AD-A056 919 MITRE CORP BEDFORD MASS F/G 15/3

TESTS RESULTS ADVANCED DEVELOPMENT MODELS OF BISS IDENTITY VERI--ETC(U)

JUL 78 A FEJFAR, P BENSON F19628-77-C-0001

UNCLASSIFIED MTR-3442-VOL-5 ESD-TR-78-150-VOL-5 NL

| OF |
ADA
056919

END
DATE
FILMED
9 78
DDC

AD A056919

LEVEL III

(18) ESD (19) TR-78-150-Vol-5

(11) SC

(14) MTR-3442-Vol-5

A056918

(6) TESTS RESULTS ADVANCED DEVELOPMENT MODELS OF BISS IDENTITY VERIFICATION EQUIPMENT. VOLUME V. MISCELLANEOUS.

(10) BY ADOLPH FEJFAR [redacted] PETER BENSON

(11) JUL 1978 (12) 47p.

Prepared for

DEPUTY FOR SURVEILLANCE AND NAVIGATION SYSTEMS
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
Hanscom Air Force Base, Massachusetts

(9) Technical rept.,

ELECTRONIC SYSTEMS DIVISION

Project No. 4130
Prepared by
THE MITRE CORPORATION
Bedford, Massachusetts
Contract No. (15) F19628-77-C-0001

Approved for public release; distribution unlimited.

235 050

78 07 27 029

REVIEW AND APPROVAL

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

THOMAS F. BASHARA
Project Manager

PAUL E. PEKO, Lt Col, USAF
Chief, Engineering Division
BIS Systems Program Office

ROGER T. KOZUMA, Col, USAF
Systems Program Director
BIS Systems Program Office
Deputy for Surveillance and Navigation Systems

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ESD-TR-78-150, Vol. V	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Test Results Advanced Development Models of BISS Identity Verification Equipment, Volume V, Miscellaneous		5. TYPE OF REPORT & PERIOD COVERED 6. PERFORMING ORG. REPORT NUMBER MTR-3442, Vol. V
7. AUTHOR(s) Adolph Fejfar and Peter Benson		8. CONTRACT OR GRANT NUMBER(s) F19628-77-C-0001
9. PERFORMING ORGANIZATION NAME AND ADDRESS The MITRE Corporation P.O. Box 208 Bedford, MA 01730		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Project No. 4130
11. CONTROLLING OFFICE NAME AND ADDRESS Deputy for Surveillance and Navigation Systems Electronic Systems Division, AFSC Hanscom Air Force Base, MA 01731		12. REPORT DATE JULY 1978 13. NUMBER OF PAGES 45
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

ACCESS CONTROL
ENTRY CONTROL
HUMAN FACTORS
INDEPENDENCE OF TECHNIQUES
TESTING
USER ACCEPTABILITY

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This volume presents general miscellaneous but related subjects concerning the results of testing the advanced development models of three identity verification (ASV, AHV and AFV) systems. The tests were conducted both at the Verification Laboratory at The MITRE Corporation and in the field at the Weapons Storage Area at Pease Air Force Base, New Hampshire. Volume I presents an executive summary

(over)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE 78 07 27 029

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

235 050

20. Abstract (continued)

and Volumes II, III and IV present the detailed test results of the speaker, handwriting, and fingerprint systems respectively.

ACKNOWLEDGMENT

This report has been prepared by The MITRE Corporation under Project No. 4130. The contract is sponsored by the Electronic Systems Division, Air Force Systems Command, Hanscom Air Force Base, Massachusetts.

TABLE OF CONTENTS

TABLE OF CONTENTS - concluded

LIST OF ILLUSTRATIONS

LIST OF TABLES

1.0 INTRODUCTION

As part of its effort to develop external physical security systems for the Department of Defense, the Electronic Systems Division (ESD) of the United States Air Force, under its Base and Installation Security System (BISS) program, acquired advanced development models of three automated identity verification systems for use in entry control.

The identity verification systems are designed to provide verification of the claimed identity of persons entering or leaving a restricted area through an Entry Control Point. The systems accomplish this verification by examining unique and measurable properties of a person's speech, handwriting, and fingerprint. The three systems that were evaluated include the Automatic Speaker Verification (ASV) system, the Automatic Handwriting Verification (AHV) system and the Automatic Fingerprint Verification (AFV) system. The MITRE Corporation was given the responsibility for installing, testing, and evaluating the three systems/techniques both at MITRE and at a field test site.

Volume I of this test report is an executive summary for the entire test program. This volume, Volume V, will address all of the remaining objectives detailed in the Test Plans that were not covered in Volumes II, III, and IV for the ASV, AHV, and AFV systems, respectively.

The remaining objectives of the Field Test Plan not previously addressed in Volumes II, III and IV are not specifically identified herein. However, because they were identical to the objectives in previous documents they are adequately covered.

2.0 DETENTION MODULE

A detention module, as it might be configured for the BISS Entry Control Subsystem, is an enclosed area (room or booth) which would be used as a deterrent to those attempting to seek unauthorized entry by detaining individuals who fail to verify their claimed identity. Use of a detention module increases the time a user must spend to complete a verification cycle. This increase in time becomes a factor in user acceptance and system cost. Increasing the verification cycle time irritates personnel and forces the use of more entry stations to maintain a given throughput. At an entry point with a detention booth, a complete verification cycle includes:

1. Door opening, entry into the booth, and door closing,
2. Claiming of an identity, i.e., entry of an ID number,
3. Verifying the individual,
4. Door opening, exit from the booth, and door closing, and
5. Dead time, i.e., unassignable causes.

Steps 2 and 3 were measured throughout the test program, while Steps 1, 4 and 5 were measured during the Field Test. Only the verification time (Step 3) changes with the technique used, and thus is used as a factor in judging the relative merit of the three techniques.

