of Technology, 2001.
- [2] Callini, G. "Assessment of an Expert System for Space Life Sciences: a Preliminary Ground-Based Evaluation of PI-in-a-Box for the Neurolab Sleep and Respiration Experiment." Unpublished Master's Thesis, Massachusetts Institute of Technology, September 1998.
- [3] Callini, G., Essig, S.M., Heher, D.M., & Young, L.R. "Effectiveness of an expert system for astronaut assistance on a sleep experiment." Aviation Space and Environmental Medicine, 71 (9): 1-10.
- [4] Franier, R.; Groleau, N.; Hazelton, L. et al. (1994) PI-in-a-box - A knowledge-based system for space science experimentation. AI Magazine 15 (1), pp. 39-51.
- [5] Kryger, M.H., T. Roth, and W.C. Dement. Principles and Practice of Sleep Medicine. Second Edition, Chapter 89. W.B. Saunders, Philadelphia, PA, 1994.
- [6] Smith, Robin. "Fault Tree Analysis and Diagnostics Development for PI-in-a-Box with the Neurolab Sleep and Respiration Experiment." Unpublished Master's Thesis, Massachusetts Institute of Technology, 1997.
- 7] Young L.R. PI-in-a-Box. Journal of the Society for Instrument and Control Engineers. (1994); 33(2):119-22

Appendix A

NASA Troubleshooting Guide

E104

2.2a [PI] DISPLAYS NO SIGNALS

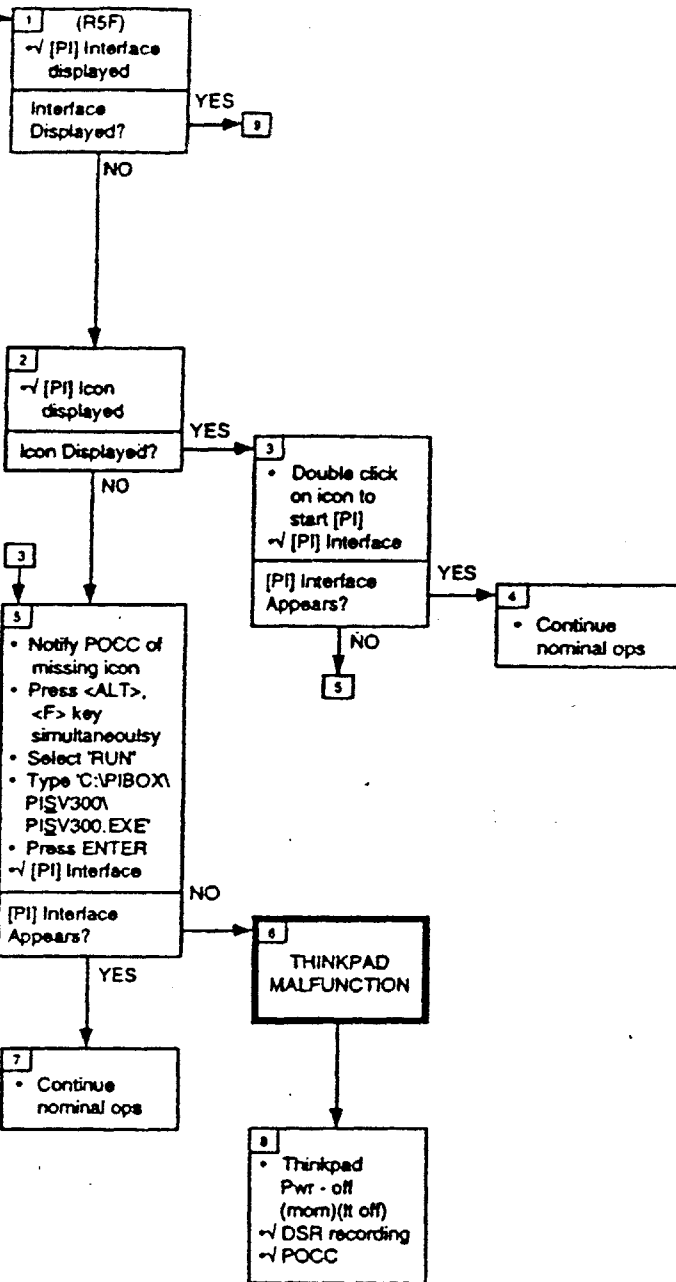
[PI] Displays No Signals

Nominal Config:

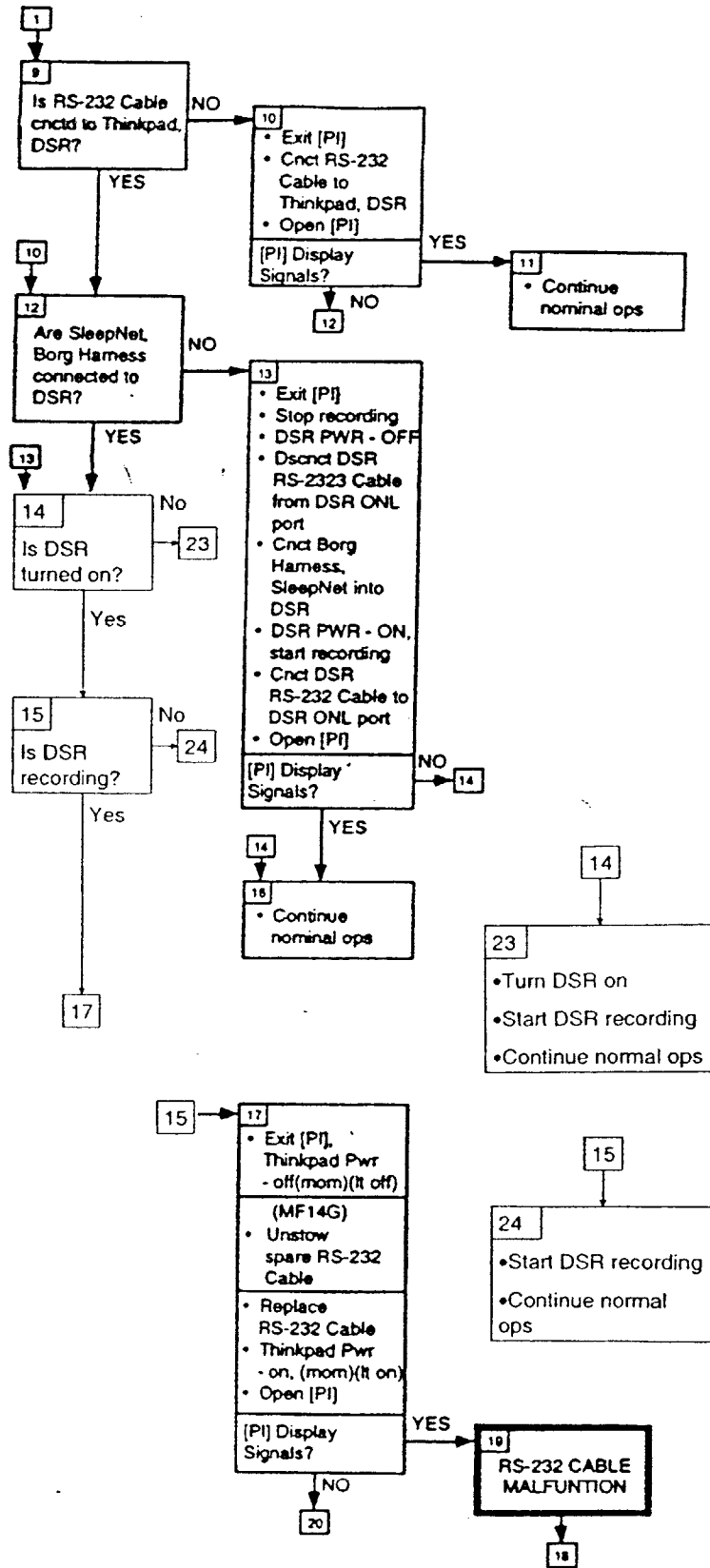
(R4W)
EPSP:
cb DC UTIL PWR - ON

(DSR)
PWR - ON
STAT It - ON
SLEEP Busy It - on
BORG HARNESS
Busy It - on
Definition File -
sleep_090597.def

PI-IN-THE-BOX
Thinkpad Pwr - on
(mom)/(t on)

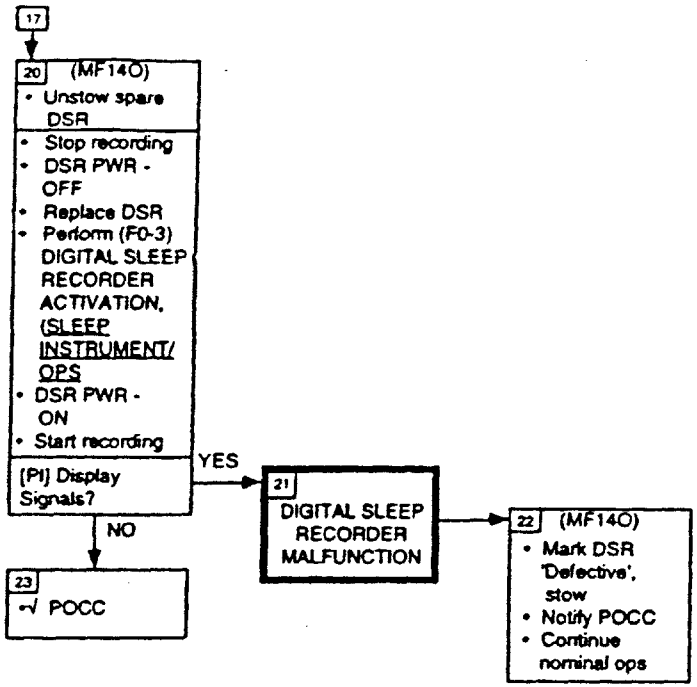


2.2a [PI] DISPLAYS NO SIGNALS (CONT'D)



E104

2.2a [PI] DISPLAYS NO SIGNALS (CONT'D)



E104

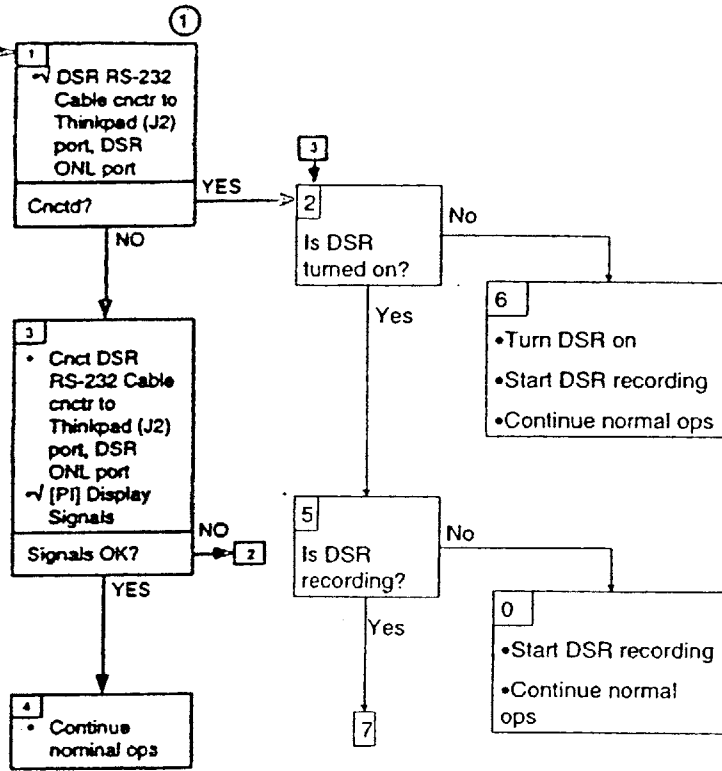
2.2b SIGNALS ON [PI] DISPLAY FREEZE

Signals on [PI] Display Freeze'

