UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DTIC FILE COPY

AD-A196 378

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER AFIT/CI/NR 88- 50	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) EFFECTS OF NOISE AND TASK LOADING ON A COMMUNICATION TASK		5. TYPE OF REPORT & PERIOD COVERED MS THESIS
		6. PERFORMING ORG. REPORT NUMBER
AUTHOR(s) DEAN H. ORRELL, II		8. CONTRACT OR GRANT NUMBER(s)
PERFORMING ORGANIZATION NAME AND ADDRESS AFIT STUDENT AT: UNIVERSITY OF DAYTON		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE 1988
		13. NUMBER OF PAGES 99
MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) AFIT/NR Wright-Patterson AFB OH 45433-6583		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

DISTRIBUTED UNLIMITED: APPROVED FOR PUBLIC RELEASE

DTIC ELECTE AUG 0 3 1988 S D H

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

SAME AS REPORT

18. SUPPLEMENTARY NOTES

Approved for Public Release: IAW AFR 190-1
LYNN E. WOLAVER Lynn Wolaver 18 July 88
Dean for Research and Professional Development
Air Force Institute of Technology
Wright-Patterson AFB OH 45433-6583

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

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

ATTACHED

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT

EFFECTS OF NOISE AND TASK LOADING ON A COMMUNICATION TASK

Dean H. Orrell II, M.A.
University of Dayton, 1987.

Major Professor: David W. Biers, Ph.D.

Previous research had shown the effect of noise on a single communication task. This research has been criticized as not being representative of a real world situation since subjects allocated all of their attention to only one task. In the present study, the effect of adding a loading task to a standard noise-communication paradigm was investigated. Subjects performed both a communication task (Modified Rhyme Test; House et al. 1965) and a short term memory task (Sternberg, 1969) in simulated levels of aircraft noise (95, 105 and 115 dB overall sound pressure level (OASPL))). Task loading was varied with Sternberg's task by requiring subjects to memorize one, four, or six alphanumeric characters. Simulated aircraft noise was varied between levels of 95, 105 and 115 dB OASPL using a pink noise source. Results show that the addition of Sternberg's task had little effect on the intelligibility of the communication task while response time for the communication task increased.

Effects of Noise and Task Loading On A Communication Task

by

Dean H. Orrell II

Thesis

Submitted to

The Graduate School of Arts and Sciences

In Partial Fulfillment of the Requirements for

The Degree of

Master of Arts in Psychology

The University of Dayton

December, 1987

APPROVED BY:

David W. Biers, Ph.D.

Chairman, Thesis Committee

Donald J. Polzella, Ph.D.

Thesis Committee Member

Thomas J. Moore, Ph.D.

Thesis Committee Member

CONCURRENCE:

Kenneth J. Kuntz

Chairman, Department of Psychology

ACKNOWLEDGEMENTS

First, I would like to thank my previous thesis chairman Dr Robert Bateman who helped generate and direct my early efforts. Thanks also to Dr Dave Biers for becoming the new chairman and for providing invaluable advice during the data analysis and writing of the thesis. I am also grateful for the reviews and input provided by Dr Don Polzella, Dr Tom Moore and Dr Don Thomas.

I owe much to Tim Anderson who helped me initiate the research and who got support from key AAMRL/BBA personnel. Also, Dr Charles Nixon, Dr Tom Moore and Rich McKinley gave freely of their advice and time to support my research. Many thanks to Tom Grace, Dave Obenshire, Doug Sauer, Alisa Workman, Ken Peachy and other members of the Systems Research Laboratory (SRL) support team for their fine assistance throughout data collection.

I would also like to thank Dr Gary Reid for his help in using SWAT and Bill Acton for his help and input in selecting the memory search task. Jeanie Bressler and Scott Potter (SRL) provided valuable help and input during the analysis of the SWAT data. Lastly, my thanks to Ed Skuya, Royal Clark and Bob Linhardt who helped me in the long process of transferring data between computers.

TABLE OF CONTENTS

METHOD (Cont)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (Continued)

ABSTRACT

EFFECTS OF NOISE AND TASK LOADING ON A COMMUNICATION TASK

Dean H. Orrell II, M.A.
University of Dayton, 1987.

Major Professor: David W. Biers, Ph.D.