2.1 DESCRIPTION AND OPERATION

In the Phase I and Phase II tests, a detention module was not used. However, in the Field Test a detention module consisting of a three-foot wide by four-foot long chain link fence extending from floor to ceiling was placed around the ASV terminal as shown in Figure 1. The detention module has two doors that were used for either entering or exiting. To gain access to the ASV terminal, a test participant pressed a button located on the door post. If the detention module was in use, or a previous request had been made from the other side, a message "PLEASE WAIT - IN USE" was displayed. Otherwise, a buzzer sounded when the door unlocked and the individual opened the door, entered, and proceeded with speaker verification. During verification, both doors were locked and no admittance was allowed; at the conclusion of the verification, the door opposite the one used to enter the booth automatically unlocked allowing the test participant to leave the detention module. For the Field Test, the detention module was not designed to detain an individual who did not verify.

In addition to the operations just described, the detention module was equipped with a time interval counter which measured the time from the sounding of the opening buzzer to the closing of the exit door. The total time from when the "PUSH TO ENTER" button was activated until the entrant verified, exited, and the exit door closed, was a complete verification cycle and was measured automatically. The timing circuitry could be set to measure a complete cycle for traffic flow in either one direction or the other but not both. When a test participant entered the detention module in the opposite direction from which the timing circuitry was set, the time measured was from when the "PUSH TO ENTER" button was pushed until the entrance door closed. This feature of the timing circuit measured the time required to enter the control element only.

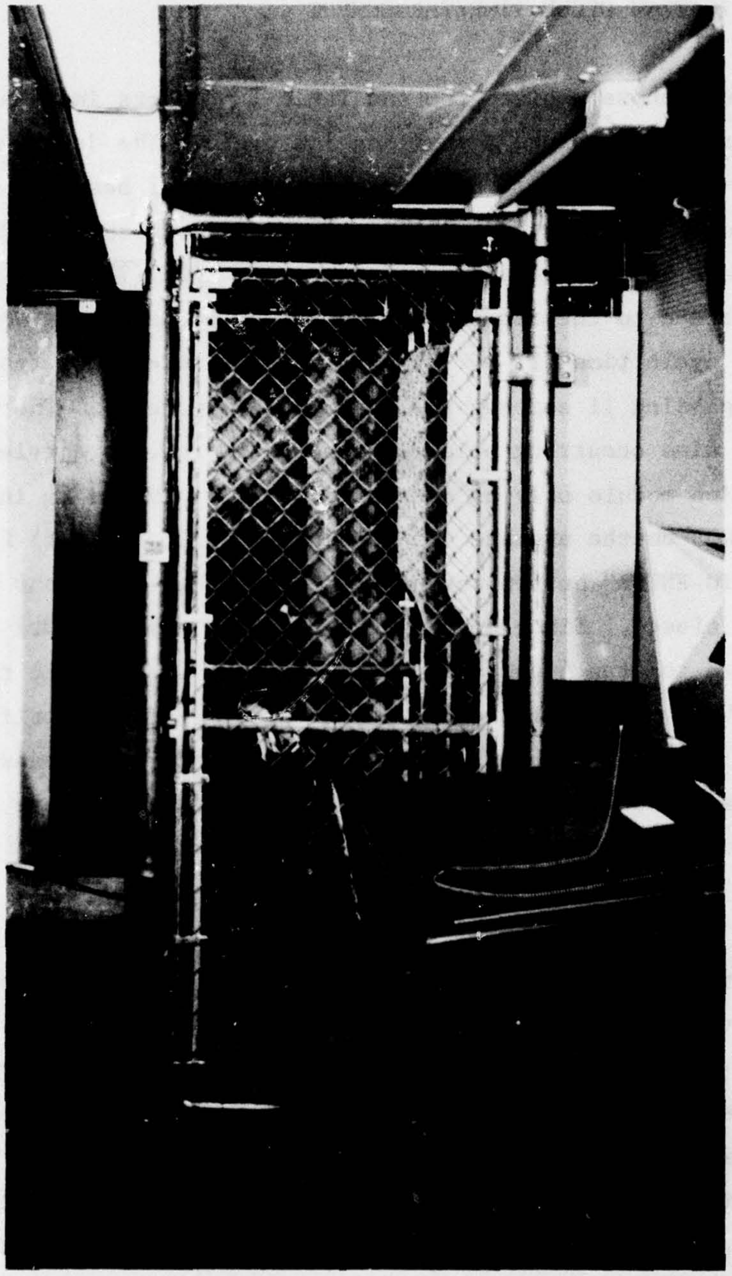

Figure 1. Detention Module

2.2 VERIFICATION CYCLE TIME/THROUGHPUT

Figure 2 shows a histogram and Table I presents in detail the time required to complete a verification cycle. The data was taken on Monday, 31 January 1977 and represents a normal busy day (224 verifications) and a day when no impostor data was collected. Had impostor data been collected, the mean cycle time $(\overline{T})$ would have been longer due to the requirement for eight phrases (see Volume II). An average cycle time $(\overline{T})$ of 18.5 seconds was determined from the data encompassing 11 seconds to 34 seconds (Figure 2). The verification cycle time occurring below 11 seconds represents entries into the detention module only as discussed in 2.1. This data indicates that it takes on the average 4.3 seconds (step 1, Table I) from when the "PUSH TO ENTER" button was pressed until the entry door was opened and closed. Since exiting the booth is simpler, 3.7 seconds was assigned to step 4 and the times for steps 2 and 3 are those reported in Volumes II, III, and IV. The reason for the verification cycle time going as high as 34 seconds (Figure 2) was because each phrase takes about four seconds to complete, and there were a few test participants who required several (a maximum of eight) phrases to verify.

The detention module could have been used with either of the other two verification systems (AHV or AFV); however, this was not required. The data obtained using the booth with the ASV system can also be used to determine the verification cycle time for the AHV and AFV systems. Analysis for each of the verification systems resulted in the estimated time required for a complete verification cycle. The results as presented in Table I show system throughput, or the number of entries which could be serviced per minute, to be 3.24, 2.33, and 2.83 for the ASV, AHV, and AFV systems, respectively.