Nominal Config:
 (R4W)
 EPSP:
 cb DC UTIL PWR - ON

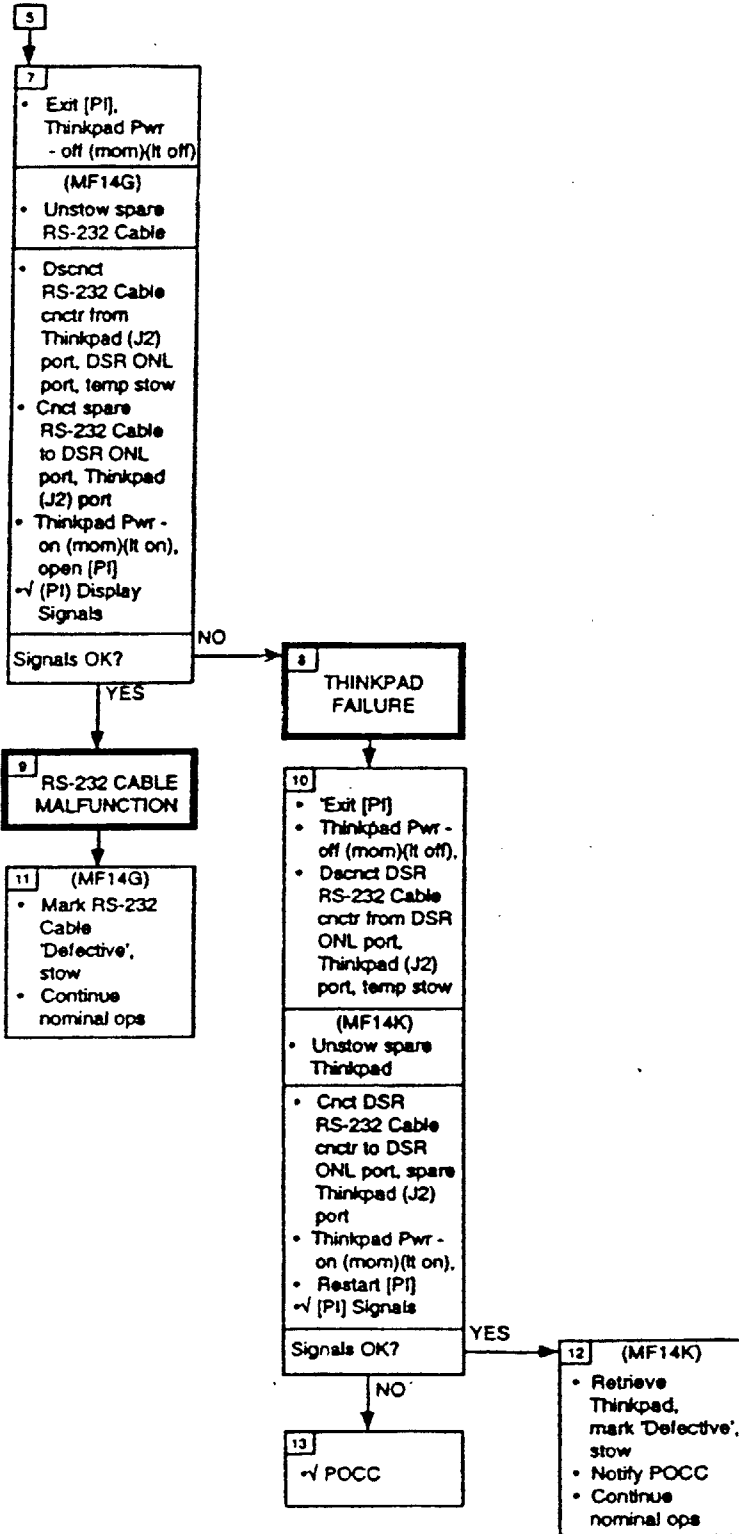
(DSR)
 PWR - ON
 STAT R - ON
 SLEEP Busy R - on
 BORG HARNESS
 Busy R - on
 Definition File -
 sleep_090597.def