Previous research had shown the effect of noise on a single communication task. This research has been criticized as not being representative of a real world situation since subjects allocated all of their attention to only one task. In the present study, the effect of adding a loading task to a standard noise-communication paradigm was investigated. Subjects performed both a communication task (Modified Rhyme Test; House et al. 1965) and a short term memory task (Sternberg, 1969) in simulated levels of aircraft noise (95, 105 and 115 dB overall sound pressure level (OASPL))). Task loading was varied with Sternberg's task by requiring subjects to memorize one, four, or six alphanumeric characters. Simulated aircraft noise was varied between levels of 95, 105 and 115 dB OASPL using a pink noise source. Results show that the addition of Sternberg's task had little effect on the intelligibility of the communication task while response time for the communication task increased.

INTRODUCTION

A variety of factors influence a pilot's performance in the cockpit and, consequently, the effective accomplishment of the mission. Some of these include cockpit design, pilot skill level, mental attitude, illumination, noise levels and weather conditions. In previous research (Moore, McKinley and Mortimer, 1979; Moore, Nixon and McKinley, 1980; McKinley, Nixon, and Moore, 1981) personnel at the Armstrong Aerospace Medical Research Laboratory (AAMRL) have examined the effect of simulated aircraft noise on one's ability to communicate in various situations. In this research, subjects were typically required to perform a listening task in simulated aircraft noise. Performance or intelligibility of speech in a given system was measured by subjects response time and percent correct.

One criticism of this research is that listeners were allowed to devote all of their available time and attention to the listening task. This scenario is not representative of a real world cockpit situation where a pilot must allocate his attention to many different events that are occurring. In an actual aircraft environment, performance on a listening task is predicted to be worse due to the multiple demands on the pilot. It is hypothesized that performance on the listening task will change if subjects are required to perform additional tasks. In the present study, the effect of adding a second task to the listening task is investigated.

Communication

Communication is of vital importance in the cockpit. It is essential for pilots to be able to understand the messages sent to them. A major requirement of any aircraft voice communication system is the ability to deliver intelligible speech. In a general sense, intelligibility may be defined as "the understanding of spoken words" (Webster, 1979).

A more formal definition can be found in the Proposed American National Standard Method for Measuring the Intelligibility of Speech Communication Systems (1983). Here, intelligibility is defined in terms of a system.

> Intelligibility is the property of speech communication systems which enables trained listeners to receive and to identify speech spoken by trained talkers or by speech synthesizers. Although the term articulation sometimes has been used to mean intelligibility, preferred practice is to reserve articulation to describe the ability of an individual to produce sounds which can be identified as speech. Similarly, discrimination is reserved to describe the ability of an individual to recognize sounds as speech. Thus articulation and discrimination are

> characteristics of individuals, and intelligibility is a characteristic of a combination of environment and equipment when used by individuals and by groups of people.

Several language factors affect intelligibility including vocabulary size, word frequency (familiarity) effects, number of response alternatives, number of syllables, phonetic elements and context (Webster, 1979). Other external considerations include equipment or design features and environment (Webster and Allen, 1972). With such a wide range of sources of influence, it is easy to understand the need for speech intelligibility testing.

Speech intelligibility testing is used in a noisy environment to determine how well speech is understood. Performance is usually measured by calculating the percent correct (also known as the word-intelligibility score) for each subject. The intelligibility of different communication systems is compared by examining percent correct for each system. Normally scores of 70% are acceptable, however when rhyme words are used scores should be above 85% (Webster, 1979).

Modified Rhyme Test

One test found to be particularly useful in studying the effects of aircraft noise has been the Modified Rhyme Test (MRT). Because of its' reliability (T. J. Moore, personal communication, March, 1985), the MRT has been proposed as a standard for use in speech intelligibility testing in Proposed American National Standard: Method for Measuring the Intelligibility of Speech Communications Systems (1983).

The MRT was developed by House, Williams, Hecker and Kryter in 1965. The MRT consists of six lists (A, B, C, D, E, F in Table 1) of 50 monosyllabic American English words. The six lists are constructed to form 50 groups of six rhyming words (e.g., group 44 in Table 1 is comprised of the words meat, feat, heat, seat, beat, neat). Each individual group is considered to be a response set. A target word from the response set is randomly chosen for verbal presentation to subjects by a trained speaker. Subjects respond by selecting the correct target word from among the six alternatives. It should be noted that House, et al. (1965) have shown the intelligibility (i.e. percent correct) of all six lists to be equivalent.

