ESD-TR-91-199 MTR-10926

NMR Display Prototype Experiment Results

By

P. J. Miller

June 1991

Prepared for
Project Director
Airspace Management Systems Program Office
Electronic Systems Division
Air Force Systems Command
United States Air Force
Hanscom Air Force Base, Massachusetts

Approved for public release;
distribution unlimited

Project No. 6210
Prepared by
The MITRE Corporation
Bedford, Massachusetts
Contract No. F19628-89-C-0001

REVIEW AND APPROVAL

This technical report has been reviewed and is approved for publication.

DONALD R. THOMPSON, CAPT, USAF
New Mobile RAPCON Program Manager

FOR THE COMMANDER

DAVID A. HERRELKO, COL, USAF
Program Director
Airspace Management Systems Program Office

UNCLASSIFIED

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
June 1991

3. REPORT TYPE AND DATES COVERED
Final

4. TITLE AND SUBTITLE
NMR Display Prototype Experiment Results

5. FUNDING NUMBERS
F19628-89-C-0001
6210

6. AUTHOR(S)
Miller, Phillip J.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The MITRE Corporation
Burlington Road
Bedford, MA 01730

8. PERFORMING ORGANIZATION REPORT NUMBER
MTR-10926

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Project Director (ESD/TGN)
Airspace Management Systems Program Office
Electronic Systems Division, AFSC
Hanscom AFB, MA 01731-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
ESD-TR-91-199

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

An experiment was conducted to determine if air traffic controllers can perform as effectively using single intensity plasma panel displays with data discriminators as they can with monochrome multi-intensity CRT dispalys. Response times and error rates of 12 active duty, fully qualified USAF air traffic controllers (subjects) were used as measures of performance. After the results were compared, subject response times on the plasma display were found to be equal to those on the CRT. Under certain overlay conditions, error rates on the plasma display were significantly less than those on the monochrome reddish-orange CRT display. A study comparing operator performance between the plasma display and a monochrome green CRT display was also conducted.

Subjective ratings on various plasma and CRT display characteristics ranged from neutral to good. One hundred percent of the subjects found both plasma and CRT displays acceptable.

14. SUBJECT TERMS
Air Traffic Control
Display
Experiment
Plasma

15. NUMBER OF PAGES
67

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

UNCLASSIFIED

EXECUTIVE SUMMARY

A New Mobile Radar approach control (RAPCON) operations subsystem packaged within a 20 foot International Standard Organization (ISO) Shelter is being developed to accomplish the safe, orderly, expeditious launch and recovery of combat and support aircraft. It includes four single-intensity, multi-function flat panel plasma display consoles for four radar air traffic controllers. The consoles will be capable of operating in several modes: surveillance, flight data, training, and maintenance. Soft switches are provided for display control, category selection, and action entry.

While plasma display technology appears attractive for tactical mobile air traffic control applications where packaging, weight, and limited maintenance are primary system considerations, large screen plasma panel displays do not afford the same inherent levels of data discrimination attributes as those multi-intensity, monochrome CRT displays currently in use. Therefore, an investigation of radar controller performance measured by comparing display data identification errors and response times between plasma and CRT capabilities was conducted. In addition, subjective feedback was solicited through several questionnaires.

Twelve active duty, fully qualified USAF air traffic controllers viewed 48 static plasma scenario screens and 48 static CRT scenario screens. Two dependent variables, error rates, and response times, were used to measure the effect on performance of the experiment independent variables: monochrome reddish-orange CRT brightness intensity versus plasma sterile and non-sterile area discrimination.[1] Error rates and response times were employed to compare operator performance between the plasma display and a monochrome green CRT display. This study is documented in appendix A of this report.

An analysis of the experimental data shows that in the aggregate (the mean of all subject's means) there was not a significant difference in response times. However, the analysis of errors yielded a significant difference at the 0.05 probability level.[2] The differences were examined further by comparing plasma and CRT response times and error rates assocated with specific display data discrimination methods.

---

1. A sterile area is achieved by establishing bit plane priorities whereby, in the event of a data block overlaying map data, the data block content is written to the display and the underlying map data is not.

2. A 0.05 probability level indicates that there is a 5 percent chance that the results are incorrect.

When alphanumeric and target overlays on the CRT were displayed at different intensities (i.e., target symbols at 100 percent of relative brightness and flight identification information at 60 percent of relative brightness) and alphanumeric and target overlays on the plasma were displayed at the same intensity, there were significantly fewer errors on the plasma display than on the CRT display. Variances between the mean response times on the plasma and the CRT display were not significant at the 0.05 probability level.

The study comparing operator performance between the reddish-orange plasma and the monochrome green CRT display indicated operator performance to be significantly better when using the green CRT. While error rates were equal, response times decreased while using the green CRT. Whether this performance difference is the result of learning or a more favorable color is subject to further investigation.

There are several general points to be made regarding the results obtained from the questionnaires. On the average, all subjects rated all three displays; reddish-orange plasma, reddish-orange CRT, and the green CRT as "good" (5) on a scale of "extremely bad" (1) to "excellent" (7). In addition, 66 percent of the subjects felt that the plasma display was better than the CRT while 100 percent of the subjects found the following to be true:

1. both the CRTs and plasma displays were acceptable;
2. training was proper and complete;
3. the experiment setting was comfortable; and
4. there was no difficulty in remembering the questions.

The experiment results indicate that controllers can perform as effectively on the plasma display as they can on the CRT display. Furthermore, under certain overlay conditions, operator performance is improved (i.e., fewer errors) while using plasma displays. Recommendations for further CRT versus plasma experimentation should include the following comparisons of operator performance: under dynamic display conditions; between a reddish-orange plasma display and a monochrome green CRT; and between a plasma display and a multi-color CRT.

ACKNOWLEDGMENTS

This document has been prepared by the MITRE Corporation under Project No. 6210, Contract No. F19628-89-C-0001. The contract is sponsored by the Electronic Systems Division, Air Force Systems Command, United States Air Force, Hanscom Air Force Base, Massachusetts 01731-5000.

A special note of thanks goes to Lt. Col. D. W. Kavanagh (TGN), Capt. D. R. Thompson (TGN), Mr. R. Mahoney (HQ TAC/SCL), and HQ Air Force Communications Command for their support to the NMR Display Prototype Equipment.

In addition, software development and system integration was provided by F. C. Wendt, J. R. Leger, and J. Nelson with other major contributions provided by T. E. LaCorti, D. M. Selvitelli, and R. E. Fontaine.

Special acknowledgment is also given to E. M. Kelly and C. L. LaCrosse for their secretarial efforts throughout the experiment and the development of this report.

Finally, to those USAF air traffic controllers who served as subjects, I extend my thanks for your cooperation, patience, and effort.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE PAGE

LIST OF TABLES

SECTION 1

INTRODUCTION

The New Mobile Radar Approach Control (RAPCON) (NMR) will be used to accomplish the safe and expeditious launch and recovery of combat and support aircraft of all types from main operating bases, collocated operating bases, and contingency airstrips. In addition, the NMR will be used to replace inoperative fixed base radar approach controls. The principal function of the NMR is to provide adequate control of and separation among aircraft in a terminal radar approach control area.

The NMR will consist of a radar subsystem (AN/TPS-73) and an operations subsystem (OPS). The latter served as the catalyst for this experiment, which was performed in support to the OPS Demonstration Prototype Program.

The NMR operations subsystem [1] supports the display of aircraft track and plot data as received from the radar subsystem, the association of this data with flight identification information, and communications with aircraft and external ground based agencies. It includes four, multi-function, flat panel plasma display consoles for four radar controllers. A surveillance mode display provides plot, track, history, and associated flight identification information overlaid on a map background consisting of vectors and special geographic symbols. When detected by the radar, weather and jammer strobe data is also displayed. Soft switches are provided for display control, category selection, and action entry. The consoles are also capable of operating in three other modes: flight data, training, and maintenance.

Where packaging, weight, and limited maintenance are primary system considerations, plasma display technology in lieu of cathode ray tube (CRT) technology appears attractive for tactically mobile air traffic control applications. However, large screen plasma panel displays do not readily afford the same inherent levels of data discrimination attributes (e.g., grey scale) as those multi-intensity, monochrome CRT displays currently in use. CRT display technology, procedures, and human engineering designs within the air traffic control environment are well documented [2]. The opposite is true of plasma panel display technology. Therefore, an investigation of radar controller performance measured by comparing display data identification errors and response times between plasma and CRT capabilities was conducted. Since the multi-intensity, monochrome CRT display is used in virtually all USAF air traffic control systems (fixed and mobile), it served as the standard for comparison.

Two dependent variables, error rates and response times, were used to measure the effect on performance of the experimental independent variables: CRT brightness intensity versus plasma sterile, non-sterile area discrimination (figure 1). The error rate was the number of incorrect responses per subject for each of the 48 screens. The response time was the time from which the static scenario screen was first displayed to the time the subject initiated a request for the answer screen.

Twelve subjects viewed 48 plasma scenario screens and 48 CRT scenario screens. To protect against learning effects and fatigue, a random order of screen presentations for each subject was used. In addition, subjects (in group visits of four) were run under counterbalanced conditions (initial subject order was also selected randomly). Two subjects started first with the CRT and then with the plasma. The other two subjects used the plasma first followed by the CRT.

A study comparing subject performances between the plasma display and a green monochrome CRT display was also conducted. This study and its results are included as appendix A.

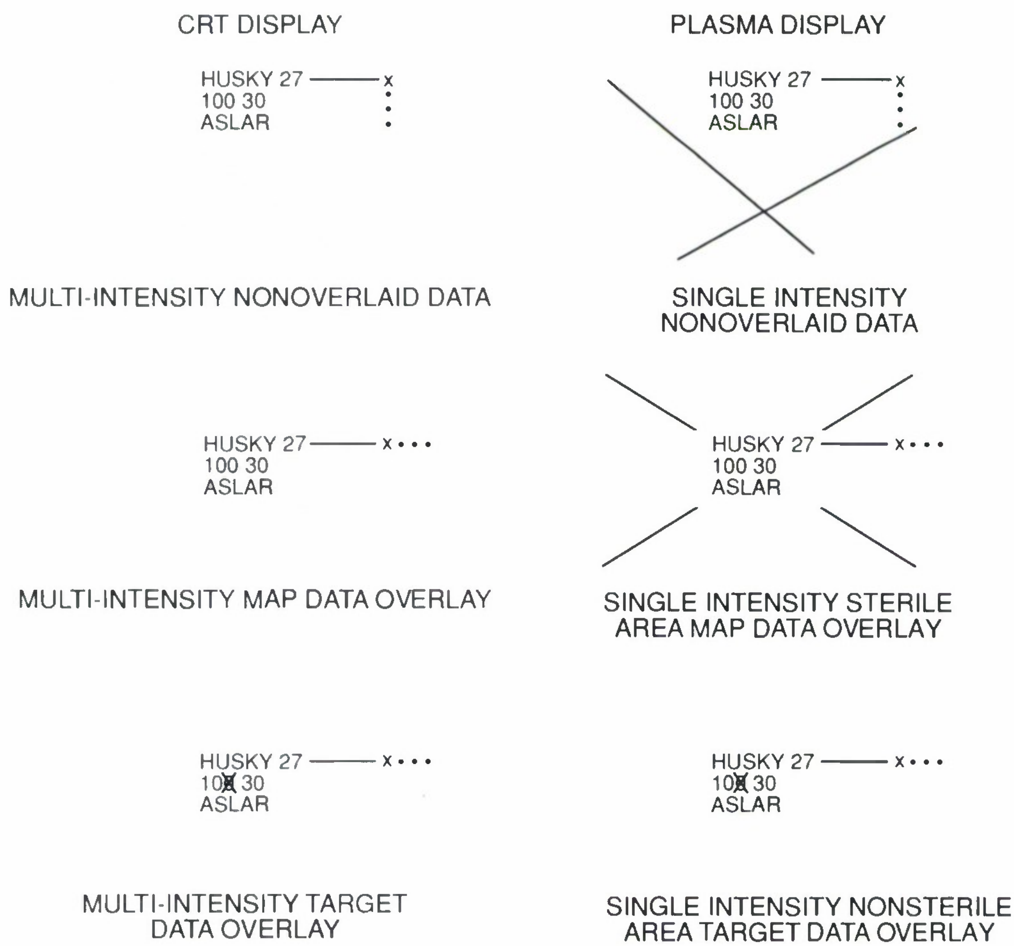


Figure 1. Discrimination Methods

SECTION 2

PURPOSE

The purpose of this experiment was to determine if radar air traffic controllers can perform as effectively using single intensity plasma panel displays with and without sterile area data discriminators as they can with monochrome reddish-orange CRTs using brightness intensity as a data discriminator.

SECTION 3

METHODOLOGY

Ninety-six static scenario screens (48 plasma, 48 CRT) consisting of overlaid and non-overlaid data were investigated under various implementations of data discrimination.

3.1 SUBJECTS

Twelve USAF radar air traffic controllers in groups of four served as subjects. These controllers were selected by Tactical Air Command's (TAC) Air Traffic Control (ATC) Division and drawn from a previously well screened population of 1,250 USAF radar air traffic controllers. This screening is accomplished as part of the medical requirements [3] for initial entry and continued performance in the USAF air traffic control field. Medical requirements include, but are not limited to, 20/40 or better near vision (natural or corrected) and the ability to distinguish red, green, and white colors. The subjects were/had recently been full performance controllers with an average of 9 years of journeyman radar approach control experience. In addition, they all had CRT and programmable indicator data processor (PIDP) experience. They were proficient in air traffic control type keyboard operations and familiar with a majority of the display formats used in the experiment.