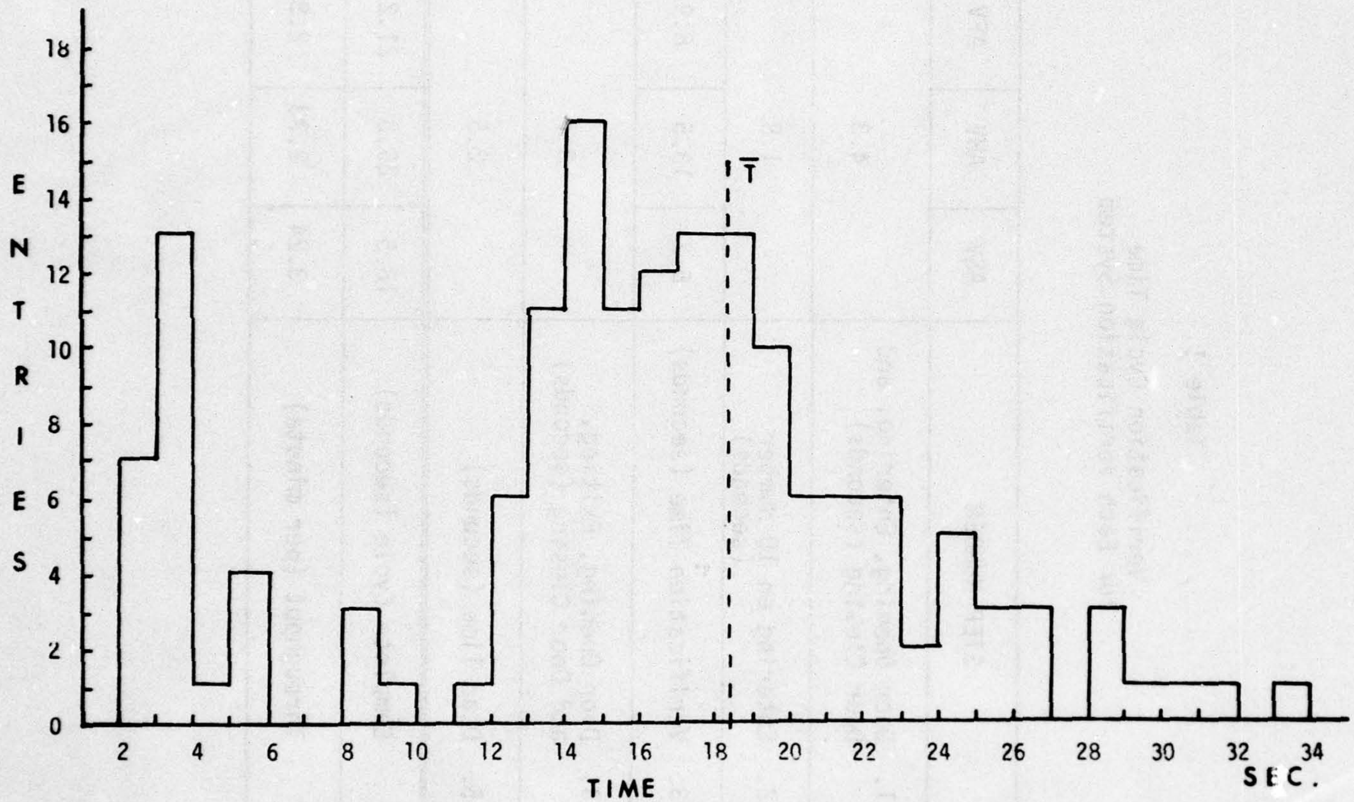

Figure 2. Verification Cycle Time vs. Number of Entries

Table I

Verification Cycle Time
for Each Verification System

STEP NUMBER	ASV	AHV	AFV
1. Door Opening, Entering, and Door Closing (seconds)		4.3	
2. Entering an ID Number (seconds)		1.8	
3. Verification Time (seconds)	6.2	13.5	8.9
4. Door Opening, Exiting, and Door Closing (seconds)		3.7	
5. Dead Time (seconds)		2.5	
Complete Cycle (seconds)	18.5	25.8	21.2
Throughput (per minute)	3.24	2.33	2.83

The BISS throughput requirement for a pedestrian entry module is an average of four (4) per minute. This throughput requirement includes both entry into, and exit from, a controlled area. Entry into a controlled area requires the user to enter an identification number plus verification, while exit from a controlled area requires the entry of an identification number only. The throughput data in Table I is for entry only.

If one considers only the verification time and allows an entry for every exit then the throughput requirement will be met only if the verification time (step 3) can be lowered to about 5.4 seconds. However, it is felt that the dead time (step 5) as measured in the Field Test is somewhat a function of the field environment, i.e., confining quarters, etc., and that the dead time would be reduced by half in an operational environment. With the above considerations, only the ASV system meets the BISS throughput requirements.

3.0 OBJECTIVES

There were several general objectives of the test program that were not covered in Volumes II, III or IV. These are discussed below.

3.1 OBJECTIVES 9, 18, AND 31 -- ANALYSIS OF THE EFFECT OF ENROLLMENT PROCEDURE ON SUCCESS OF VERIFICATION

Statement of Objective

To analyze the enrollment procedure and Post Enrollment performance, and to determine if the procedure can be optimized.

3.1.1 Enrollment

The enrollment process for Phase I and Phase II was identical. Upon arriving at the Verification Laboratory, the test participants were shown, on a portable theater, a 20-minute slide show with taped narration (audiovisual briefing) that described the test program and the enrollment and verification procedures. After the briefing, the test participants provided, via an enrollment card, personal data required for the verification process and data analysis. The test participants then proceeded to the ASV terminal to begin the enrollment sequence on the three Identity Verification Systems as described in Volumes II, III and IV. The enrollments were done by appointment in groups of three or four in a relatively comfortable and noise free environment.

The Field Test enrollment was similar to Phase I and II but the conditions were not as favorable. The audiovisual briefing was shown in a room which was part of the guard house. Viewing and listening to the briefing was not optimum because of telephone distractions and security police moving about performing their daily tasks. To proceed from the audiovisual briefing to the test van, the enrollees had to go outside in sub-freezing weather and pass through their own security system (badge exchange). The actual enrollment of the test participants onto the ASV, AHV and AFV systems was very similar to the procedures in Phases I and II. The main difference being that in the Field Test the test van was more noisy and confining than the Entry Control Laboratory at MITRE.

The Type I error rate in Post Enrollment for both the ASV and AHV systems was much higher in the Field Test than in Phases I and II. Most of this change was probably due to the change in the user population. Some of the change may have been due to the change in the enrollment conditions.