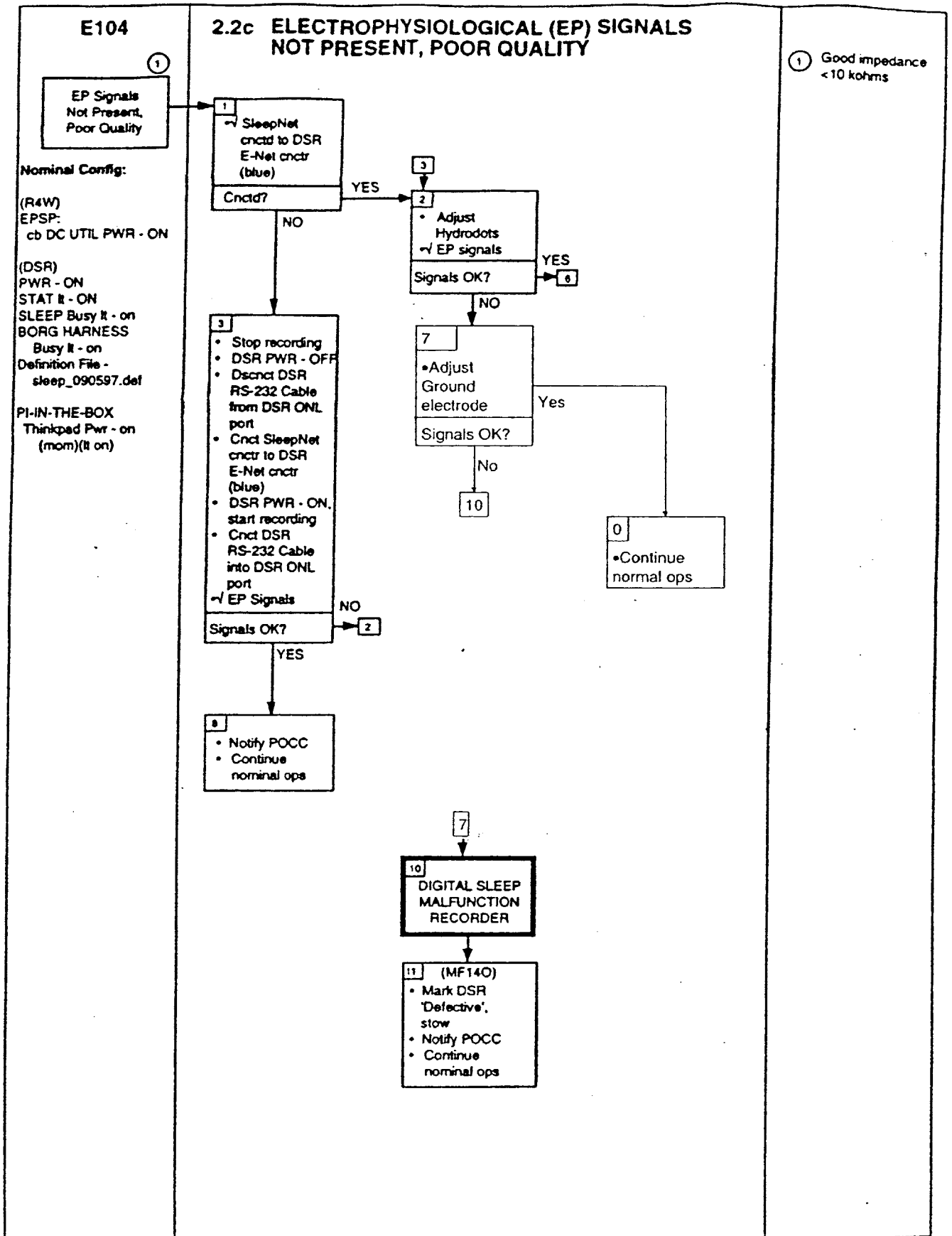
PI-IN-THE-BOX
 Thinkpad Pwr - on
 (mom)(It on)



① DSR will read 'REMOTE VIEW REC' if good communication between DSR, Thinkpad

2.2b SIGNALS ON [PI] DISPLAY FREEZE (CONT'D)





E104

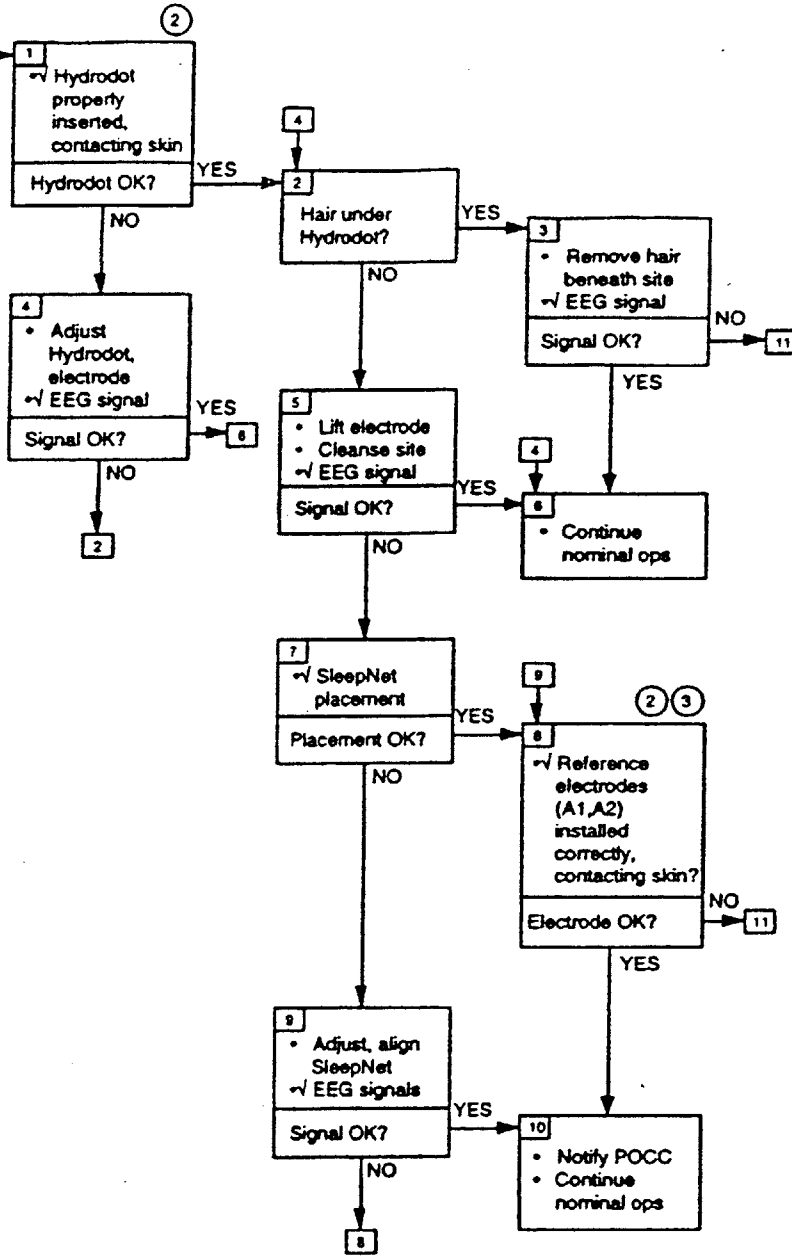
2.2e EEG SIGNAL NOT PRESENT, POOR QUALITY

EEG Signal Not Present, Poor Quality

Nominal Config:
 (R4W)
 EPSP:
 cb DC UTIL PWR - ON

(DSR)
 PWR - ON
 STAT II - ON
 SLEEP Busy II - on
 BORG HARNESS
 Busy II - on
 Definition File -
 sleep_090597.def