3.2 APPARATUS

The experiment was conducted at The MITRE Corporation, Bedford, MA. As shown in figure 2, two PC Designs, GV386 personal computers were used to drive the CRT and plasma displays, to control all sequences of display events, and to perform data recording. The CRT was a 19-inch diagonal Hitachi color monitor operated as a multi-intensity monochrome monitor with reddish-orange color (1931 CIE Chromaticity Coordinates X=0.640, Y=0.350). It had an addressable resolution of 1280 x 1024 pixels at 93 pixels per inch with an active area of 11 inches x 11 inches. The AC plasma panel was a Photonics 14-inch x 14-inch monochrome reddish-orange (1931 Chromaticity Coordinates X=0.592, Y=0.353) display configured with an active area of 11 inches x 11 inches. It had an addressable resolution of 1024 x 1024 pixels at 73 pixels per inch. There was no perceived difference in color between the two displays. Each display had anti-reflective filters. Symbols and vectors were generated on each display with dedicated MATROX PG-1281 graphics processors capable of addressing 1280 x 1024 pixels in a double buffered display memory mode. Two QWERTY style keyboards (PC Designs), one for each display, were used for subject input.

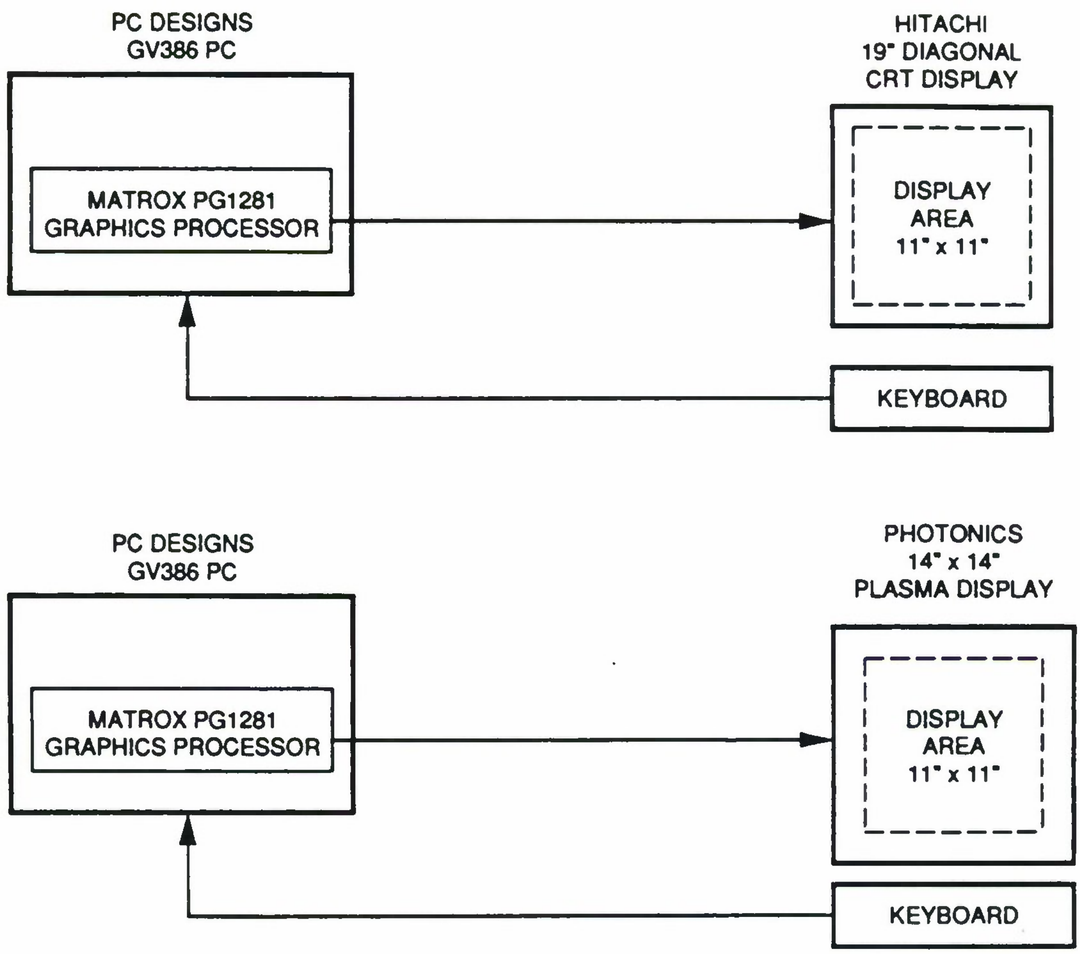


Figure 2. Experiment System Configuration

3.3 DISPLAY CONDITIONS

The CRT and plasma displays presented reddish-orange data on a black background. The CRT contrast ratio was 13:1 with 0.44 footcandles of light incident to the display surface. The plasma contrast ratio was 12:1 with 0.44 footcandles of light incident to the display surface. By using equivalent filters, and controlling for the experiment illumination environment, the display reflectance was minimized and held constant between the two displays.

Since the 19-inch CRT presented display information in an 11-inch x 11-inch square display area (1024 x 1024 pixels) centered on the CRT the plasma display presentation was scaled to display the same information in an 11-inch x 11-inch centrally located area (804 x 804 pixels) on the display.

To keep uppercase characters and special symbols for each display the same physical size, 9 x 11 (alphanumerics) and 12 x 12 (target and geographic symbols), CRT cell matrices were used versus 7 x 9 (alphanumeric) and 10 x 10 (target and geographic symbols) cell matrices for the plasma display. To minimize intra-symbol confusion [4], the Lincoln/MITRE font was used. A summary of display characteristics and provisions for their control are identified in table 1.

Table 1. Display Characteristics

Characteristics	CRT	Plasma
Addressable Pixels	1280 x 1024	1024 x 1024
Lines Per Inch	93	73
Contrast	13:1	12:1
Brightness	3.25 fL	2.0 fL
Frame Rate	60 Hz	60 Hz
Size	19" diagonal	14" x 14"
Active Area	11" x 11"	11" x 11"
Color	reddish-orange	reddish-orange
Alphanumeric Symbol Size	0.090" x 0.110" (9 x 11)	0.091" x 0.117" (7 x 9)
Special Symbol Size	0.120" x 0.120" (12 x 12)	0.100" x 0.100" (10 x 10)
Font	Lincoln/MITRE	Lincoln/MITRE
Off-Axis Viewing	>60°	>60°

The following groups of data were displayed at varying brightness intensities on the CRT. This variance was held constant throughout the experiment.

1. Map symbols and lines, weather areas, range rings, cursor	20% of selected brightness
2. Data block alphanumerics	60% of selected brightness
3. Plot, track, and history symbols	100% of selected brightness

3.4 DESCRIPTION OF DISPLAYED DATA

Each subject viewed 48 screens of static scenario data on each display. Heavy and average display loads were divided equally among the 48 screens. Table 2 gives the display loads. Heavy display loads were representative of anticipated wartime terminal approach control environments. The average display load was one half of a heavy display load. Figure 3 shows the display symbology, and table 3 the distribution of each target symbol for each scenario screen.

The 48 screens consisted of 16 screens on which targets and/or associated flight identification information did not overlay other data and 32 screens on which targets and/or associated flight identification information overlaid other data. Of the 32 screens, 16 included overlays of alphanumeric (flight identification information) over map data (lines) and 16 included instances of alphanumeric (flight identification information) over target data. As illustrated in figure 4, associated flight identification information consisted of a full or limited data block. Test stimuli (data blocks, map data and/or, target data, and associated overlays) were identical on each display. However, sterile area discriminators were used for instances of overlaid map and data block information on the plasma display and multiple intensity discriminators for instances of overlaid map and data block information on the CRT display. Figure 5 and table 4 define the specific test screen stimuli used for each of the 96 test screens (48 plasma, 48 CRT).

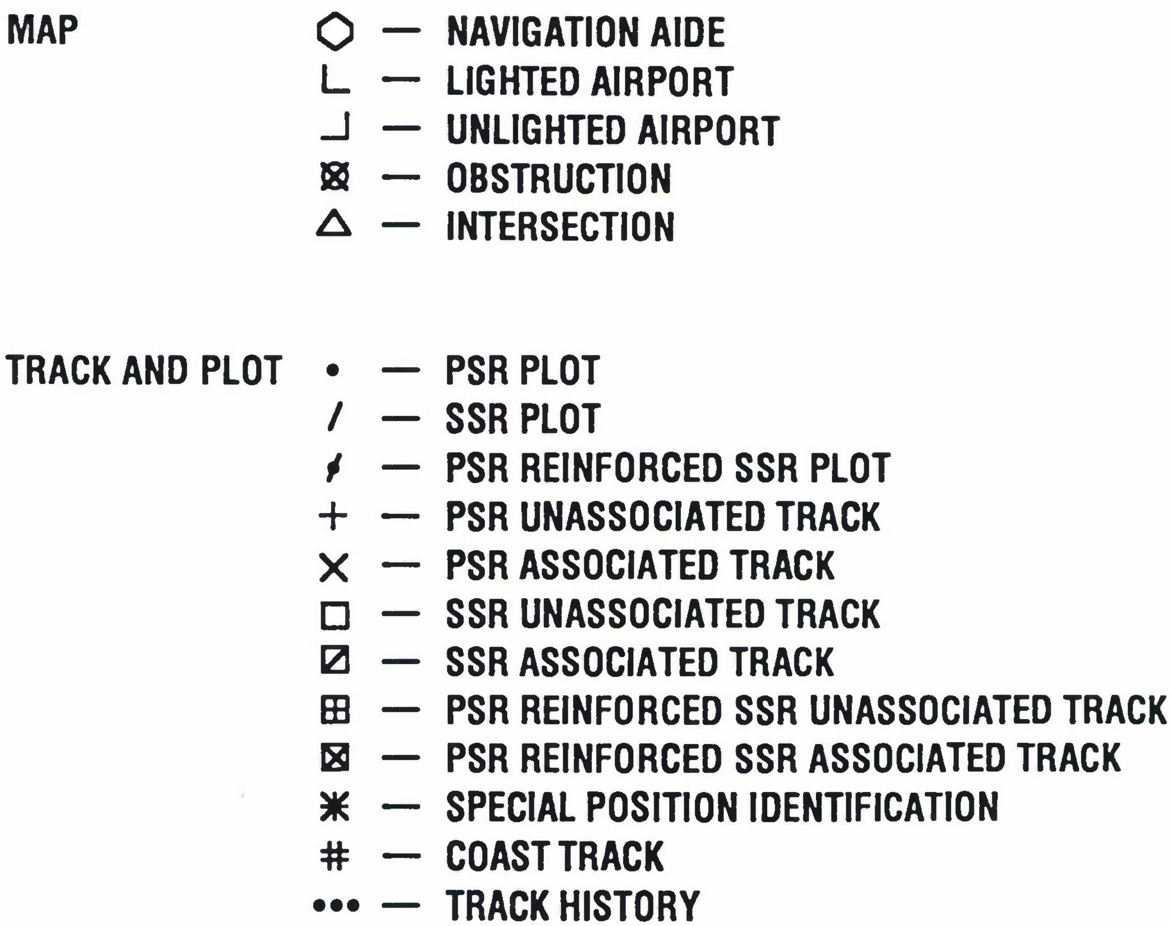
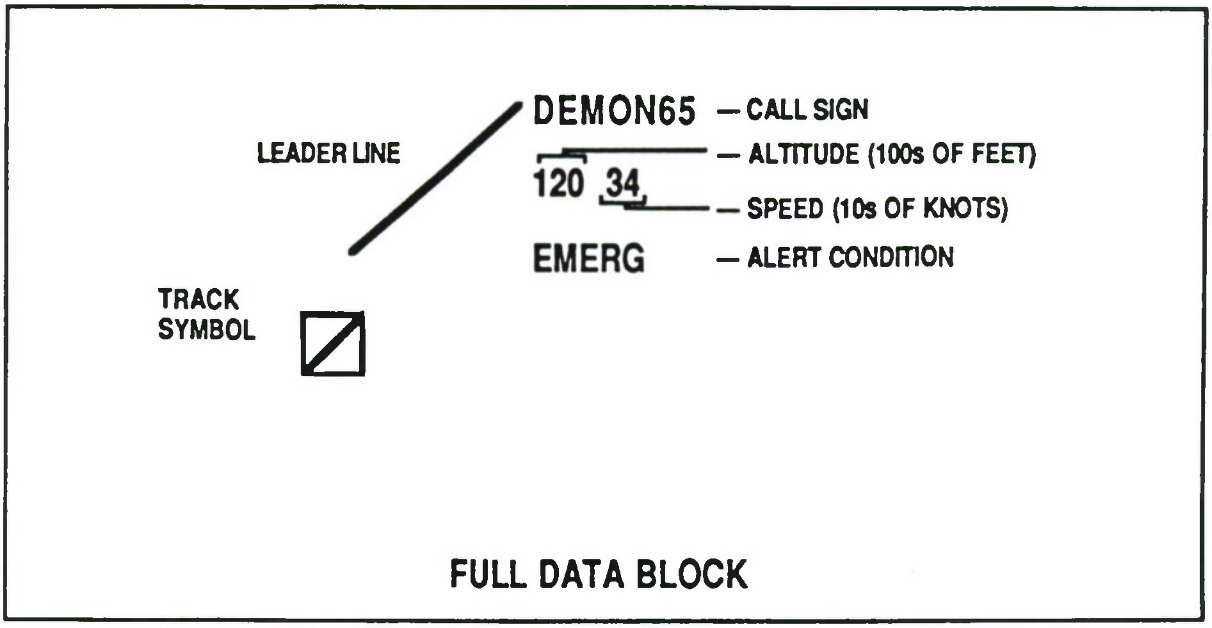