3.1.2 User's Recommendations/Comments

As part of the questionnaire handed out to the Field Test participants, the question was asked, "Did you like the preenrollment audiovisual briefing?" Of the 67 people responding, 57 answered yes, 9 answered no, and 1 could not remember.

There were several recommendations made by the users as to how the enrollment process might be improved. Several suggested that the enrollees be shown or demonstrated the verification equipment prior to the briefing and that the briefing be shown in close proximity to the equipment. Several users suggested that since the operator instructs the user on the enrollment process for each verification system that the briefing could be done away with entirely. Others suggested showing movies on the use of the verification systems instead of slides and using a better means of film projection.

3.1.3 Summary

Experience and user reaction indicate that a preenrollment briefing on the use of the verification equipment should be given to all enrollees. It appears, however, that it would be helpful to show and perhaps demonstrate the equipment to the enrollees prior to or just after the audiovisual briefing, but before enrollment.

Concerning the three verification systems, the enrollment on the AHV and AFV systems went well. There were some difficulties, however, with the ASV system. As is discussed in 3.4.2.1, many of the users had difficulty understanding the phrases prompted. This was especially true for enrollment. It is suggested, therefore, that just prior to an individual enrolling on the ASV system, that the individual be given a chance to listen and respond to the ASV prompting.

A disadvantage in the Field Test was that the new enrollees were briefed on three new systems in one session. This undoubtedly reduced the effectiveness of the briefing for any one of the systems. In an operational environment where only one technique would be used, the briefing would be shorter and would concentrate on only that one technique. This change plus a hands-on preenrollment familiarization should improve the enrollment session.

3.2 OBJECTIVES 19 AND 32 - INDEPENDENCE OF ERRORS AMONG THE SYSTEMS

Statement of Objective

To determine the independence of the Type I and Type II errors occurring during simultaneous ASV, AHV, and AFV verification attempts.

3.2.1 Purpose of Test

There are applications for automated entry control systems in which the error rates achievable with a single system will not be low enough. In these cases, a hybrid system combining two or more of the techniques could be implemented. To predict the hybrid performance, the degree of independence between the techniques must be known. To determine the degree of independence, an individual was requested to use all of the systems during each verification session.

3.2.2 Independence of Type I Errors

For each session in the laboratory that a user attempted verification on two or more systems, data for testing Type I independence was collected. Type I errors which occurred on two or all three systems during the same session were then collated.

Table II shows the Type I data observed for all combinations available during each phase of testing. The 90% confidence interval is based on the Chi-squared test using the observed errors. The upper and lower limits were rounded to obtain the integer values shown.

Table II

Type I Independence Data

	OBSERVED			90% CONF. INTERVAL ON ERRORS *		PREDICTED	
COMBINATION	ERRORS	ATTEMPTS	ERROR RATE (%)	LOWER	UPPER	ERRORS	ERROR RATE (%)
a. Phase I							
ASV & AHV	7	2645	0.265	4	13	5	0.170
b. Phase II							
ASV & AHV	0	603	0	0	3	0	0.015
ASV & AFV	1	5091	0.027	0	5	1	0.017
AHV & AFV	0	603	0	0	3	1	0.245
ASV & AHV & AFV	0	603	0	0	3	0	0.001
c. Field							
ASV & AHV	5	7129	0.070	3	11	3	0.043
ASV & AFV	4	7126	0.056	2	9	9	0.129
AHV & AFV	12	7071	0.170	8	19	10	0.142
ASV & AHV & AFV	0	7018	0	0	3	0	0.003

*Using the Chi-squared distribution with $\nu = 2$ (errors +1)

Assuming independence between two systems A and B, the joint error rate for A and B is the probability of an error on system A, p(A), times the probability of an error on system B, p(B), i.e., p(AB) = p(A) p(B). Thus, the predicted error rate is obtained by multiplying the Type I error rates (obtained from Volumes II, III and IV) for the systems involved. This combined rate is then multiplied by the observed attempts to obtain the number of predicted errors. The error rates used for each system included all users, periods, and errors (including assignable errors).

3.2.2.1 Phase I Results. The AFV system was not available in Phase I, so only the ASV-AHV combination was possible (Table II.a). Three of the seven errors observed were due to one user who should have been reenrolled on both systems. When all pre-reenrollment errors and attempts were removed from the data base, there were four errors with a confidence interval of 2 to 9 errors in 2502 trials as compared to a predicted 2 errors. In both cases, the predicted errors fell within the 90% confidence interval. Thus, the assumption of independence was demonstrated.

3.2.2.2 Phase II Results. Table II.b shows the results for the four possible combinations when three systems are available. The AFV data is the combination of the two test periods discussed in Volume IV. None of the combinations in Phase II had enough attempts to predict more than one error because the ASV system had a low error rate (0.64%), and the AHV system was available for only two weeks. Since all of the predicted values fall within the confidence interval, there was no observed violation of the assumption of independence.

3.2.2.3 Field Results. There was sufficient data collected in the Field Test (Table II.c) for evaluating the two-way combinations. At least five times as much data would be required to predict one error for the three-way combination. Once again, the predicted errors fall within the 90% confidence interval of the observed data.

3.2.3 Independence of Type II Errors

Table III presents the results of the Type II data analysis for independence between the techniques. Table III was derived in the same way as was Table II, except that playback data was used resulting in many more attempts. Whenever an impostor was available on the systems being considered, a check was made to see if that impostor passed against the same reference file on both systems. Since an impostor was rarely an impostor on two systems during the same visit to the laboratory, this was done in chronological sequence. That is, when impostor N was found in system A, the system B data was scanned until impostor N was found, and then the reference file ID's were checked to see if there was a common Type II error.

3.2.3.1 Phase I Results. Since the AFV system was not available during Phase I, only the ASV-AHV combination was possible. The predicted errors fell within the 90% confidence interval, and the results in Table III.a indicate independence between the two techniques.

3.2.3.2 Phase II Results. All three systems were available in Phase II, and all four configurations are shown in Table III.b. Because of the limited data available on the AHV system, sufficient data was not available to predict errors for the three-way combination. Otherwise, the predicted errors always fell within the 90% confidence interval, and the assumption of independence was supported.