Single task communication paradigm

The MRT has been used extensively by the Bioacoustics Branch of the Armstrong Aerospace Medical Research Laboratory (AAMRL) at Wright-Patterson Air Force Base in Dayton OH to test the intelligibility of aircraft communication systems in noise environments. They have

Table 1

Modified Rhyme Test (MRT) Lists

The stimulus words are arranged acccording to lists A-F. Each row represents a response set. In the first 25 sets, the final consonantal element is varied; in the second 25 sets, the reverse is true. These lists are in quasialphabetic order, not in order of use. (Source: House, et al. 1965)

	Forms					
	A	B	C	D	E	F
1	bat	bad	back	bass	ban	bath
2	bean	beach	beat	beam	bead	beak
3	bun	bus	but	buff	buck	bug
4	came	cape	cane	cake	cave	case
5	cut	cub	cuff	cup	cud	cuss
6	dig	dip	did	dim	dill	din
7	duck	dud	dung	dub	dug	dun
8	fill	fig	fin	fizz	fib	fit
9	hear	heath	heal	heave	heat	heap
10	kick	king	kid	kit	kin	kill
11	late	lake	lay	lace	lane	lame
12	map	mat	math	man	mass	mad
13	page	pane	pace	pay	pale	pave
14	pass	pat	pack	pad	path	pan
15	peace	peas	peak	peal	peat	peach
16	pill	pick	pip	pig	pin	pit
17	pun	puff	pup	puck	pus	pub
18	rave	rake	race	rate	raze	ray
19	sake	sale	save	sane	safe	same
20	sad	sass	sag	sack	sap	sat
21	seep	seen	seethe	seed	seem	seek
22	sing	sit	sin	sip	sick	sill
23	sud	sum	sub	sun	sup	sung
24	tab	tan	tam	tang	tack	tap
25	teach	tear	tease	teal	team	teak
26	led	shed	red	bed	fed	wed
27	sold	told	hold	fold	gold	cold
28	dig	wig	big	rig	pig	fig
29	kick	lick	sick	pick	wick	tick
30	book	took	shook	cook	hook	look
31	hark	dark	mark	lark	park	bark
32	gale	male	tale	bale	sale	pale
33	peel	reel	feel	heel	keel	eel
34	will	hill	kill	till	fill	bill
35	foil	coil	boil	oil	toil	soil
36	fame	same	came	name	tame	game
37	ten	pen	den	hen	then	men
38	pin	sin	tin	win	din	fin
39	sun	nun	gun	fun	bun	run
40	rang	fang	gang	bang	sang	hang
41	tent	bent	went	dent	rent	sent
42	sip	rip	tip	dip	hip	lip
43	top	hop	pop	cop	mop	shop
44	meat	feat	heat	seat	beat	neat
45	kit	bit	fit	sit	wit	hit
46	hot	got	not	pot	lot	tot
47	nest	vest	west	test	best	rest
48	bust	just	rust	must	gust	dust
49	raw	paw	law	jaw	thaw	saw
50	way	may	say	gay	day	pay

They have built a unique facility called Voice Communication Research and Evaluation System (VOCRES). This facility has been specifically designed to test a listener's comprehension of various communication messages under varying conditions of aircraft noise.

The general approach when using VOCRES involves the participation of volunteers who communicate as talkers and listeners under controlled conditions that replicate the specific communication environments being evaluated. Subjects are stationed at custom-designed consoles and communicate with standardized or special purpose (speech) vocabulary materials. Selected system and environmental characteristics or equipment are varied and the resulting communication effectiveness is measured. (McKinley, 1980). A more detailed description of VOCRES is found in Appendix A - see pages 58-68.

In AAMRL's research studies, two important parameters, 1) the level of the sound (overall sound pressure level (OASPL)) in decibels (dB) and 2) the spectrum of the sound were recorded to insure that future researchers would be able to replicate the study (C. W. Nixon, personal communication, March, 1985). The levels of unweighted noise (85 to 115 dB) used in these studies are representative of noise levels found in most military aircraft while in operation. The spectrum of sound was usually shaped to be representative of one particular type of aircraft noise.