Figure 3. Display Symbology



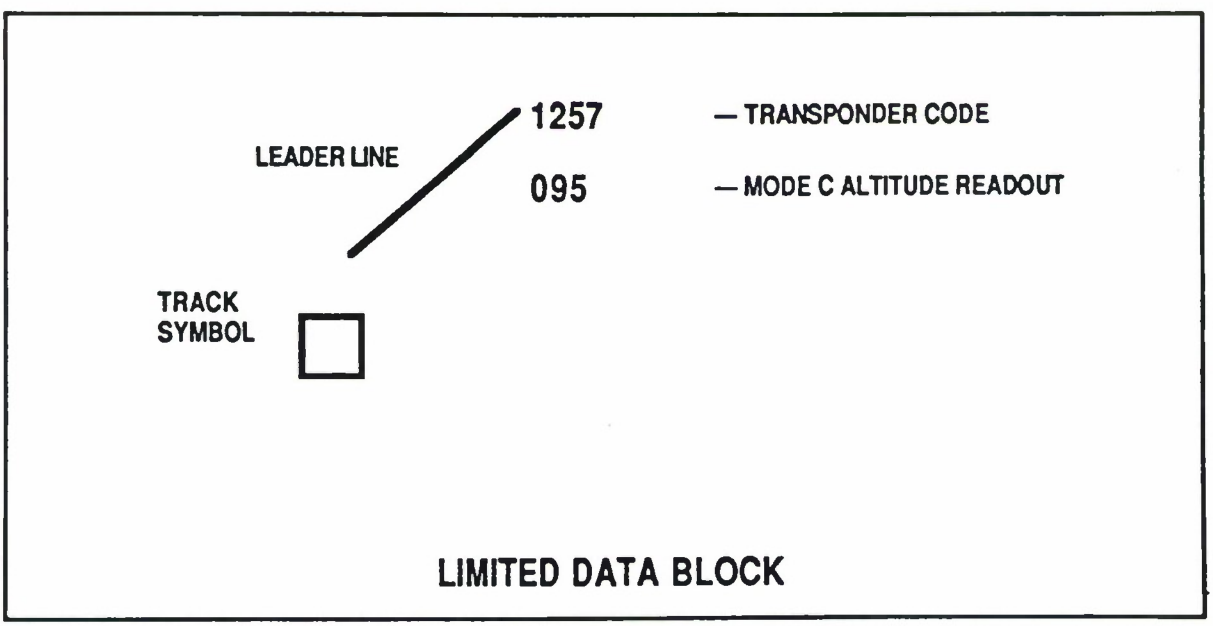


Figure 4. Full and Limited Data Blocks

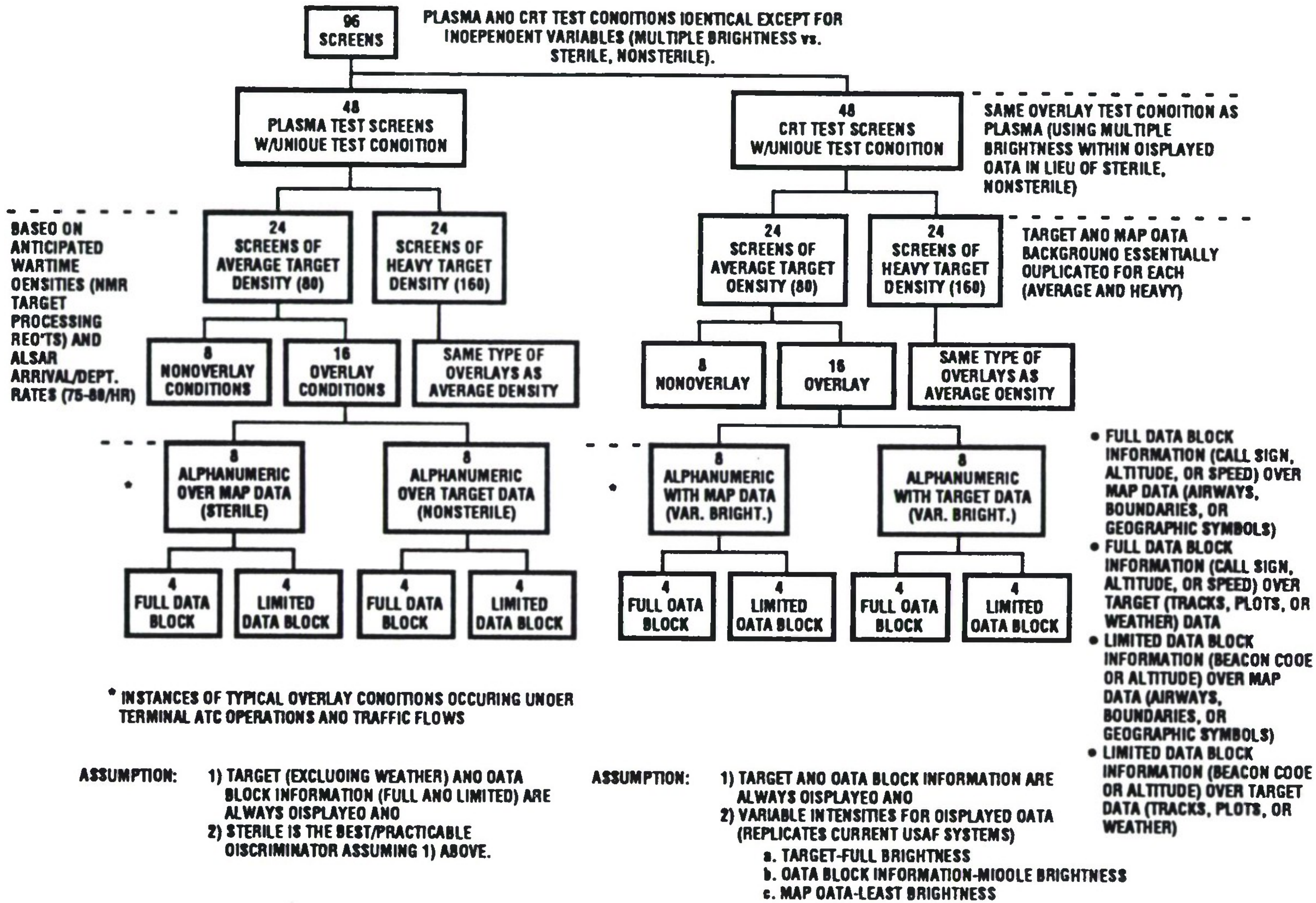


Figure 5. Test Screen Stimuli Hierarchy

Table 2. Display Loads

Data Type	Heavy	Average
Full Data Blocks (FDBs)	26	13
Limited Data Blocks (LDBs)	64	32
Alphanumeric Symbols for FDBs and LDBs[1]	708	484
Plot Symbols	50	25
Track Symbols	110	55
History Symbols	324	162
Leader Lines	70	35
Geographic Map Symbols	20	20
Map Lines	15	15
Weather Areas	1	1
Range Rings	3	3
Special Position Identification	2	1

Table 3. Target and History Symbol Frequency Per Scenario Screen

Symbol Type	Heavy	Average
PSR Plot	20	10
SSR Plot	20	10
PSR - Reinforced SSR Plot	10	5
PSR Unassociated Track	40	20
PSR Associated Track	4	2
SSR Unassociated Track	40	20
PSR - Reinforced SSR Unassociated Track	4	2
SSR Associated Track	14	7
PSR - Reinforced SSR Associated Track	4	2
Special Position Identification	2	1
Coast Track	2	1
History	330	165

1. Dissimilar flight identification information resulted in an average alphanumeric symbol display load exceeding the one half value of the heavy display load.

Table 4. Plasma and CRT Test Stimuli

96 Screens
48 Plasma and 48 CRT
(Same Test Stimuli)

Average Load (80 Targets)

A. Nonoverlay Conditions
Using Elements of 8 FDBs

1. 040
2. HUSKY 27
3. 20
4. A15721
5. 040
6. HUSKY 77
7. 100
8. MUGSY 27

B. 16 Overlay Conditions

o 8 Alphanumeric Over Map Data

- Using Elements of 4 FDBs

1. 020
2. 28
3. HUSKY 59
4. 050

- Using Elements of 4 LDBs
(Code or Altitude)

1. 160
2. 0301
3. 070
4. 065

Heavy Load (160 Targets)

A. Nonoverlay Conditions
Using Elements of 8 FDBs

1. DEMON 29
2. 080
3. 30
4. 100
5. R58764
6. 130
7. 30
8. 100

B. 16 Overlay Conditions

o 8 Alphanumeric Over Map Data

- Using Elements of 4 FDBs

1. 080
2. ASLAR
3. 135
4. 210

- Using Elements of 4 LDBS
(Code or Altitude)

1. 0341
2. 130
3. 080
4. 3423

Table 4 (Continued)

o 8 Alphanumeric Over Target Data	o 8 Alphanumeric Over Target Data
- Using Elements of 4 FDBS	- Using Elements of 4 FDBs
1. 018	1. 067
2. 30	2. DUSTY 22
3. 33	3. 30
4. TIG 66	4. 110
- Using Elements of 4 LDBs (Code or Altitude)	- Using Elements of 4 LDBs (Code or Altitude)
1. 075	1. 080
2. 190	2. 111
3. 133	3. 2224
4. 0345	4. 0433

3.5 PROCEDURE

Each group received four hours of training prior to the experiment. Training consisted of NMR display prototype system familiarization and a presentation on the purpose of the experiment activity. The experiment facilities were viewed, and the interactive portions of the experiment hardware were explained in detail. Representative experiment screens were presented, and practice runs were taken by each subject on each display. Subject training continued until each subject had reached a level of performance defined as 1) 100 percent correct scores on the display format test (appendix B) and 2) subject feedback that indicated comfortableness with the experiment system. In no case did subject training exceed four hours.

After the training sessions, subjects were exposed to the actual experiment scenarios for the first time, and measures were taken as to their performance. On average, each experimental run (48 screens) lasted approximately 30 minutes. Two subjects were run in parallel, one subject per display. Each subject, display, and associated equipment were located in separate labs. Subjects were seated facing the display at a distance of approximately 16 inches. Although instructed not to move the displays, physical arrangements permitted the subjects to view the displays from as close as desired. Chairs were adjusted to provide optimum positioning of the subjects. Keyboards were placed 3 inches from the front edge of the work surface.

Immediately prior to the presentation of each static scenario screen, the subject was prompted with a question screen on the display (see figure 6) to identify, at a given location (azimuth and range), an element of displayed data (e.g., call sign, altitude, speed, etc.). The question screens were structured to be consistent with standard ATC phraseology for conducting radar hand-offs. The key elements of these questions were annunciated by using uppercase characters. The question screens were timed (30 seconds) and controlled by the computer. If the question screen remained on the display for 30 seconds, the computer automatically displayed the scenario screen. Once subjects felt they had memorized the question, they then pressed the enter key. Following this keystroke, 1 of 48 static scenario screens appeared on the display (see figure 7). When the solicited displayed data was found by the subject, the enter key was depressed which caused the static scenario screen to disappear and a response screen (see figure 8) to be displayed. The subject then entered the answer into the system for data recording. Response times were measured from the time the static scenario screen was displayed to the subject to the time the subject requested the display of the response screen. Static scenario screens and response screens were also timed out after 30 seconds. The above sequence was repeated 48 times per subject per display.

Pretests with ten novice volunteers established that 30 seconds for each screen was sufficient to memorize the question, find the item of interest (stimuli), and answer the question. In addition, it provided the means to control for aberrant subject behavior (e.g., inattentiveness, leisureliness).

3.6 DESIGN AND ANALYSIS

Subject error rates and response times between the plasma and CRT displays were assessed by comparing each subject's performance on the plasma with their performance on the CRT (a within-subject design). Subject performance was compared in the aggregate (without regard to the specific discrimination method) and with respect to the discrimination method (see tables 5 and 6).

An analysis of variance (ANOVA) was used to test the significance of the error rate and the response time data.

Position: 260 at 18 miles

What is the ALTITUDE of HUSKY 27?