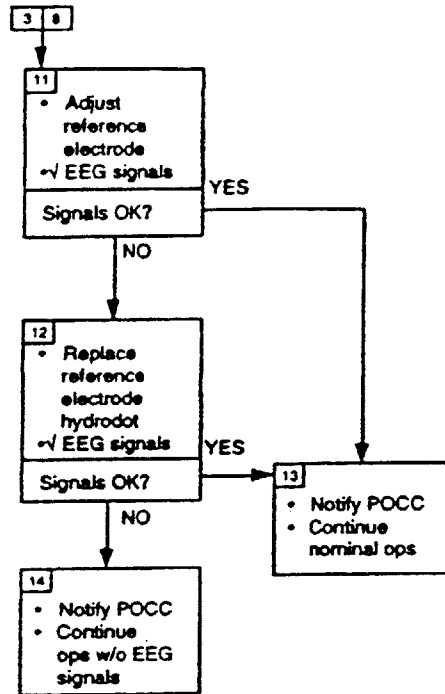
PI-IN-THE-BOX
 Thinkpad Pwr - on
 (mom)(It on)



- ① Good impedance < 10 ohms
- ② Hydrodot must be fully in socket, with gel contacting skin
- ③ O2, C3 reference is A2
O2, C4 reference is A1

E104

2.2e EEG SIGNAL NOT PRESENT, POOR QUALITY (CONT'D)



E104

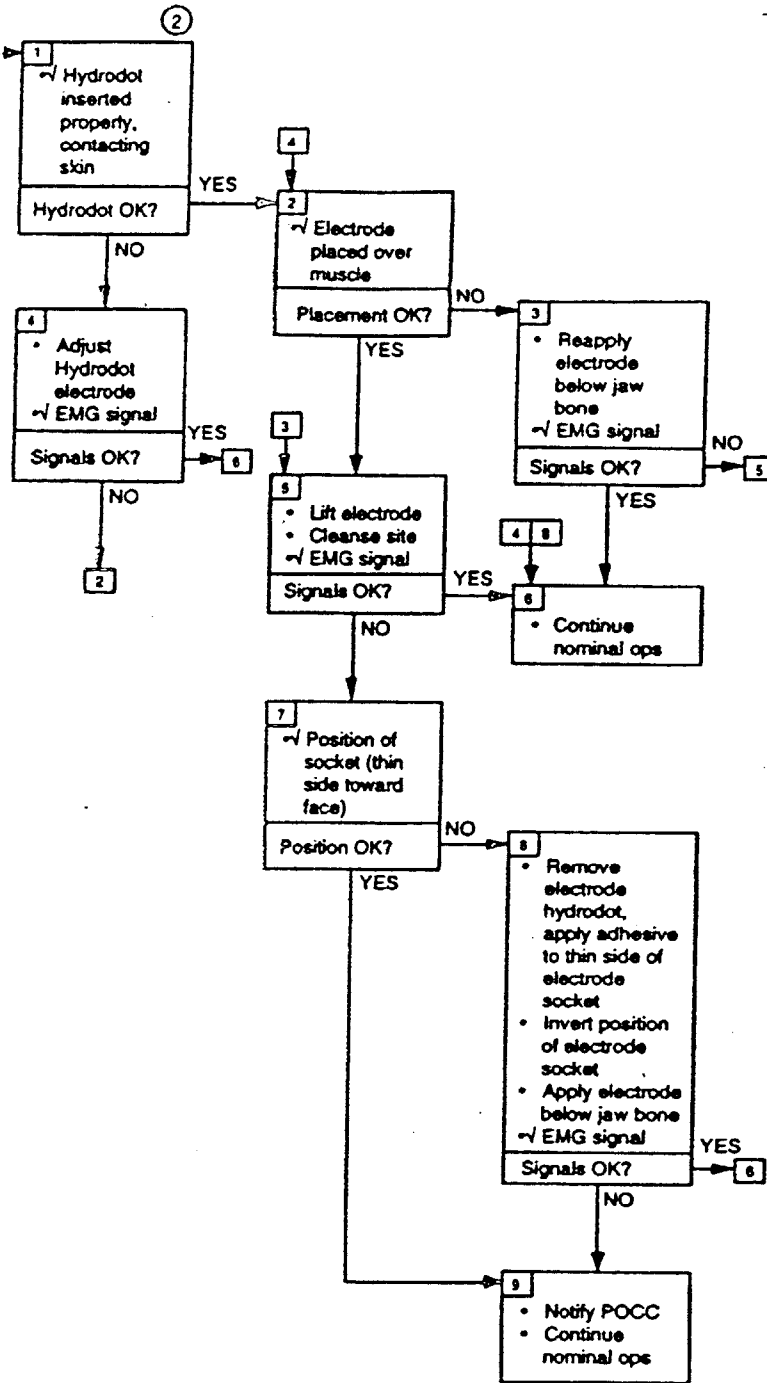
2.2f EMG SIGNAL NOT PRESENT, POOR QUALITY