Table III

Type II Independence Data

COMBINATION	OBSERVED ERRORS	OBSERVED ATTEMPTS	OBSERVED ERROR RATE (%)	90% CONF. INTERVAL ON ERRORS * LOWER	90% CONF. INTERVAL ON ERRORS * UPPER	PREDICTED ERRORS	PREDICTED ERROR RATE (%)
a. Phase I							
ASV & AHV	13	79,058	0.016	8	21	16	0.021
b. Phase II							
ASV & AHV	2	7,597	0.026	1	6	5	0.069
ASV & AFV	45	51,971	0.087	35	58	48	0.092
AHV & AFV	2	11,266	0.018	1	6	3	0.030
ASV & AHV & AFV	0	7,597	0	0	3	0	0.001
c. Field							
ASV & AHV	54	57,145	0.094	43	68	55	0.097
ASV & AFV	16	51,105	0.031	11	24	20	0.040
AHV & AFV	77	57,982	0.133	64	93	71	0.123
ASV & AHV & AFV	1	51,105	0.002	0	5	1	0.002

*Using the Chi-squared distribution with $\nu = 2$ (errors + 1)

3.2.3.3 Field Results. Sufficient data was available in the Field Test (Table III.c) for the two-way combinations, but was marginal for the three-way combination. The assumptions of independence were supported in that the predicted errors always fell within their respective confidence intervals. This was in spite of the fact that the ASV and AHV systems had unusual Type II error rates. The ASV system had an unexpectedly high Type II error rate in the Normal mode (see Volume II), and the AHV system had a high Type II error rate due, in large part, to the saturation in the surface transducer (see Volume III).

3.2.3.4 Summary. Both the Type I and Type II analysis supported the fact that the three systems are independent. Although there was not enough data in some test periods, all of the results indicate independence of the techniques. In addition, since each of the paired systems are independent, the three systems are also independent. Thus, combining the techniques appears to be a practical way to reduce Type I and Type II error rates.

An interesting occurrence of the Type I and Type II data in the Field Test (Tables II.c and III.c) is that the observed error rate on the AHV and AFV combination is larger than that observed for either the ASV and AHV and the ASV and AFV combinations. If further side-by-side testing of these three techniques is ever conducted, the trend should be looked for again to answer the question, "Is speaker verification more independent of handwriting and fingerprint than handwriting and fingerprint are of each other?"

3.3 OBJECTIVES 20 AND 33 - ANALYSIS OF HYBRID SYSTEM ERROR RATES

Statement of Objective

To determine the Type I and Type II Error rates for the ASV, the AHV and the AFV techniques operating together as a hybrid system.

3.3.1 Hybrid Combinations

Many different combinations of hybrid systems are possible with three systems. These are:

ASV and AHV	ASV or AHV
ASV and AFV	ASV or AFV
AHV and AFV	AHV or AFV
ASV and AHV and AFV	ASV or AHV or AFV
(ASV and AHV) or AFV	(ASV or AHV) and AFV
(ASV and AFV) or AHV	(ASV or AFV) and AHV
(AHV and AFV) or ASV	(AHV or AFV) and ASV

Each would have to be analyzed in terms of the thresholds, i.e., error rates for each component, to achieve the desired hybrid error rate and the resulting throughput.

3.3.1.1 Example of an AND System. One example of an AND system is presented to illustrate the approach required. For this combination, the user must verify on both techniques before gaining entry.

3.3.1.1.1 Hypothesized AND System. From section 3.2, it was shown that all of the techniques are independent, and the performance of an AND hybrid system can be hypothesized. Table IV shows the error rates for males in the Normal mode of operation. As an example, the ASV and AHV systems will be combined, and the performance will be computed using the Field Test results. Because the user must be verified on both systems to gain entry, the Type I error rates add and the Type II error rates multiply. The results are a 3.0% Type I error rate and a 0.18% Type II error rate. Eighteen Type II attempts in 10,000 are successful but 300 authorized people are rejected. As is obvious, the thresholds used for the stand-alone systems are not the ones that would be used for a hybrid system.

Table IV

ERROR RATES FOR ALL THREE SYSTEMS IN THE NORMAL MODE

	ASV Performance		AHV Performance		AFV Performance	
	Type I	Type II	Type I	Type II	Type I	Type II
Phase I (%)	0.9	1.0	6.3	3.0	-	-
Phase II (%)	0.2	4.4	3.2	1.7	4.6	2.2
Field Test (%)	1.1	3.3	1.9	5.6	6.5	2.3
Verification Time (sec)	6.2	-	13.5	-	8.9	-

3.3.1.1.2 <u>Optimum Thresholds</u>. By changing any threshold on the ASV or AHV systems, the Type I and Type II error rates of the systems will be changed and thus the error rates of the hybrid system will be changed. To determine how a hybrid system's error rates vary with any threshold or combination of thresholds requires that the threshold be varied and the Type I and Type II rates be determined for each system. The hybrid system error rates are then determined by a single addition and a single multiplication. An estimate of the new thresholds and error rates can be done by using the sensitivity analysis sections in Volumes II, III, and IV. Better estimates can then be determined by running a special program which recomputes Type I and Type II errors from the digitized data obtained during the test. Such a special program was written for the ASV system but none were written for the AHV and AFV systems.

Assume that a 1% Type I error rate is required for an ASV-AHV hybrid system. Reducing the ASV Type I error rate to 0.5% may result in a Type II error rate of 5%. Similarly, for the AHV system, a Type I error rate of 0.5% may yield a Type II error rate of 10%. The resulting hybrid Type I error rate is 1% (0.5 + 0.5) and the Type II error rate is 0.5% (5% . 10%).

3.3.1.1.3 <u>Cycle Time</u>. The cycle time of the AND system will include keyboard time, door operation and dead time, plus time to verify on each system. If the keyboard digits from the ASV system were passed over to the AHV system, and the person were in Normal verification, then the average time would be:

ASV keyboard (from Table I)	1.8 seconds
Door operation and Dead Time (from Table I)	10.5 seconds
ASV normal verification	6.2 seconds
AHV verification	13.5 seconds
Cycle Time	32.0 seconds

Obviously an AND system has lower throughput than any single system.

3.3.1.2 Example of an OR System. A similar example of an OR system will be presented. In this case, the user need only to verify on either of the two techniques to gain entry.