Figure 6. Sample Question Screen

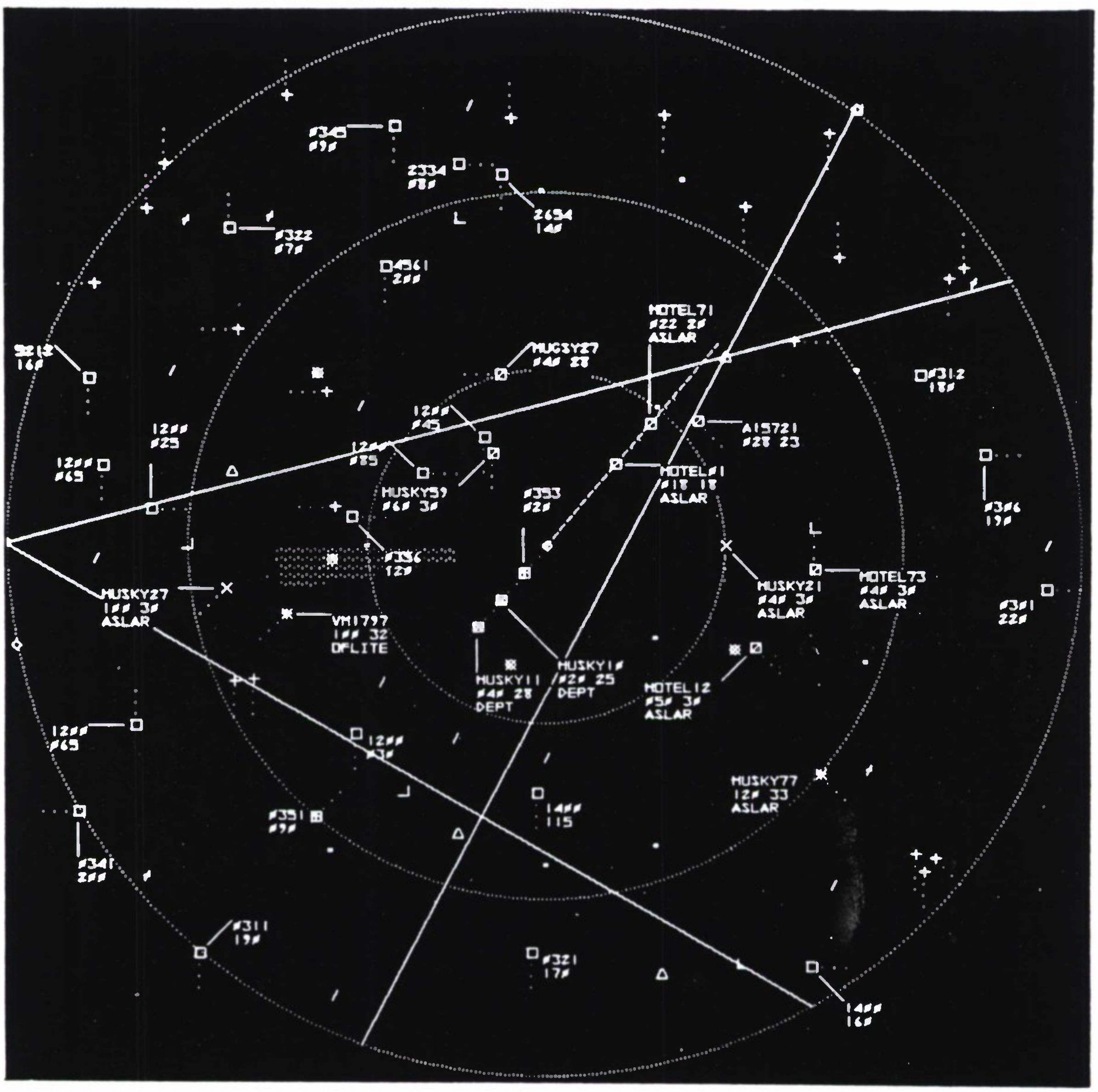


Figure 7. Sample Static Scenario Screen

Position: 260 at 18 miles

What is the ALTITUDE of HUSKY 27?

Answer: ________

Figure 8. Sample Answer Screen

Table 5. Experiment Design

DISCRIMINATION METHOD	48 PLASMA SCREENS	48 CRT SCREENS
METHOD 1	Single Intensity No Overlays (16 Screens)	versus Single Intensity No Overlays (16 Screens)
METHOD 2	Sterile Area Map Overlays (16 Screens)	versus Multi-Intensity Map Overlays (16 Screens)
METHOD 3	Single Intensity Target Overlays (16 Screens)	versus Multi-Intensity Target Overlays (16 Screens)

Table 6. Subject Sequence

GROUP	SUBJECTS	SEQUENCE	
		PLASMA FIRST CRT SECOND	CRT FIRST PLASMA SECOND
1	A, B, C, D	A, C	B, D
2	E, F, G, H	E, G	F, H
3	I, J, K, L	I, K	J, L

3.7 DATA COLLECTION, REDUCTION

The primary method of data collection was via software data recording. The software recorded all subject inputs to the system. This capability provided for off-line analysis of the subject's responses.

The software data recording function determined and recorded response times and error rates for each subject. Time-outs associated with an answer screen were recorded as errors. This data was then merged into a common data base from which raw data and statistical information was derived.

Subjective data was also compiled and compared. The subjective data, as derived from questionnaires, are contained in section 4.3.

SECTION 4

RESULTS

The results are presented in three parts:

1. Those dealing with aggregate error rates and response times.

2. Those dealing with error rates and response times associated with the following discrimination methods:

   a. Method 1 - single intensity non-overlaid plasma data versus single intensity non-overlaid CRT data;

   b. Method 2 - single intensity sterile area plasma data versus multi-intensity overlaid CRT data; and

   c. Method 3 - single intensity non-sterile area overlaid plasma data versus multi-intensity overlaid CRT data.

3. Those dealing with subjective data obtained from questionnaires.

4.1 AGGREGATE ERROR RATES AND RESPONSE TIMES

Figure 9 summarizes the mean response time and mean error rate performance for all subjects for all plasma and CRT static scenario screens. These aggregate error rates and response time mean scores are derived from the sum of individual subject error rates and the mean of subject's scores in the case of response times.

Mean response times for all practical purposes were identical. They were 6.7 seconds on the CRT and 6.1 seconds on the plasma. The standard deviations for response times on the plasma and CRT displays were 2.5 and 2.4 respectively. Mean error rates on the CRT were 4.2 incorrect responses compared to mean error rates on the plasma of 2.0.

A simple analysis of variance (appendix C) was performed for both the response time and error rate scores. The analyses for response times yielded an F value of 0.79 which was not significant at the 0.05 probability level. The analyses for error scores yielded an F value of 9.72 which was significant at the 0.05 probability level. In section 4.2, these differences are examined further by considering the association of response times and error rates with discrete discrimination methods.

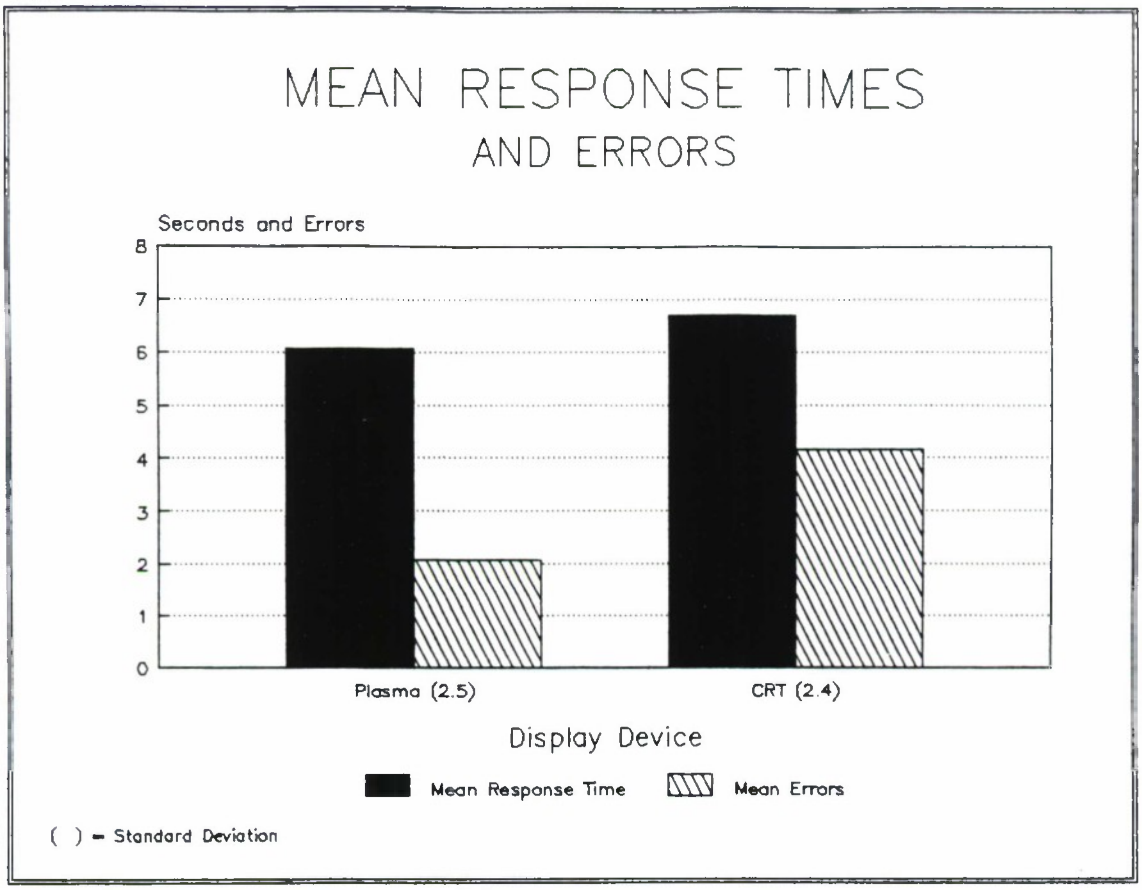


Figure 9. Mean Response Times and Errors

Individual error rates for each display are illustrated in figure 10. Subject mean response times for each display are illustrated in figure 11.

4.2 DISCRETE ERROR RATES AND RESPONSE TIMES

Categorized by data discrimination method, figures 12 through 17 show individual subject error rates and mean response times. Figures 18 and 19 show the mean error rates and the mean response times as derived from the sum of subject mean scores and categorized by data discrimination methods.

Grouped according to data discrimination method, plasma and CRT mean error rates and mean response times and their significance at the 0.05 probability level are summarized in table 7.

Table 7. Experiment Results

DISCRIMINATION CATEGORY	PLASMA	CRT	F VALUE	SIGNIFICANCE AT THE 0.05 PROBABILITY LEVEL
METHOD 1				
Mean Error	0.3	0.7	0.32	Not Significant
Mean Response Time	5.2	6.0	4.14	Not Significant
METHOD 2				
Mean Error	0.3	0.8	2.10	Not Significant
Mean Response Time	5.5	6.0	0.43	Not Significant
METHOD 3				
Mean Error	1.5	2.8	4.88	Significant
Mean Response Time	7.5	8.1	0.83	Not Significant

As illustrated in table 7, the analysis of variance (appendix C) for error scores yielded F values of 0.32 and 2.10 for methods 1 and 2 above, which were not significant at the 0.05 probability level and 4.88 for method 3 above which is significant at the 0.05 probability level.

Furthermore, an analysis of variance for the response times yielded F values of 4.14, 0.43, and 0.83 for methods 1 through 3, respectively, which were not significant at the 0.05 probability level.