①
EMG Signal
Not Present,
Poor Quality

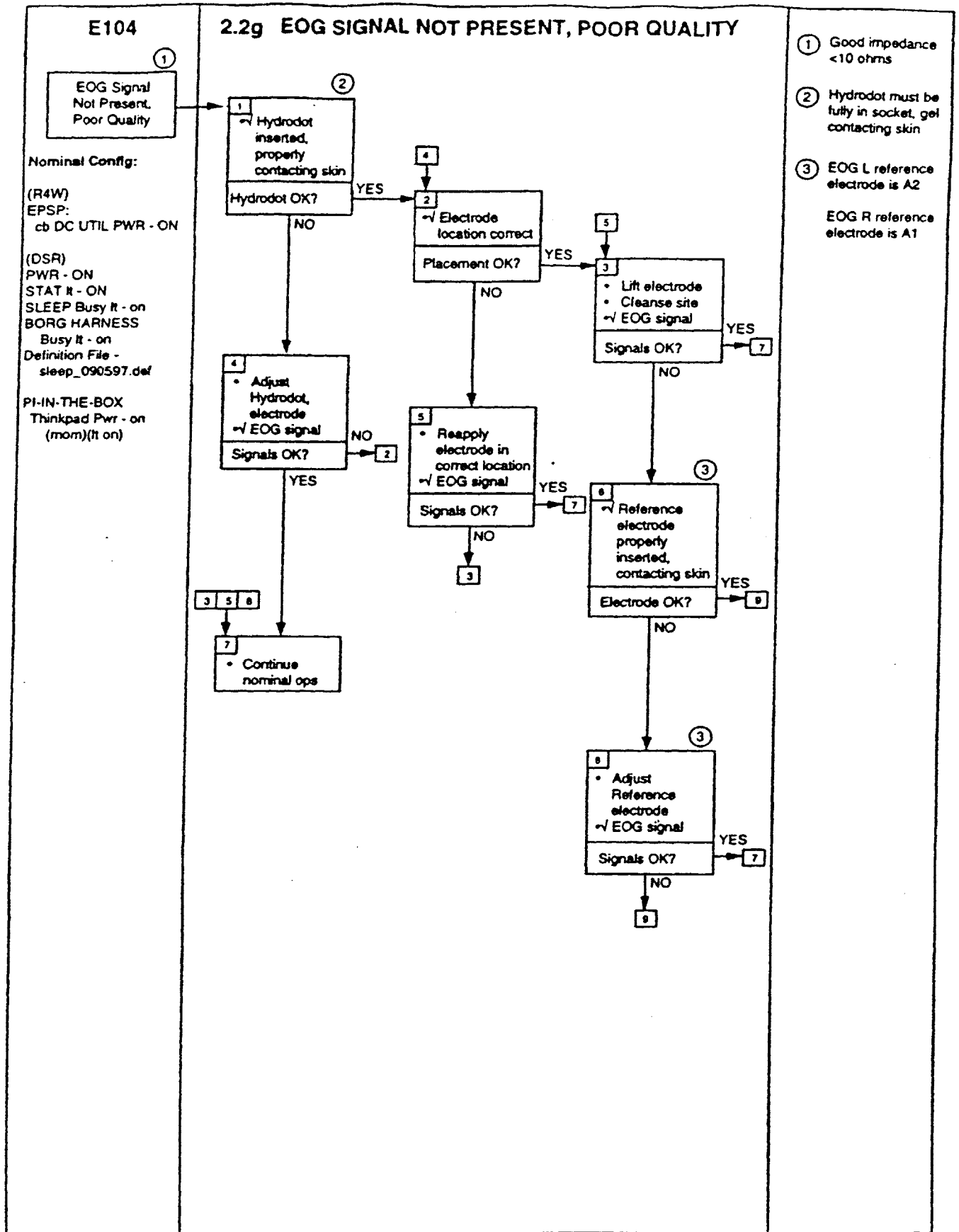
Nominal Config:
(RAW)
EPSP:
cb DC UTIL PWR - ON

(DSR)
PWR - ON
STAT R - ON
SLEEP Busy It - on
BORG HARNESS
Busy It - on
Definition File -
sleep_090597.def

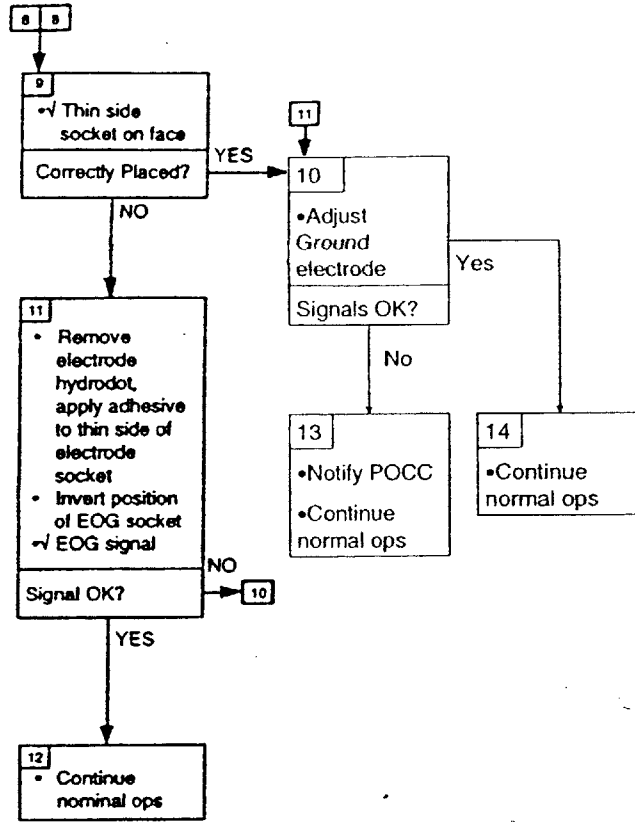
PI-IN-THE-BOX
Thinkpad Pwr - on
(mom)(ll on)



- ① Good impedance <10 ohms
- ② Hydrodot must be fully in socket with gel contacting skin



2.2g EOG SIGNAL NOT PRESENT, POOR QUALITY (CONT'D)



Appendix B

Pre-experiment questionnaire

Subject Number: _____

Date: _____

Principal-Investigator-in-a-Box Ground Based Evaluation Study Year 2 January 2000

Subject Name: _____

Home Phone Number: _____ MIT ID#: _____

E-Mail Address: _____

Please fill out the following for payment purposes:

Social Security Number (for payment): _____

Address: _____

Country of Citizenship: _____

Age (as of today): _____ Gender: Male Female

Year in School (freshman= 1): 1 2 3 4 Graduate

Field of Study: _____

How many hours per week do you use a personal computer or workstation? _____

Have you ever seen or worked on experiments involving electrophysiological signals such as EEGs, EKGs, etc ? YES NO

(If YES, please elaborate.) _____

Do you have any experience as a repair or support technician? YES NO

(If YES, please elaborate.) _____

Are you color blind? YES NO

Do you wear corrective lenses? YES NO

If YES, are you currently wearing GLASSES or CONTACTS (circle one)

Are you right or left-handed? LEFT RIGHT

Appendix C

Informed Consent Forms

C.1 INFORMED CONSENT FORM FOR TEST SUBJECTS

NSBRI PI-IN-A-BOX GROUND STUDY

Purpose

We would like permission to enroll you in a research study. The purpose of this study is to evaluate the efficacy of an expert system called PI-in-a-box in identifying the presence of artifacts in sleep data and suggesting corrective procedures to eliminate these artifacts. A version of PI-in-a-box has already been developed to assist astronauts in performing a sleep experiment in space. This experiment is designed to quantify the effectiveness of an expert system in a laboratory environment in terms of both time and accuracy.