A pink noise source, rather than the familiar white noise source, was used to generate these noise levels and spectrums. White noise is defined to have equal energy per frequency while pink noise is defined to have equal energy per octave band (See Figures 1 and 2 for a comparison of the spectrum levels for pink and white noise). The pink noise generator and the spectrum shaper allowed AAMRL to accurately reproduce real world noise conditions found in specific operational situations, which is essential for valid communication testing. Pink noise has been used extensively in acoustical research (C. W. Nixon, personal communication, March, 1985) and also better represents generic aircraft noise than does white noise (McKinley and Carr, 1984).

In the following studies multiple male and female talkers presented MRT lists in noise. Typically, five talkers (three male and two female) verbally presented the six MRT lists to a group of ten male and female subjects who responded by selecting the MRT target word from among six rhyming words presented on a cathode ray tube. Subjects wore helmets which typically attenuated each

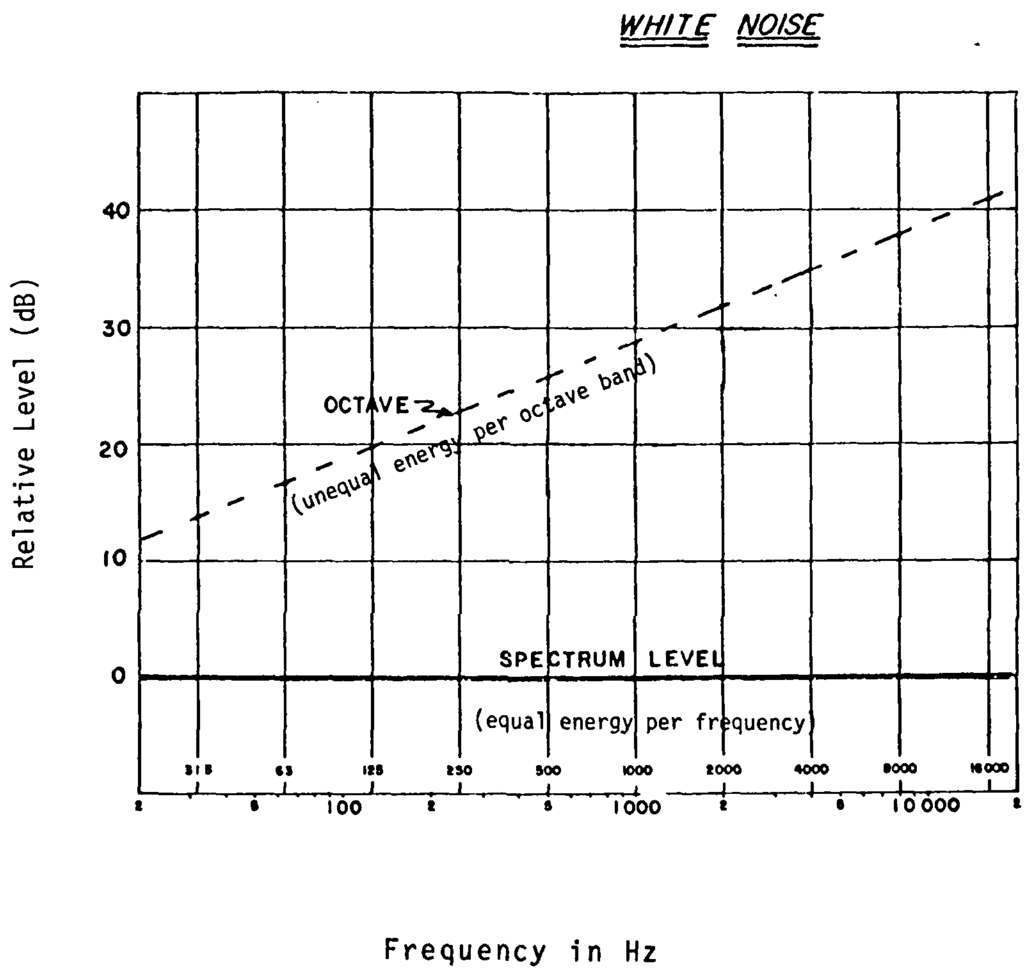

Figure 1. White noise, shown above, has equal energy levels for each frequency but has increasing levels of energy for higher octave bands. Contrast this with pink noise shown in Figure 2.