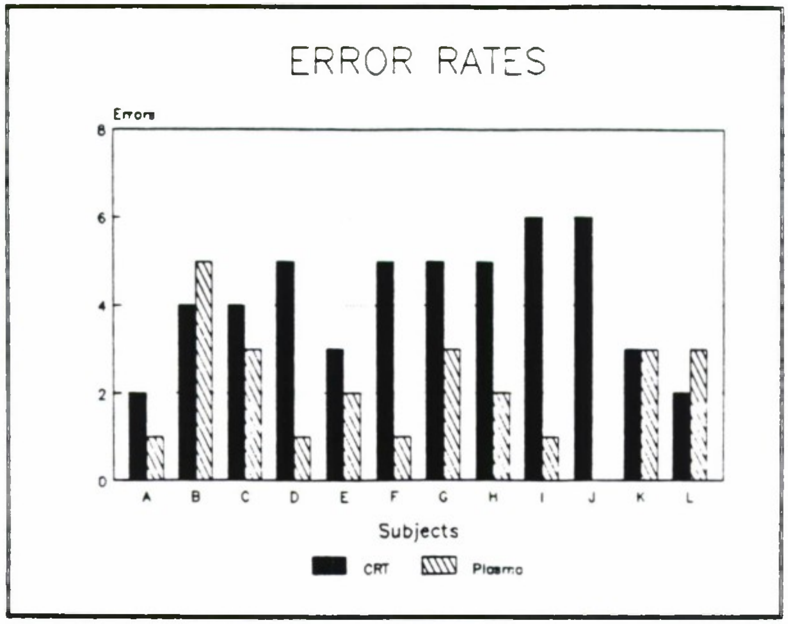


Figure 10. CRT versus Plasma Subject Errors

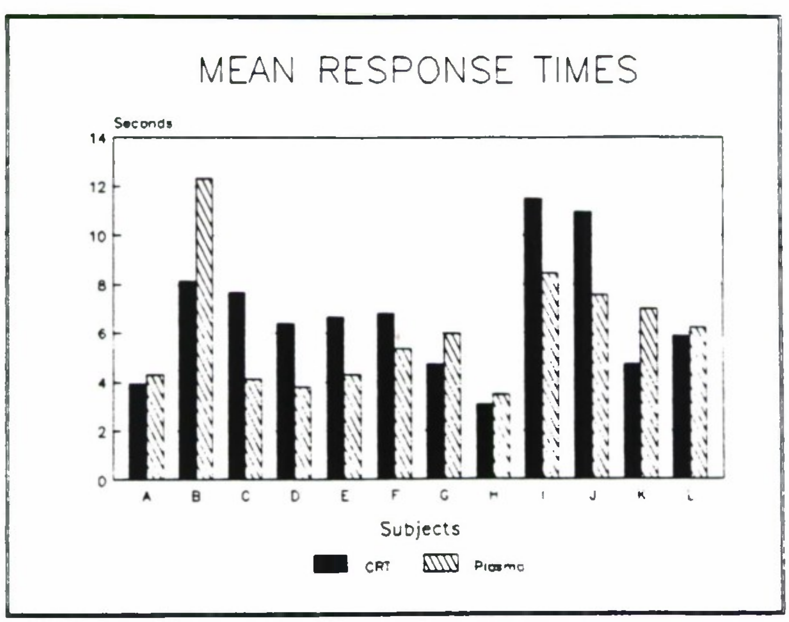


Figure 11. CRT versus Plasma Subject Mean Response Times

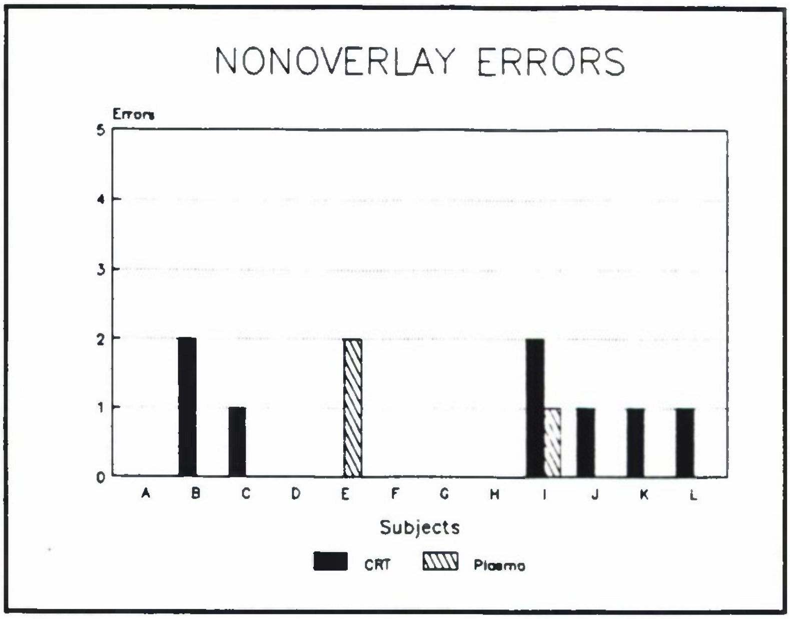


Figure 12. Discrimination Method 1 - Errors by Subject

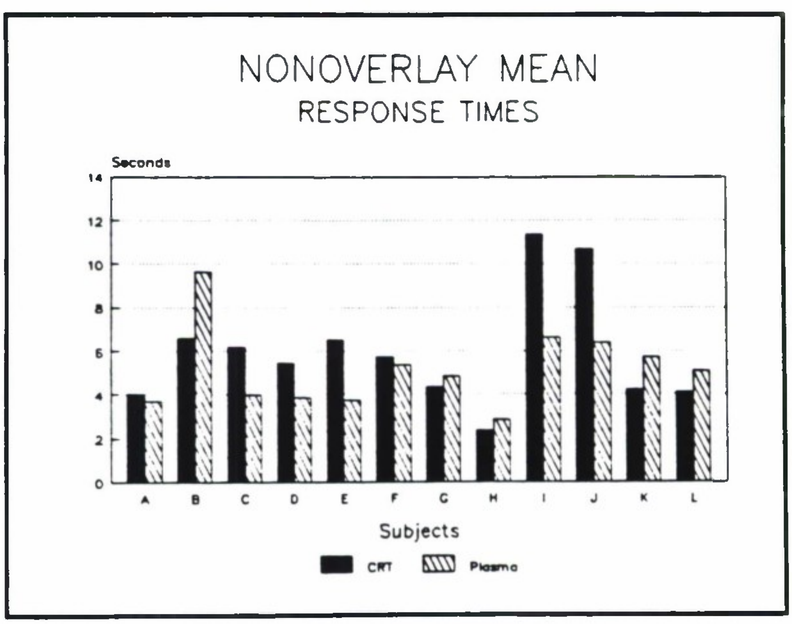


Figure 13. Discrimination Method 1 - Mean Response Times by Subject

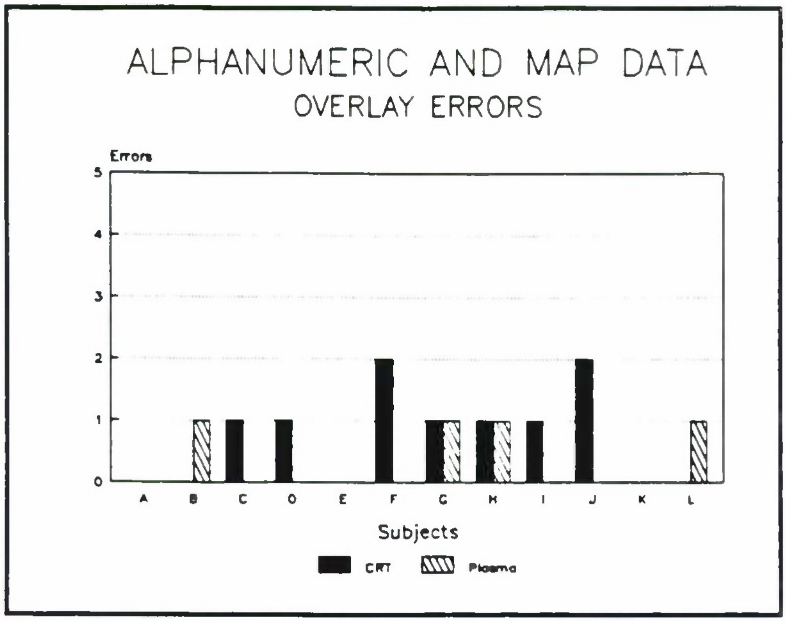


Figure 14. Discrimination Method 2 - Errors by Subject

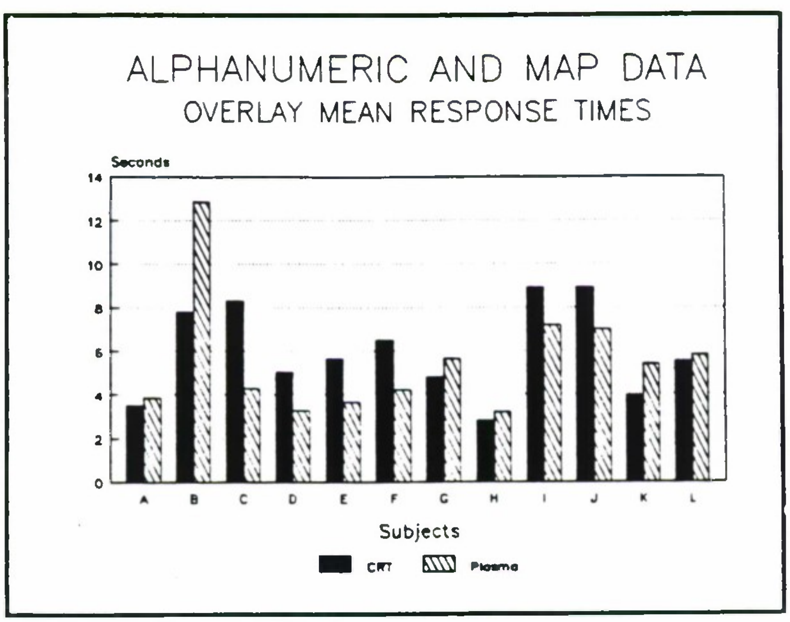


Figure 15. Discrimination Method 2 - Mean Response Times by Subject

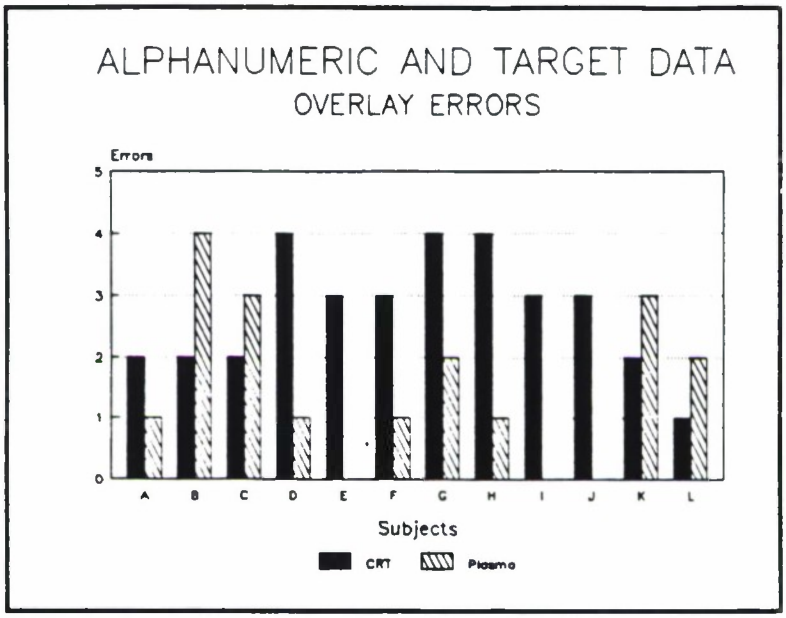


Figure 16. Discrimination Method 3 - Errors by Subject

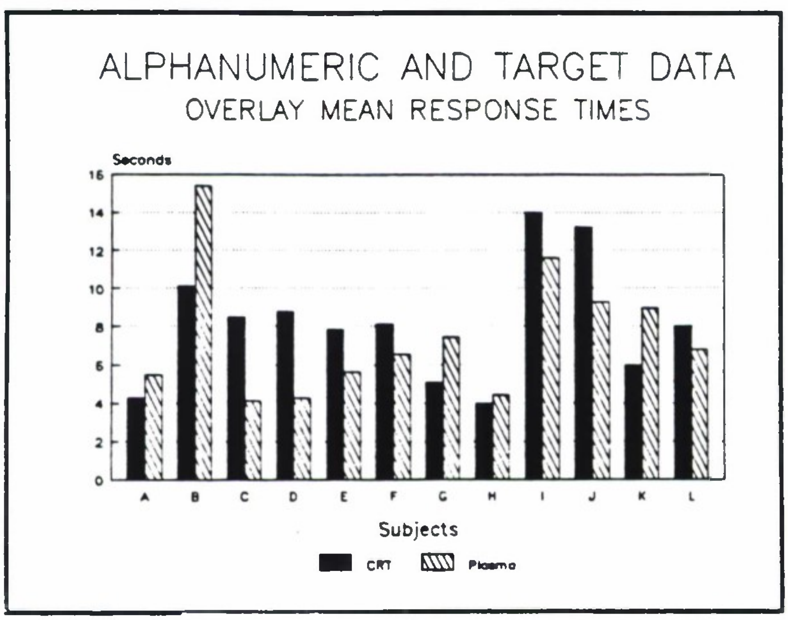


Figure 17. Discrimination Method 3 - Mean Response Times by Subject

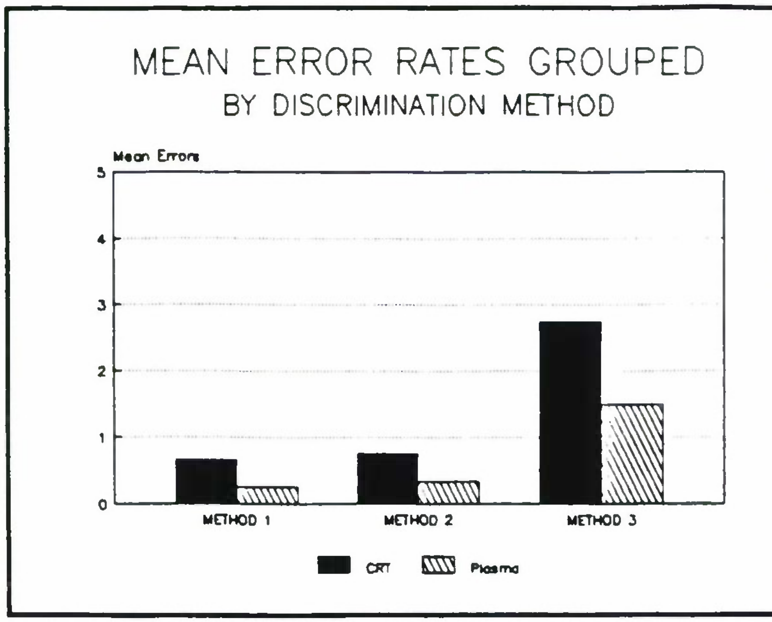


Figure 18. Mean Error Rates Grouped by Data Discrimination Method

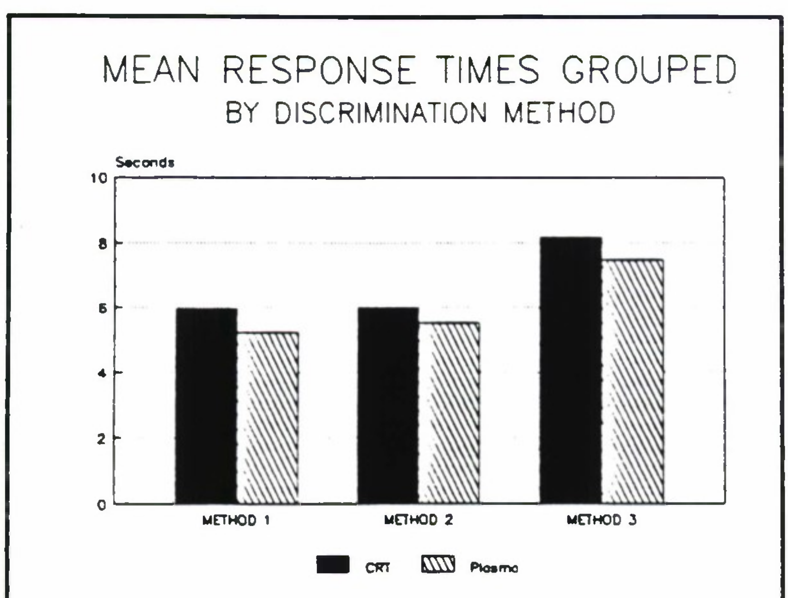


Figure 19. Mean Response Times Grouped by Data Discrimination Method

4.3 SUBJECTIVE DATA

Several questionnaires were used to investigate issues relative to the experiment design and to CRT (reddish-orange and green) and plasma display technologies. They included an agreement/disagreement questionnaire and a display characteristics rating questionnaire. The agreement/disagreement questionnaire, with the percentages of subject disagreement and agreement, is presented in table 6. There are several general points to be made regarding the obtained results. They include the following:

a. sixty-six percent agreed (from moderately to definitely) that the plasma display, overall, was better than the CRT;

b. one hundred percent of the subjects found both the CRT and plasma displays acceptable;

c. fifty-eight percent of the subjects rated the green CRT superior to the orange CRT;

d. fifty-nine percent reported that the color of the display does matter; and

e. sixteen percent reported fatigue with either the plasma or CRT display.

Subject responses to the display characteristics rating questionnaire are summarized in figure 20. On the average, all subjects rated all three displays: reddish-orange plasma, reddish-orange CRT, and green CRT as "good" (5) on a scale of "extremely bad" (1) to "excellent" (7).

Table 8. Agreement/Disagreement Questionnaire

1. I think the plasma is superior to the CRT

Definitely Disagree	Generally Disagree	Moderately Disagree	Moderately Agree	Generally Agree	Definitely Agree
	17%	17%	17%	41%	8%

2. I think the CRT is superior to the plasma

Definitely Disagree	Generally Disagree	Moderately Disagree	Moderately Agree	Generally Agree	Definitely Agree
8%	25%	34%	25%	8%	

Table 8 (Continued)

3. I think plasma and CRT displays are equal

Definitely Disagree	Generally Disagree	Moderately Disagree	Moderately Agree	Generally Agree	Definitely Agree
17%	33%	17%	8%	25%	

4. I found the plasma acceptable

Definitely Disagree	Generally Disagree	Moderately Disagree	Moderately Agree	Generally Agree	Definitely Agree
			34%	41%	25%

5. I found the CRT acceptable

Definitely Disagree	Generally Disagree	Moderately Disagree	Moderately Agree	Generally Agree	Definitely Agree
			34%	41%	25%

6. I believe the green CRT is superior to the orange CRT

Definitely Disagree	Generally Disagree	Moderately Disagree	Moderately Agree	Generally Agree	Definitely Agree
	25%	17%	17%	25%	16%

7. The color of the display does not matter

Definitely Disagree	Generally Disagree	Moderately Disagree	Moderately Agree	Generally Agree	Definitely Agree
25%	17%	17%	17%	16%	8%

8. I experienced fatigue with the plasma

Definitely Disagree	Generally Disagree	Moderately Disagree	Moderately Agree	Generally Agree	Definitely Agree
59%	25%		8%	8%	

Table 8 (Concluded)

9. I experienced fatigue with the CRT

Definitely Disagree	Generally Disagree	Moderately Disagree	Moderately Agree	Generally Agree	Definitely Agree
67%	17%		8%	8%	

10. I believe the experiment ran smoothly

Definitely Disagree	Generally Disagree	Moderately Disagree	Moderately Agree	Generally Agree	Definitely Agree
			8%	17%	75%

11. I was trained properly and completely

Definitely Disagree	Generally Disagree	Moderately Disagree	Moderately Agree	Generally Agree	Definitely Agree
				8%	92%

12. Traffic densities were representative of anticipated wartime environments

Definitely Disagree	Generally Disagree	Moderately Disagree	Moderately Agree	Generally Agree	Definitely Agree
	17%		33%	33%	17%

13. There was sufficient time to read and understand the question

Definitely Disagree	Generally Disagree	Moderately Disagree	Moderately Agree	Generally Agree	Definitely Agree
				8%	92%

14. I did not have any difficulty remembering the questions presented

Definitely Disagree	Generally Disagree	Moderately Disagree	Moderately Agree	Generally Agree	Definitely Agree
		8%	8%	50%	34%

15. I felt comfortable with the experiment setting

Definitely Disagree	Generally Disagree	Moderately Disagree	Moderately Agree	Generally Agree	Definitely Agree
			8%	17%	75%

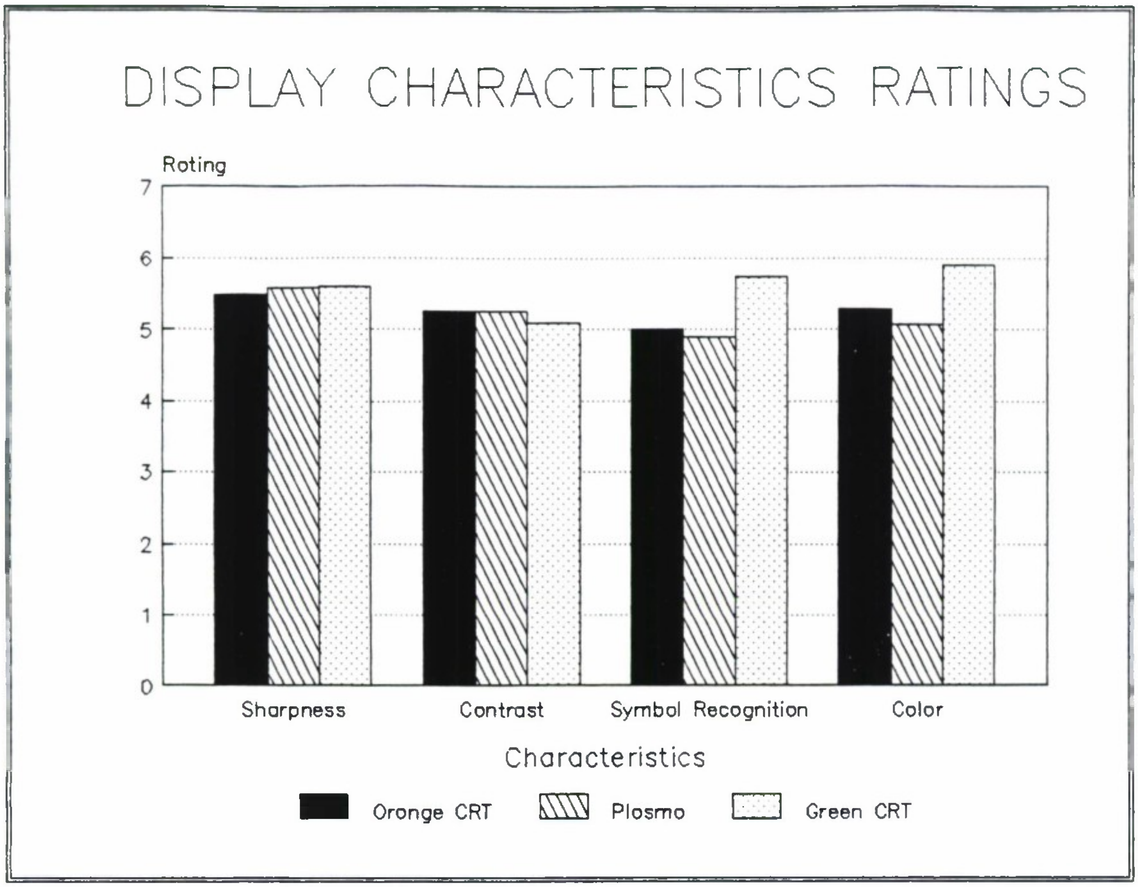


Figure 20. Display Characteristics Ratings

SECTION 5

DISCUSSION

The experiment showed that subject performance on the plasma is at least equal to, and in the case of alphanumeric and target symbol overlays, enhanced when compared to the CRT.

As noted previously, subject performance on the plasma, as measured by response times, was equal to subject performance on the CRT. Analyzing the response time data in the aggregate and across discrimination methods served to reinforce this conclusion.

As measured by error rates, subject performance on the plasma was equal to subject performance on the CRT under the following discrimination condition comparisons:

a. single intensity non-overlaid plasma data versus single intensity non-overlaid CRT data (Method 1); and

b. single intensity sterile area plasma data versus multi-intensity overlaid CRT data (Method 2).

However, as a result of this experiment, there is significant evidence that subject error rates, under certain conditions, decrease with the plasma. This variance is significant under Method 3 when alphanumeric data overlays target data at the same intensity on the plasma is compared with alphanumeric data at one intensity level (60 percent of selected brightness) overlaying target data at another intensity level (100 percent of selected brightness) on the CRT. There is a need to examine why subject performance is enhanced on the plasma without using sterile area discrimination and without the inherent CRT benefit of variable intensity discrimination within displayed data. This dilemma presents both a challenge and an opportunity for further research.

Finally, while the present experiment considered the effectiveness of multi-intensity discriminators on the CRT display versus sterile and non-sterile discriminators on the plasma display, there remains the need to systematically study the effects of the same independent variables within the context of dynamic scenario screens. In addition, with the USAF's eventual move to replace their fixed base monochrome CRT displays, the need to compare controller performance between monochrome plasma and multi-color CRT displays appears warranted.

REFERENCES

*1. Bloom, P. J., K. W. Eiben, and P. J. Miller, 1987, *New Mobile RAPCON (NMR) Operations Subsystem Segment Specification,* MTR 10287, The MITRE Corporation, Bedford, MA.

2. Kircher, J. H. and W. Laurig, 1971, *The Human Operator in Air Traffic Control,* Ergonomics (14) (15), 549-556.

3. Shurtleff, D. A., 1980, *How to Make Displays Legible,* Human Interface Design, La Miranda, CA.

*4. AFR 160-43, *Medical Examination and Medical Standards,* Department of the Air Force, Washington, DC.

* These documents have not been reviewed by the Directorate for Security Review and are therefore not available for public dissemination.

APPENDIX A

PLASMA VERSUS GREEN CRT STUDY

A.1 INTRODUCTION

A study was conducted to determine if radar air traffic controllers can perform as effectively using single intensity plasma panel displays, with and without data discriminators (sterile areas), as they can with monochrome green CRTs using brightness intensity as a data discriminator.

The methodology for this study was almost identical to the plasma versus reddish-orange CRT experiment methodology. Therefore, for the purpose of this report only differences with the experimental methodology will be discussed.

The green CRT data collected as a result of this study was compared with the plasma display data obtained from the plasma versus orange CRT experiment.

A.2 METHODOLOGY

Ninety-six static scenario screens (48 plasma, 48 CRT) consisting of overlaid and non-overlaid data were investigated under various implementations of data discrimination.

A.2.1 APPARATUS

The CRT (the same one that was used for the experiment) was operated as a multi-intensity monochrome monitor with green color (1931 CIE chromaticity coordinates X=0.200, Y=0.590). One PC design, the GV386 personal computer, was used to drive the CRT target display, control all sequences of display events, and to perform data recording.

A.2.2 DISPLAY CONDITIONS

The CRT presented green data on a black background.

A.2.3 DESCRIPTION OF DISPLAYED DATA

The same subjects used in the experiment viewed 48 screens of static scenario data on the green CRT display. Question, scenario, and answer screens were identical to those used during the reddish-orange plasma versus reddish-orange CRT display experiment.

A.2.4 PROCEDURE

Subjects in each group were run in succession. Their order was randomly determined.

A.2.5 DESIGN AND ANALYSIS

Subject error rates and response times between the reddish-orange plasma display and the green CRT display were compared in the aggregate and across discrimination methods.

An ANOVA was used to test the significance of the error rate and response time data.

A.3 RESULTS

The results are presented in two parts:

1. Those dealing with aggregate error rates and response times.

2. Those dealing with error rates and response times based upon the following discrimination methods:

   a. Method 1 - single intensity non-overlaid plasma data versus single intensity non-overlaid CRT data;

   b. Method 2 - single intensity sterile area plasma data versus multi-intensity overlaid CRT data; and

   c. Method 3 - single intensity non-sterile area overlaid plasma data versus multi-intensity overlaid CRT data.

A.3.1 AGGREGATE ERROR RATES AND RESPONSE TIME

Response time mean scores derived from the sum of subject's mean scores and mean error rates are shown in figure A-1.