Participation in this study is voluntary and you are free to withdraw your consent and discontinue participation in the experiment at any time without prejudice.

Procedures

You will be given approximately five to seven hours of training on a “training day.” It is intended to provide an overview of the equipment used in sleep recordings, and the characteristics of each signal recorded. You will be trained in the possible problems that can arise with the instrumentation of a Sleep*Net, a web-like cap used to record electrophysiological signals. Another volunteer will be wearing the Sleep*Net on their head. You will learn how to detect, troubleshoot and, correct problems which will occur in the instrumentation session. You will also be trained on the use of a computer decision aid called PI-in-a-box, which runs on a laptop computer. It will display the signals and use color-coded lights to indicate the quality of each signal. In addition, PI-in-a-box displays a “diagnostics” window which contains procedures for correcting poor quality signals.

Testing will take place over the course of three to four days, with one “training day” of five to seven hours, and two test days which will involve no more than one and a half hours on each day. Total testing time will be between 8-10 hours. During the test sessions, your task will be to detect problems in the sleep signal system, troubleshoot the problem to find the cause of these problems, and instruct a sleep technician to fix the problem and restore the quality of the sleep signals. One test day will be performed with the decision aid, and one without the aid.

Risks and Discomforts

No known risks associated with this component of the experiment.

Benefits

A prorated payment of \$7.00 per hour will be provided to participants.

In the unlikely event of a physical injury resulting from participation in this research, I understand that medical treatment will be available from the MIT Medical Department, including first aid emergency treatment and follow-up care as needed, and that my insurance carrier may be billed for the cost of such treatment. However, no compensation can be provided for medical apart from the foregoing. I further understand that making such medical treatment available; or providing it, does not imply that such injury is the Investigator's fault. I also understand that by my participation in this study, I am not waiving any of my legal rights.

I understand that I may also contact the Chairman of the Committee on the use of Humans as Experimental Subjects, MIT 253-6787, if I feel I have been treated unfairly as a subject.

Signature

I have been fully informed as to the procedures to be followed, including those which are investigational, and have been given a description of the attendant discomforts, risks, and benefits to be expected. In signing this consent form, I agree to participate in the project and I understand that I am free to withdraw my consent and have this study discontinued at any time. I understand also that if I have any questions at any time, they will be answered.

Subject's Signature

Date

C.2 INFORMED CONSENT FORM FOR SLEEP SUBJECTS

NSBRI PI-IN-A-BOX GROUND STUDY

Purpose

We would like permission to enroll you in a research study. The purpose of this study is to evaluate the efficacy of an expert system called PI-in-a-box in identifying the presence of artifacts in sleep data and suggesting corrective procedures to eliminate these artifacts. A version of PI-in-a-box has already been developed to assist astronauts in performing a sleep experiment in space. This experiment is designed to quantify the effectiveness of an expert system in a laboratory environment in terms of both time and accuracy.

Participation in this study is voluntary and you are free to withdraw your consent and discontinue participation in the experiment at any time without prejudice.

Procedures

You will act as a sleep subject in two instrumenting test sessions per day in a span of two to three days. You will don a Sleep*Net, a web-like cap used to record electrophysiological signals such as EEG, EOG and EMG signals. A mildly abrasive cream will be used to scrub each electrode site prior to electrode application. Small adhesive discs will be used to apply the facial electrodes. During the test session, the technician will loosen or remove electrodes in the setup to deliberately introduce problems in the sleep signals.

Risks and Discomforts

A mild, abrasive cream will be used to scrub each electrode site prior to applying the electrodes. Minor irritation may result from this cleansing process.

Facial electrodes will be applied to the skin using small adhesive discs. The glue on these adhesives may cause minor discomfort or skin irritation.

Discomfort may be experienced as electrodes are poked out and put back into the Sleep*Net, but there will be breaks between test sessions, and you will not be required to wear the Sleep*Net for more than one hour at a time.

Benefits

A prorated payment of \$7.00 per hour will be provided to participants.

In the unlikely event of a physical injury resulting from participation in this research, I understand that medical treatment will be available from the MIT Medical Department, including first aid emergency treatment and follow-up care as needed, and that my insurance carrier may be billed for the cost of such treatment. However, no compensation can be provided for medical apart from the foregoing. I further understand that making such medical treatment available; or providing it, does not imply that such injury is the Investigator's fault. I also understand that by my participation in this study, I am not waiving any of my legal rights.

I understand that I may also contact the Chairman of the Committee on the use of Humans as Experimental Subjects, MIT 253-6787, if I feel I have been treated unfairly as a subject.

Signature

I have been fully informed as to the procedures to be followed, including those which are investigational, and have been given a description of the attendant discomforts, risks, and benefits to be expected. In signing this consent form, I agree to participate in the project and I understand that I am free to withdraw my consent and have this study discontinued at any time. I understand also that if I have any questions at any time, they will be answered.