3.3.1.2.1 Hypothesized OR System. If a person passes on one part of an OR system, say ASV, he is allowed to enter, while if he fails, he must try the other part of the system -- in this case, AHV. Passing either ASV or AHV allows entry.

The Type I error rate in an OR system is the product of the Type I error rate for each part of the system and the Type II error rate is the sum of the Type II error rates. Using the Field Test data from Table IV, Type I error rate for an OR system is 0.02% (1.1% times 1.9%) while the Type II error rate is 3.3% plus 5.6%, or 8.9%.

3.3.1.2.2 Optimum Thresholds. In the OR system, the thresholds of the individual systems should be changed to reduce Type II error rates at the price of somewhat higher Type I error rates. The OR system error rates versus threshold can be determined, as in the AND system, from the error rate versus threshold of the individual verification systems.

Again, assume that a hybrid Type I error rate of 1% is desired. The ASV threshold may be changed to obtain a Type I error rate of 10% with a corresponding Type II error rate of 0.01%. Likewise, for AHV, a Type I error rate of 10% may bring a corresponding Type II error rate of 0.05%. For the hybrid system, the Type I error rate is 1% (10% · 10%) and the Type II error rate is 0.06% (0.01% + 0.05%).

3.3.1.2.3 Cycle Time. The cycle time of an OR system is not much worse than that of whichever system is used first. If the ASV is used first, then the time required to key in an identification will be the same as before and the time to verify will be only slightly more than with a stand-alone system due to the change in threshold causing higher Type I errors. Assuming that the ASV Type I error is 10%, then only 10% will need to try the AHV system. Thus, for the person in normal verification:

ASV keyboard time (from Table I)	1.8 seconds
Door Operation and Dead Time (from Table I)	10.5 seconds
ASV normal verification	6.2 seconds
(AHV verification time) · (10%)	1.4 seconds
Cycle Time	19.9 seconds

This verification cycle is only marginally slower than the ASV stand-alone cycle.

Even with free usage, i.e., not directing the user to the ASV system first, verification time is much shorter than the AND system. From the verification times shown in Table IV, ASV is twice as fast as AHV so assume that 67% of the verifications occur on ASV and the other 33% on AHV. In this case, verification time for each component of the hybrid system is:

$$\text{ASV: } [0.67 + (0.33)\ (0.1)]\ 6.2 = 4.4 \text{ seconds}$$

$$\text{AHV: } [0.33 + (0.67)\ (0.1)]\ 13.5 = 5.4 \text{ seconds}$$

For the total verification cycle, 22.2 seconds are now required. Thus, only 2.3 seconds are sacrificed for freedom of choice of which system to use.

3.3.1.3 <u>Summary</u>. A variety of hybrid systems are possible. For an AND combination, the Type I error rates and the verification times add while the Type II error rates multiply. For an OR combination, the Type I error rates multiply, the Type II error rates add, and verification times add as shown in the second example for the OR system.

The OR system would appear to have more utility than an AND system for four reasons. One, a user having difficulty on one technique can avoid it in favor of the other. Two, in the general Type I/Type II error rate tradeoff, a lower Type II error rate can be obtained for a given Type I error rate (4.5% Type I - 0.1% Type II) than vice versa (0.6% Type I, - 4.5% Type II)[1]. Three, it has the higher throughput, and four, it has better reliability than the AND system.

3.4 OBJECTIVE 21 AND 34 - INVESTIGATE THE HUMAN ENGINEERING ASPECTS OF THE EQUIPMENT

<u>Statement of Objective</u>

To investigate the human engineering aspects of the equipment.

3.4.1 <u>Field Test Survey</u>

For the Field Test only, a survey in the form of a questionnaire was used to elicit comments on comfort, difficulties, opinions and appeal with regard to the identity verification equipment and detention module. This survey was distributed to the operators and all test participants who used the equipment on 17 January 1977. Of the 74 questionnaires that were distributed there were 68 replies.

3.4.2 Findings of Questionnaire

3.4.2.1 Automatic Speaker Verification System. For the ASV system 18 (26%) of the 68 people replying answered "No" to the question, "Is the speaker terminal easy to use?"

The reasons for this negative feeling is perhaps found in the replies to other questions;

12 (18%) of the respondees thought the phrases were spoken too fast.

15 (22%) of the respondees had difficulty in remembering the four words to repeat.

25 (37%) of the respondees did not feel comfortable speaking into the microphone.

28 (41%) of the respondees had difficulty in understanding the prompted phrase.

Also, the ASV terminal was unique in that it was enclosed by the detention module and this may have added to an overall negative feeling for the ASV system. This is suggested by the response to the question, "Do you feel comfortable using the speaker booth?" whereby 23 (34%) of the respondees replied, "No."

Much of the negative response may also have been due to the experiences encountered by the users during their first few weeks of enrollment and verification. It took about four weeks for the users to become accustomed to using the ASV system. During this learning period, when Type I errors were highest, feelings were acquired that may not have completely disappeared by the time the questionnaire was distributed.

3.4.2.1.1 User's Recommendations/Comments. Of the 22 people responding to the question, "What recommendations do you have for improving the speaker terminal?" 14 (64%) of those responding commented about not being able to understand the prompting voice. Three people thought the prompting was too fast, two suggested the prompting voice should be changed from a male's to a female's and two felt that reducing external noise would help them to better understand the prompted words.

It should be mentioned that the noise level within the van was fairly high at times especially during the rush hour. There was general background noise due to the cooling fans within the various equipments and added to this was an occasional aircraft flying overhead, people talking and the teletypewriter printing messages.

Other suggestions for improving the ASV terminal included:

"Have the system say words more personal before the test, maybe 'Good morning'."

"Provide a digital readout for the ID number so that one knows when a proper ID has not been entered." (This was suggested for all three terminals.)

"Install more than one terminal so that during busy times people don't have to wait."

3.4.2.1.2 <u>Operator's Recommendations/Comments</u>. Three Air Force personnel were trained as system operators and their comments are discussed separately from the rest of the user comments. Section 4.0 discusses operator training.

Operator No. 1 - "I think that a set of well insulated headphones should be installed to give the person verifying a better understanding of the phrases he or she is required to repeat."