Mean response times were 4.7 seconds on the green CRT display and 6.1 seconds on the plasma display. The standard deviation for response times on the plasma and CRT displays were 2.5 and 1.5 respectively. Mean error rates on the green CRT display were 2.9 incorrect responses, while mean error rates on the plasma display were 2.0 incorrect responses.

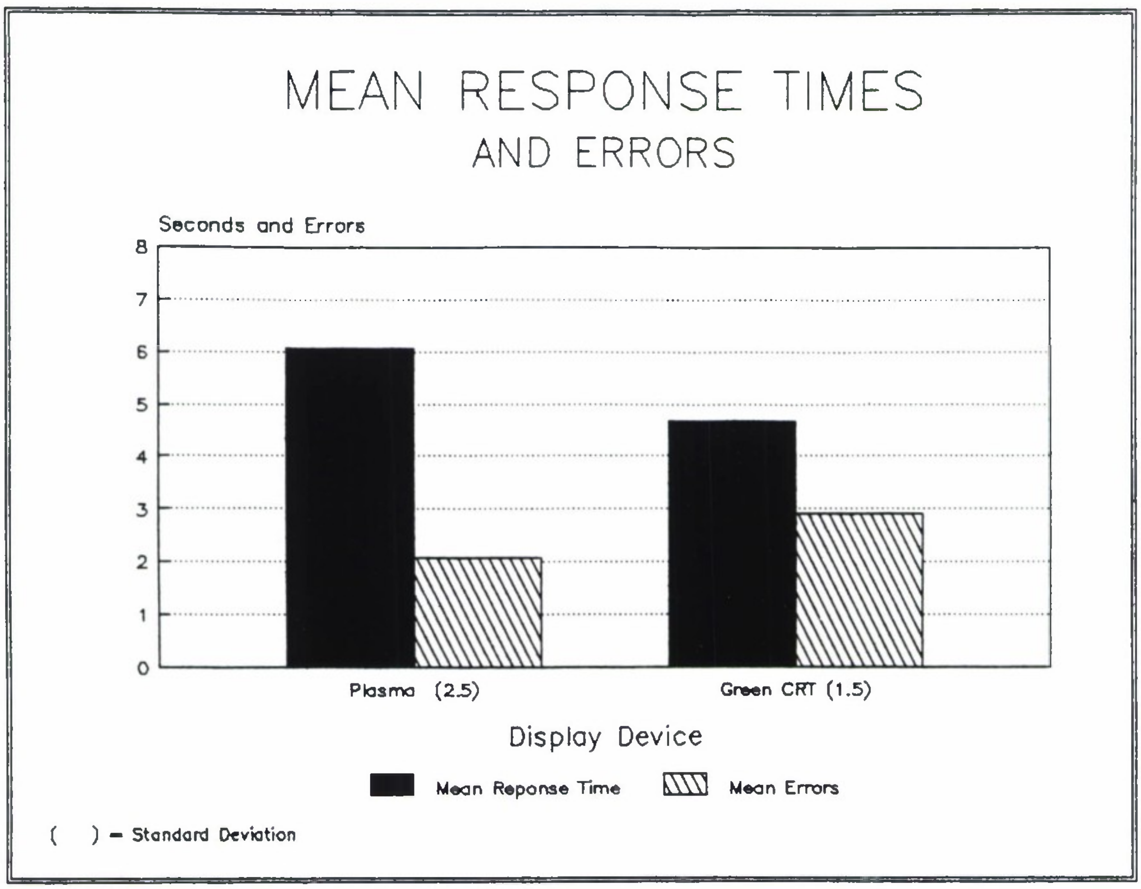


Figure A-1. Mean Response Times and Errors

A simple ANOVA was performed in both the response time and error rate scores. The analyses for response times yielded an F value of 7.02 which was significant at the 0.05 probability level. The analyses for error scores yielded an F value of 2.31, which was not significant at the 0.05 probability level. This trend is examined further by considering the association of response times and error rates with discrete data discriminators.

Subject mean response times for each display are illustrated in figure A-2. Individual error rates are illustrated in figure A-3.

A.3.2 DISCRETE ERROR RATES AND RESPONSE TIMES

Categorized by data discrimination method, figures A-4 through A-9 show individual subject error rates and mean response times. Figures A-10 and A-11 show the mean error rates and mean response times as derived from the sum of subject mean scores and categorized by data discrimination methods.

Grouped according to data discrimination method, plasma and CRT mean error rates and mean response times and their significance at the 0.05 probability level are summarized in table A-1.

Table A-1. Study Results

DISCRIMINATION CATEGORY	PLASMA	CRT	F VALUE	SIGNIFICANCE AT THE 0.05 PROBABILITY LEVEL
METHOD 1				
Mean Error	0.3	0.3	2.10	Not Significant
Mean Response Time	5.2	4.3	3.92	Not Significant
METHOD 2				
Mean Error	0.3	0.3	0.013	Not Significant
Mean Response Time	5.5	4.2	7.17	Significant
METHOD 3				
Mean Error	1.5	2.3	3.09	Not Significant
Mean Response Time	7.5	4.6	13.18	Significant