Subject's Signature

Date

Appendix D

Subjective questionnaire

Please circle the level of rating you think best represents the following statements

STATEMENT	RATING											
<p>The level of accuracy of monitoring signal quality:</p> <p>Q1. Using [PI] assessments (colored lights)</p> <p>Q2. Observing the signal waveforms</p> <p>Q3. Using both [PI] assessments and observing signal waveforms</p>	Poor		Satisfactory		Good	1	2	3	4	5	6	7
<p>The effectiveness of figuring out the cause of the problem:</p> <p>Q4. Using [PI] diagnostics messages</p> <p>Q5. Using the troubleshooting procedures manual</p> <p>Q6. Using both [PI] messages and the troubleshooting procedures</p>						1	2	3	4	5	6	7
<p>Q7. How well did you understand the troubleshooting directions given in the procedures?</p> <p>Q8. How closely did you follow the troubleshooting directions given in the NASA guideline?</p> <p>Q9. How helpful were the [PI] diagnostics instructions in determining how to correct the problem?</p> <p>Q10. How effective was the training session in preparing you for the experiment?</p> <p>Q11. Describe the usefulness of [PI] as a completely autonomous decision-making tool</p> <p>Q12. Describe the usefulness of [PI] as a troubleshooting advisory tool</p>						1	2	3	4	5	6	7

Were you able to rectify the problem using the procedures?

If NO, please explain:

Please feel free to include any comments you may have about our experiment

Thank you for your time ...

Appendix E

Subject Comments

Subjects included the following comments in their subjective questionnaires:

- I think the biggest problem you will find is eliminating learning curve effects between first and second sessions.
- The PI diagnostics were much easier to use than the NASA Troubleshooting Guide, probably because the computer was already in front of me.
- I think that actual astronauts will tend to be more comfortable with the troubleshooting guide since most of their training is with it. Just a random thought.
- PI-in-a-box seems to provide very limited utility. More useful would be reasonably accurate estimates as to the cause of the problem, based on analysis of all the signals- e.g. A1 electrode problem when O2,C4, and EOGR are bad. Also, from a software design perspective, the interface is abysmal with inconsistent usage of interface elements.
- EMGs always yellow confusing.
- I wasn't able to do it well when [PI] displayed no signals and I was using the NASA Troubleshooting Guide only; in all other instances, I was able to rectify the problem using the procedures.
- I think that on a whole [PI] is a very good tool and certainly could be a beneficiary tool. I also think that the project was organized well and I didn't find the task of finding and solving the problems given very hard at all.
- I thought that using the [PI] was a lot easier than flipping through the manual. It was faster and the colored lights helped me to diagnose the problem sooner.
- For me it was easier to fix the problem by using the NASA Troubleshooting Guide. The [PI] messages were sort of confusing to me.
- Very well organized and easy to follow.
- I didn't catch some problems because there were multiple, unintended problems at times.
- I much prefer the [PI] guide to the NASA Troubleshooting Guide.
- [PI] does not use all available data to prioritize troubleshooting actions. For example, when the A1 or A2 reference electrodes are off, three signals show problems. Interchangeable use of hydrodot and electrode. Apparently many intermittent problems indicated by PI (false failures) that did not require corrective action. (threshold issue?). Procedures not consistent with diagnostics procedures vs. corrective actions.
- It seems like a good system. There were a few cases where there would be a flat line displays with a green light.
- It's been fun.
- Thanks, it was fun. Do we get any feedback on results?
- NASA Troubleshooting Guide doesn't have questions to ask, but procedures, so sometimes it was hard to figure out what to ask. [PI] was better in that it always gave you questions, but it is not smart enough to reorder the questions if multiple signals are bad, ex. Reference electrode affect three signals but [PI] can't tell.

- I noticed [PI] was a little slow sometimes in determining if a signal was bad, but it usually got a red light eventually. Also, the EMG light were never green, but always yellow. Also, the sleepnet seemed to behave better the second time - the signals stayed really steady except for the ones that are supposed to go bad; the first day the signals would go haywire every time they tried to make an error.
- Sometimes incorrectly diagnosed the problem, both with and without [PI] diagnostics.
- NASA Guide difficult to incorporate into the software procedures. There was too much going back and forth. The [PI] would perhaps be more helpful if it gave conditional suggestions dependent on the color of the indicator lights.
- Include the response time as a control factor, i.e. 30 seconds to detect the problem, and not more than that. I would say a 15 second window would make this more interesting.
- Diagnostics were clear. No errors were difficult, signals freeze or flat were too easy. Null error was most interesting error.
- Manual procedures not homogeneous (some require action and check, some require only check). Manual procedures less fit for 1 monitoring / 1 assisting than [PI] diagnostics
- It seemed like there was a lot of extra steps, probably for procedural sake: the troubleshooting manual was very confusing, it seems like that would be simplified.
- For some problems, the procedures are easy to follow, but for some it's rather confusing.
- Good experiment - it was hard to remember to use the event marker. The windows (to check) were possibly more useful than the NASA procedures - there are only a limited # of possible errors.
- I thought the training exercise/direction made the experiment pretty simple, and the signal quality was usually pretty good which made troubleshooting with [PI] and the NASA guide easy.
- [I was able to rectify the problem] every time except once, and then I misread the NASA flowchart, problem was fixable.
- I think [PI] should be able to identify when the A1 or A2 ref electrode is not functioning.
- Didn't need manual for A2.
- [I was able to rectify the problem] using the [PI] yes, without it, there were a couple of problems that I couldn't figure out what was wrong.
- Fun and interesting!
- Noticed EMGs were always yellow. Diagnostics were clear - but 'radio button' idea - so you cannot have more than one or two sets of diagnostics up. No errors too difficult - hair was hard to tell, not scrubbed. All signals flat was easy.
- There were a few times I had trouble determining what the problem was based on waveforms and indicator lights.
- This is a neat project... good luck!
- Sometimes I couldn't find the problem. NASA guidebook was tedious to flip through. Also, once I had gone through it, the second [time] was easier because I could almost tell what was wrong without guide/[PI]. But getting used to the kind of problems possible was the most helpful.
- Might want to remind us of pushing event marker. Could tell errors today were same as last time, but didn't know order.