Operator No. 2 - "A more understanding voice should be used."

Operator No. 3 - "The phrases should be made clearer."

3.4.2.2 <u>Automatic Handwriting Verification System</u>. In replying to questions concerning the handwriting terminal, there was almost unanimous agreement that the terminal was easy to use and the instructions easy to read.

3.4.2.2.1 <u>User's Recommendations/Comments</u>. There were relatively few recommendations made by the users as to how the terminal might be improved. Of those that did offer suggestions, several suggested that the smooth MYLAR writing surface made it somewhat difficult in signing their name and that a different surface would be better. Several mentioned that the writing surface was not conveniently positioned for both right and left handed people, and suggested that several writing surfaces be made available to accommodate various writing positions.

3.4.2.2.2 Operator's Recommendations/Comments.

Operator No. 1 - "Add a tension device to the MYLAR paper advance mechanism so that the take-up spool is more neatly rolled with less slack and that the removal of the spool be made easier."

Operator No. 2 - "Have it at an easier writing angle."

Operator No. 3 - "The smooth writing surface took a while to get used to."

3.4.2.3 Automatic Fingerprint Verification System. For the fingerprint terminal, 11 (16%) of the 68 people replying answered, "No" to the question, "Is the fingerprint terminal easy to use?" All of the respondees found the displayed instructions easy to read and understand and only four found it difficult to remember which was their Primary and which was their Alternate finger. However, 14 people (21%) found the finger positioner, used to align their finger, uncomfortable to use with eight complaining that the position was at the wrong angle and two complaining that the positioner was the wrong size.

3.4.2.3.1 User's Recommendations/Comments. The users had few recommendations on how the AFV terminal might be improved; those that were submitted are presented below:

"Place buttons (key pad) in the middle so left/right hand people can reach the buttons easier with the hand not being finger printed."

"Make the finger positioner easier to use."

"Adjust contact terminal so that contact (with the finger) is made easier."

"Have it (the terminal) so you can use the whole hand instead of one finger."

"The angle at which you must place your hand is not comfortable. Perhaps a flat table or an adjustable table could be used."

3.4.2.3.2 Operator's Recommendations/Comments.

Operator No. 1 - "For the area in which the test took place, the system is not that effective. Participants in the test were constantly coming through with new cuts on their fingers because of the type of work they do. This resulted in a lot of failures. Also, the cold temperature accounts for more than half of the failures. To me, the system would be better used inside a building instead of where the temperature is sometimes 10° or lower."

(It should be noted that the correlation between failure and cold fingers is a subject for discussion and that the above comments are only the opinion of the operator.)

Operator No. 2 - "No comment - A well engineered terminal."

Operator No. 3 - "No comment."

3.4.2.4 <u>Detention Module</u>. As has been previously discussed in the ASV section, 23 (34%) of the respondees did not feel comfortable using the speaker booth.

3.4.2.4.1 <u>User's Recommendations/Comments</u>. There were 21 responses to the question, "What recommendations do you have for improving the speaker booth?" Of those 21, six wanted to do away with the booth completely, four suggested making the booth larger and two people did not want the gates locked.

There was one person out of all those that were enrolled in the ASV system who absolutely refused to enter the detention module. He said that it gave him claustrophobia. This person attempted in the beginning to enroll in the ASV system, but had a difficult time and required reenrollment. When the person was asked to reenroll the operator found out about his dislike for the detention module.

3.4.2.4.2 <u>Operator's Recommendations/Comments</u>.

Operator No. 1 - "I think the booth should be made higher. Anyone 6'3" or higher has a hard time positioning himself in the cage."

Operators No. 2 and 3 - No comment.

3.4.2.5 <u>General Comments about the Field Test</u>.

Of the 68 people responding to the questionnaire;

52 responded that they used the system every day,
11 responded that they used the system every other day,
1 responded that he used the system one day per week,

4 responded that they used the system less than one day per week, and

13 responded that they would have preferred not to participate and considered the test a "waste of time."

Question 3 asked the participants to identify the systems they would like or not like to use each time they entered or exited a secure area. The question and number of responses for each case are provided below:

Identify the systems you would like or not like to use each time you enter or exit a secure area.

		LIKE	NOT LIKE
Consider that a booth similar to the speech booth is used with all three systems.	SPEECH	32	32
	FINGERPRINT	40	24
	HANDWRITING	36	29

		LIKE	NOT LIKE
Consider that for all three systems, no booth is used.	SPEECH	32	23
	FINGERPRINT	49	12
	HANDWRITING	44	16

3.4.3 Summary

Considering either configuration, with or without the booth, the Fingerprint System was best liked with Handwriting second and Speaker third. The dislike for the detention module is evident in the totals; there were a total of 85 responses in the negative for systems using a detention module whereas there were only 51 responses in the negative when no booth was used.

When the Type I and Type II error rate performance is compared with these human factor results, it is obvious that they do not correlate. The familiarity and acceptability of fingerprints made it acceptable even though the AFV system forced users at that terminal to do the maximum six scans very often.

Being spoken to and speaking to a machine was foreign to all users at the beginning of the Field Test. In spite of the difficulty with the prompting voice, the ASV system had the best overall performance (similar reactions to the three systems were expressed by users in Phase I and Phase II). The users were "impostors" (i.e., extra data samples were required) every eighth time on all three systems. Yet, they only complained about the ASV system forcing the maximum tries.

The ASV system was the only one of the three that bystanders were easily aware of. Thus, when someone required several phases to verify, his buddies were aware of it and they would often give the user the "Bronx" cheer, thus embarrassing him. On the AHV and AFV system, the bystander was not generally aware of what the user at the terminal was doing; thus, embarrassment in front of peers or subordinates was minimized.

The detention module is intimidating and so is the ASV system. The mere presence of either or both of these units would appear to be an effective deterrent to many potential casual intruders.

3.5 OBJECTIVES 22 AND 35 - ASSESS THE NEED FOR CALIBRATION AND FIELD ADJUSTMENT

Statement of Objective

To assess the need for calibration and field adjustment.

3.5.1 Estimation of Equipment Reliability

The reliability of the AHV, ASV and AFV systems was analyzed based on equipment or component failures.