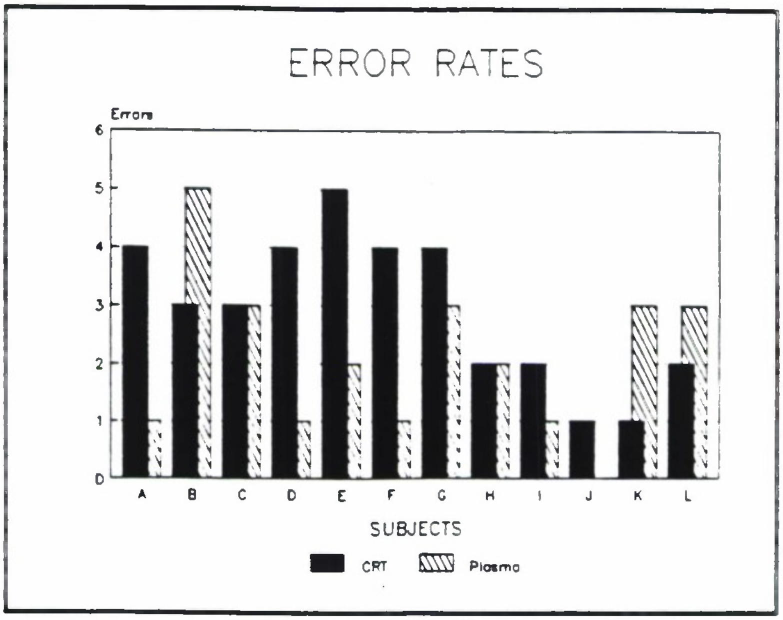


Figure A-2. CRT versus Plasma Subject Error Rates

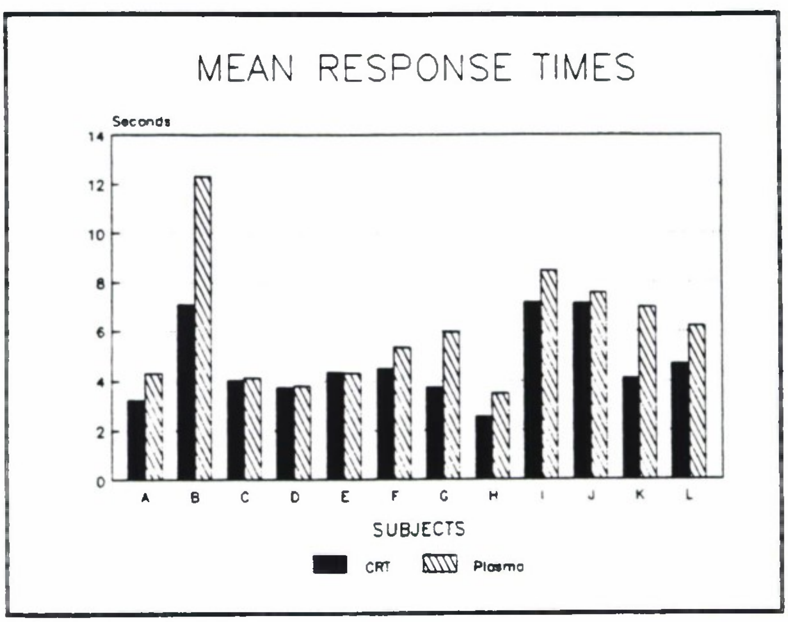


Figure A-3. CRT versus Plasma Subject Mean Response Times

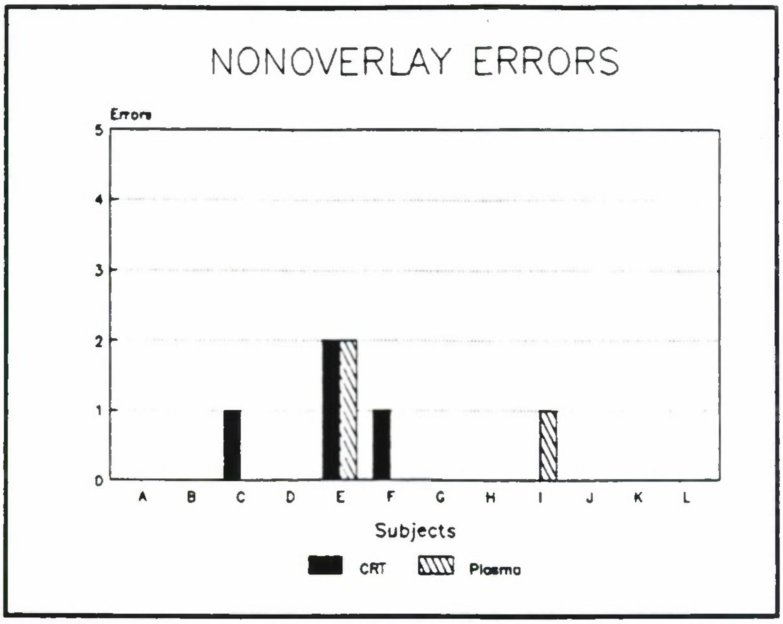


Figure A-4. Discrimination Method 1 - Errors by Subject

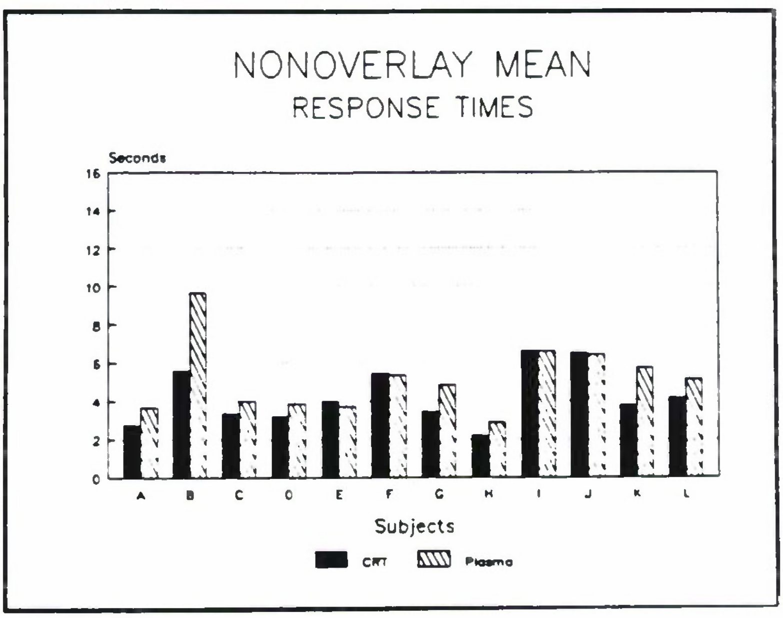


Figure A-5. Discrimination Method 1 - Mean Response Times by Subject

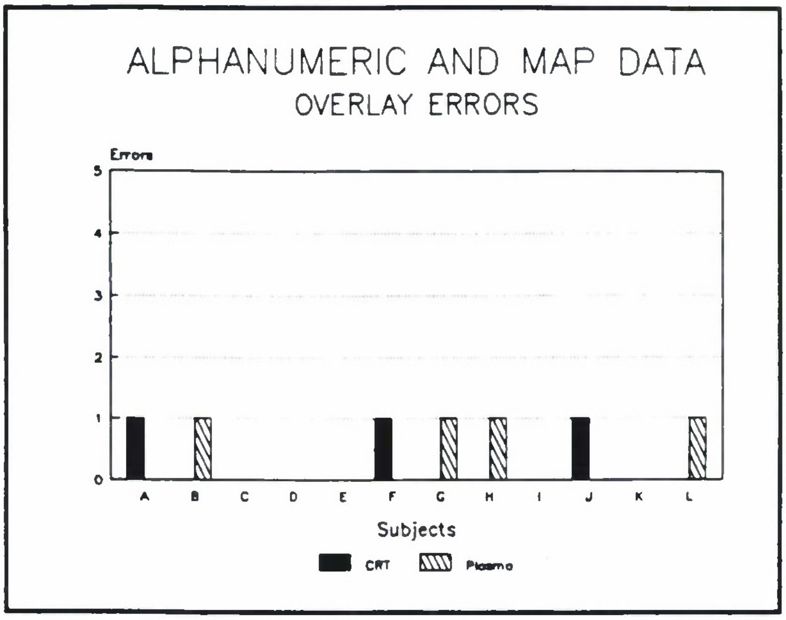


Figure A-6. Discrimination Method 2 - Errors by Subject

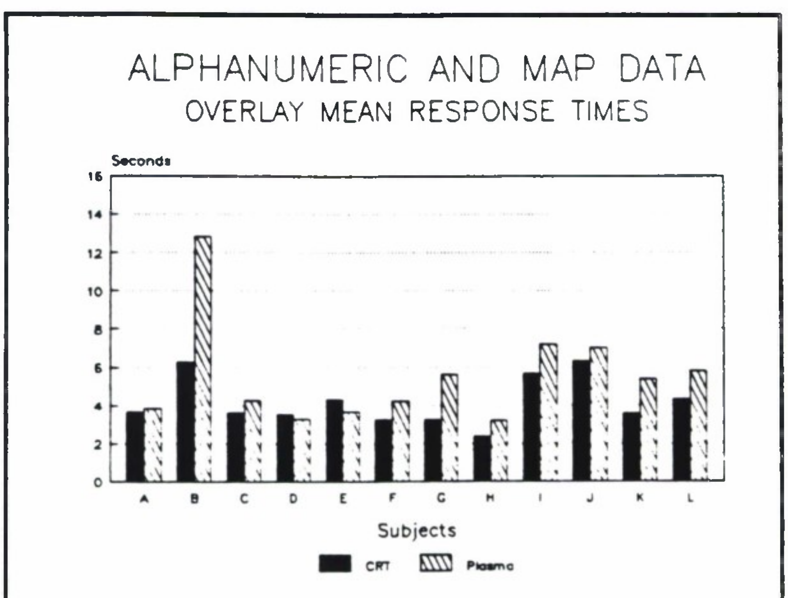


Figure A-7. Discrimination Method 2 - Mean Response Times by Subject

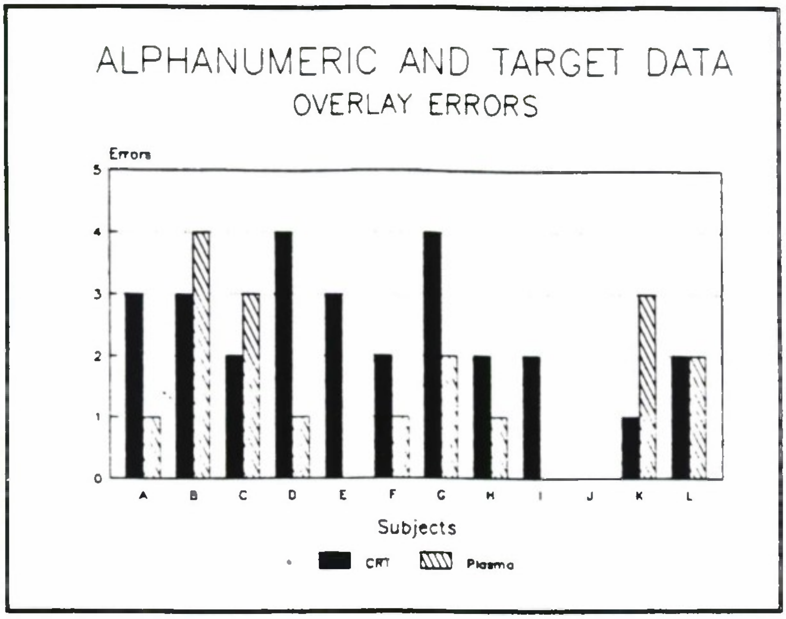


Figure A-8. Discrimination Method 3 - Errors by Subject

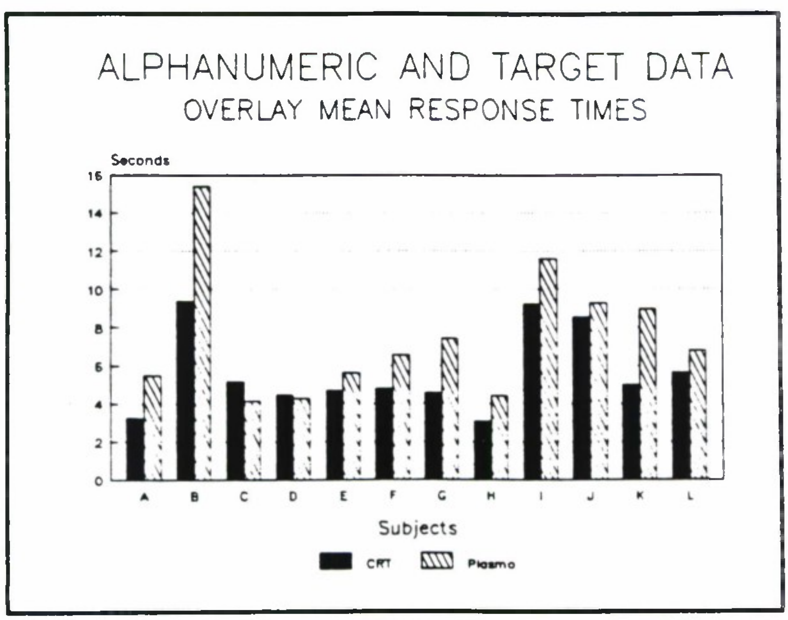


Figure A-9. Discrimination Method 3 - Mean Response Times by Subject

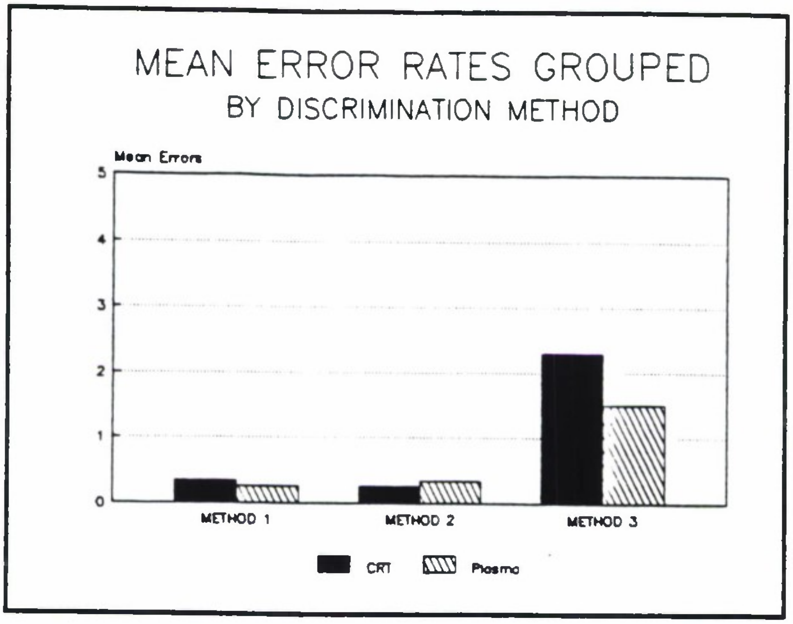


Figure A-10. Mean Error Rates Grouped by Discrimination Method

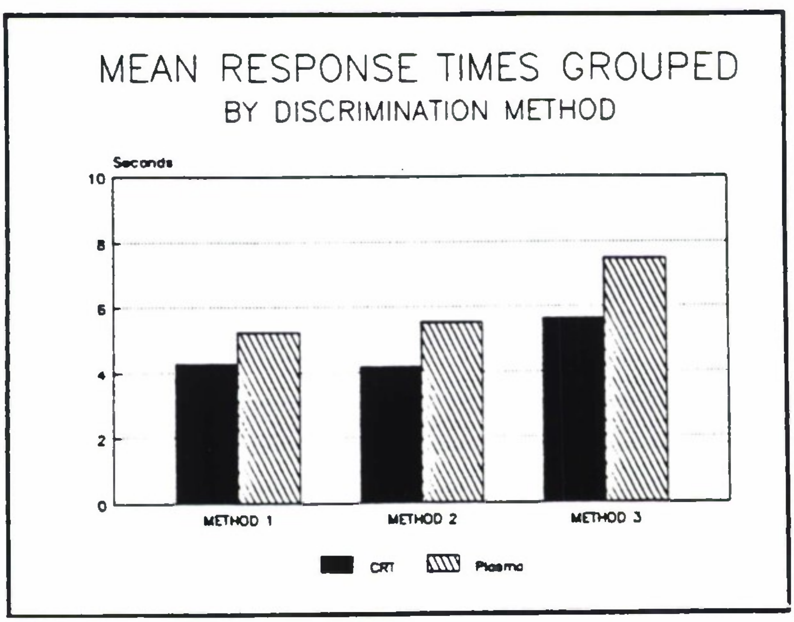


Figure A-11. Mean Response Times Grouped by Discrimination Method

As illustrated in table A-1, the ANOVA of the error scores yielded F values of 2.10, 0.013, and 3.09 respectively for methods 1, 2, and 3 above. These F values are not significant at the 0.05 probability level.

Furthermore, an ANOVA of the response time scores yielded F values of 3.92, 7.17, and 13.18 respectively for methods 1, 2, and 3 above. The F value of 3.92 is not significant at the 0.05 probability level. However, F values 7.17 and 13.18 are significant at the 0.05 probability level. In addition, 13.18 is significant at the 0.01 probability level.

A.4 DISCUSSION

The study showed that subject performance (errors) on the plasma display is equal to subject performance on the green CRT display. However, subject response times on the green CRT are significantly less than those on the plasma under the following conditions:

1. alphanumeric data at one intensity (60 percent) overlays map data at another intensity (20 percent) on the CRT compared with single intensity alphanumeric data displayed within a sterile area on the plasma (method 2); and

2. alphanumeric data at one intensity (60 percent) overlays target data at another intensity (100 percent) on the CRT compared with alphanumeric and target overlays at the same intensity on the plasma (method 3).

The operator performance with the green CRT is better (faster response time with equal error rate) than with the plasma display. Whether this performance difference is the result of learning or a more favorable color is subject to further investigation.

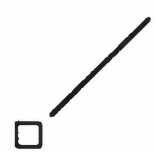
APPENDIX B

DISPLAY FORMAT TEST

Instructions:

Please identify the appropriate flight identification elements of each data block.

FULL DATA BLOCK



LIMITED DATA BLOCK

APPENDIX C

ANOVA CALCULATIONS

The following tables summarize the ANOVAs performed in the experiment. They show whether there were statistically significant effects of the independent variables upon subject error rates and response times.

ANOVAs calculate the probability that the differences between measured conditions are due to chance. If the F values in these tables were equal to or greater than 4.84, the differences between measures were considered to be statistically significant at the .05 probability level.

C.1 AGGREGATE ERROR RATES

Source	df	SS	MS	F
Conditions (c)	1	26.041	26.041	9.724
Subject(s)	11	13.125	1.193	
C x S	11	29.459	2.678	
Total	23	68.625		

C.2 AGGREGATE RESPONSE TIMES

Source	df	SS	MS	F
Conditions (c)	1	2.426	2.426	0.789
Subject(s)	1	109.877	9.988	
C x S	11	33.822	3.074	
Total	23	146.125		

C.3 PLASMA AND ORANGE CRT SI NON-OVERLAY

Source	df	SS	MS	F
Conditions (c)	1	0.0423	0.0423	0.320
Subject(s)	11	7.459	0.678	
C x S	11	1.457	0.132	
Total	23	8.959		

C.4 PLASMA SA OVERLAY VERSUS ORANGE CRT MI OVERLAY ERRORS

Source	df	SS	MS	F
Conditions (c)	1	1.042	1.042	2.10
Subject(s)	11	3.459	0.314	
C x S	11	5.458	0.496	
Total	23	9.959		

C.5 PLASMA SI OVERLAY VERSUS ORANGE CRT MI OVERLAY ERRORS

Source	df	SS	MS	F
Conditions (c)	1	9.375	9.375	4.882
Subject(s)	11	8.125	0.738	
C x S	11	21.125	1.920	
Total	23	38.625		

C.6 PLASMA AND ORANGE CRT SI NON-OVERLAY RESPONSE TIMES

Source	df	SS	MS	F
Conditions (c)	1	12.808	12.808	4.139
Subject(s)	11	82.472	7.497	
C x S	11	34.038	3.094	
Total	23	129.318		

C.7 PLASMA SA OVERLAY VERSUS ORANGE CRT MI OVERLAY RESPONSE TIMES

Source	df	SS	MS	F
Conditions (c)	1	1.203	1.203	0.426
Subject(s)	11	97.317	8.847	
C x S	11	31.051	2.822	
Total	23	129.571		

C.8 PLASMA SI OVERLAY VERSUS ORANGE CRT MI OVERLAY RESPONSE TIMES

Source	df	SS	MS	F
Conditions (c)	1	3.995	3.995	0.827
Subject(s)	11	182.506	16.591	
C x S	11	53.126	4.829	
Total	23	239.627		

GLOSSARY

ACRONYMS

AFR	air force regulation
ATC	air traffic control
ANOVA	analysis of variance
C	Condition
CRT	cathode ray tube
df	degrees of freedom
FDB	full data block
fL	foot Lamberts
G	green
Hz	hertz
ISO	International Standard Organization
LDB	limited data block
NMR	New Mobile RAPCON
MI	multiple intensity
MS	mean square
O	orange
OPS	operations subsystem
PIDP	programmable indicator data processor
PSR	primary surveillance radar
RAPCON	radar approach control
S	subject
SA	sterile area
SI	single intensity
SS	within subjects
SSR	secondary surveillance radar
TAC	Tactical Air Command
USAF	United States Air Force