Throughout the period of the Field Test records were maintained on the downtime and repair of each identity verification system. Much of the downtime and maintenance was due to software glitches and other anomalies associated with data transmission protocol between the systems. These problems were aggravating but were usually fixed quickly with minimum loss in operating time.

For the Field Test, there were 43 days in which AF personnel used the equipment for verification. There were 11 1/2 hours of operation in each day, giving a total of approximately 495 operational hours. In addition, the equipment was not turned off except on weekends allowing an additional 537 hours of running time to accumulate.

Of the total 1,032 hours of running time, the failures that can be directly assigned to each system accounted for the following number of hours in downtime: ASV, 120 hours; AHV, 39 hours; AFV, 28 hours.

The 120 hours of downtime for the ASV system occurred during the start up process just after moving the equipment to Pease AFB. Not all of the 120 hours were used trying to bring the ASV system on line but several days were spent by MITRE personnel diagnosing the system. The problem was found to be particles of metal in the microphone electrical socket and a defective circuit board in the CPU. The particles were deposited in the electrical socket from the many adjustments made to the microphone to place it in a convenient position by the test participants in Phase I and Phase II. The microphone connector was cleaned and the circuit board replaced and the ASV system worked well throughout the remainder of the test.

For the AHV system there were two major failures, both on the user terminal. The first occurred during the first days of test when a PROM malfunctioned; the second occurred halfway through the test period when a voltage regulator required replacement. In each instance Veripen, the manufacturer, was on site within 12 hours to make the necessary repairs.

For the AFV system, there was one major failure occurring halfway through the test. The manufacturer, Calspan, was on site within 24 hours to make the necessary repairs.

3.5.2 Calibration Requirements

The ASV terminal did not need any periodic calibration. There were two activities required for the AHV terminal. One, the MYLAR material had to be replenished once a week, more or less, depending on usage, and two, the calibration of the pen (Phases I and II) or surface (Field Test) was checked once a week.

For the AFV terminal, Calspan checked the electronic, mechanical and optical alignments on a monthly basis. From their comments, this appeared to be adequate.

The other components of the systems, e.g., minicomputers, power supplies, disc drives, teletypes, and magnetic tape units, required periodic maintenance and cleaning. This was done either during other maintenance by contractor personnel or else on a scheduled basis by the site operators.

4.0 OPERATOR TRAINING

As part of the overall objectives for the Field Test, it was desirable to find out how effective Air Force personnel would be in operating the identity verification equipment and the training they would require.

For the Field Test, three Airmen 1st Class were provided for the duration of the Field Test by the 509th Security Police Squadron at Pease AFB, New Hampshire. Because of time constraints and the few operators to train, the training was limited to on the job training with no classroom activities. However, a training manual was provided that described the complete operation of the equipment. As it turned out, three weeks was ample time for the Air Force operators to learn to use and operate a small computer and associated peripherals such as a teletypewriter, teleprinter, and computer tape and disc drives. Their duties involved running various computer programs necessary for the operation and data collection process of the test program. In addition, they learned to keep daily logs and records and provide routine maintenance and adjustment to the equipment.

The operators were able to enroll over three hundred Air Force personnel on the three systems and to provide assistance to personnel using the equipment on a daily basis.

As a result of the excellent performance of the Air Force operators, it is felt that Air Force personnel would be very effective in operating identity verification equipment/systems similar to the equipment described in Volumes II, III and IV of this report.

5.0 THE EFFECT OF THE OPERATIONAL ENVIRONMENT ON THE TYPE I ERROR RATE

5.1 WEATHER EFFECTS

Throughout the test period, local weather conditions at the test site were recorded, i.e., temperature, wind speed, and precipitation. No direct day-to-day correlation could be found between the weather data and the Type I error rates of the three systems. However, there were some general observations. The test participants, because of the time of year, usually had cold hands when they operated the user terminals. As was discussed in Volume IV, this may have accounted for the generally higher AFV Type I error rate for the Field Test. Also, during periods of heavy precipitation, the AFV and AHV terminals required wiping to remove deposits of water. This was especially true of the AFV terminal as the surface of the finger scanner would become covered with water droplets. No correlation between water on the terminals and Type I errors was found; perhaps because the operators wiped the terminals often.

Finally, there were problems associated with the detention module due to water. As part of the control circuitry for the module, a switch mat was used on the floor as a sensor to indicate whether or not the module was occupied. This switch mat malfunctioned many times from water resulting from the melted snow carried in on the feet of test participants. Because of the weight of the water, the switch mat would activate, indicating that the module was occupied and, in addition, the excess water entered the mat and shorted the switch element. When the switch mat malfunctioned, the doors of the detention module were kept unlocked so that speech verification could continue. There was no indication that this condition affected the Type I error rate.

5.2 IDENTIFICATION NUMBERS

As was discussed in Volumes III and IV, the distribution of ID numbers for the 270 Field Test participants were closely distributed between 0003 and 0491. In Phase I and II, the ID number of the participants were more widely dispersed between 0052 and 9415. Intuitively, one can see that a Field Test participant, when making a mistake, had a greater chance of entering another ID than the Phase I or II participants. Table V presents data on the Type I error rates for the three systems due to wrong ID.

Table V

TYPE I ERROR RATES BECAUSE OF WRONG ID - FIELD TEST

System	No. of Verifications	Wrong ID	Type I Error Rates
ASV	7242	11	0.152%
AHV	7167	13	0.181%
AFV	7235	13	0.180%

The data was collected from the observations of the operators on duty who logged the information in the daily log book. Type I errors due to wrong ID were not observed in the Phase I or Phase II tests.

5.3 NOISE

As was indicated in 3.4.2.1.1 of this volume, several test participants complained that they had a difficult time understanding the prompted phrases of the ASV system because of the noise level within the van. The van was confining and at times noisy with the operation of the teleprinter, conversation of test participants waiting to use the system, the sound of equipment cooling fans and occasional jet aircraft noise. However, no Type I errors were observed that could be directly attributed to the noise level. After the test participants had visited the trailer a few times, they appeared to get accustomed to the noise level.

REFERENCES

1. Doddington, G.R. and B.M. Hydrick, Speaker Verification II, RADC-TR-75-274, Texas Instruments, P.O. Box 6015, Dallas, Texas, 75222, November 1975.