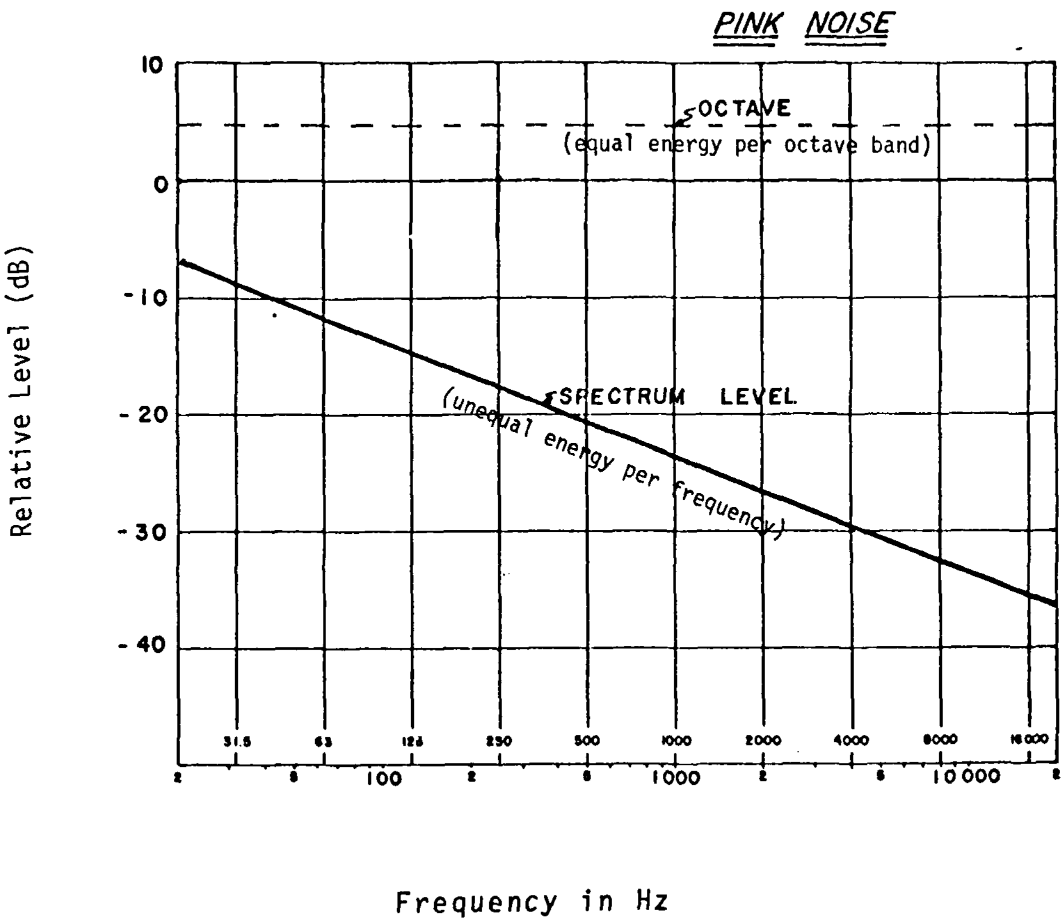

Figure 2. Pink noise, shown above, has equal levels of energy in each octave band but decreasing levels of energy at the higher frequencies.

Moore, McKinley and Mortimer (1979) examined the effect of aircraft noise on performance using two standard Air Force radios (i.e., ARC-34 and ARC-164). A pink noise source generated three levels of aircraft noise (95, 105 and 105 dB OASPL). Ten subjects were given the MRT in the VOCRES listening center. A randomized within subjects design was used for each radio with each subject participating daily in four 45 minute listening sessions. The ARC-34 radio was evaluated first and the ARC-164 radio was evaluated second. Data was also collected for the ARC-34 under an ambient noise (85 dB OASPL) condition. Figure 3 summarizes the results. Inspection of this figure shows that when aircraft noise was increased intelligibility (percent correct) of the communications signal decreased.

Moore, Nixon and McKinley (1980) also compared performance of the ARC-34 and ARC-64 with a third radio and an intercom system. The MRT was used to determine intelligibility under three noise conditions using an intercom system (AIC-25) and three aircraft radios (ARC-164, ARC-150 and ARC-34). Again, a pink noise source generated three levels of aircraft noise (95, 105 and 115 dB OASPL) in the VOCRES listening center. Ten subjects participated in four daily 45 minute listening sessions.

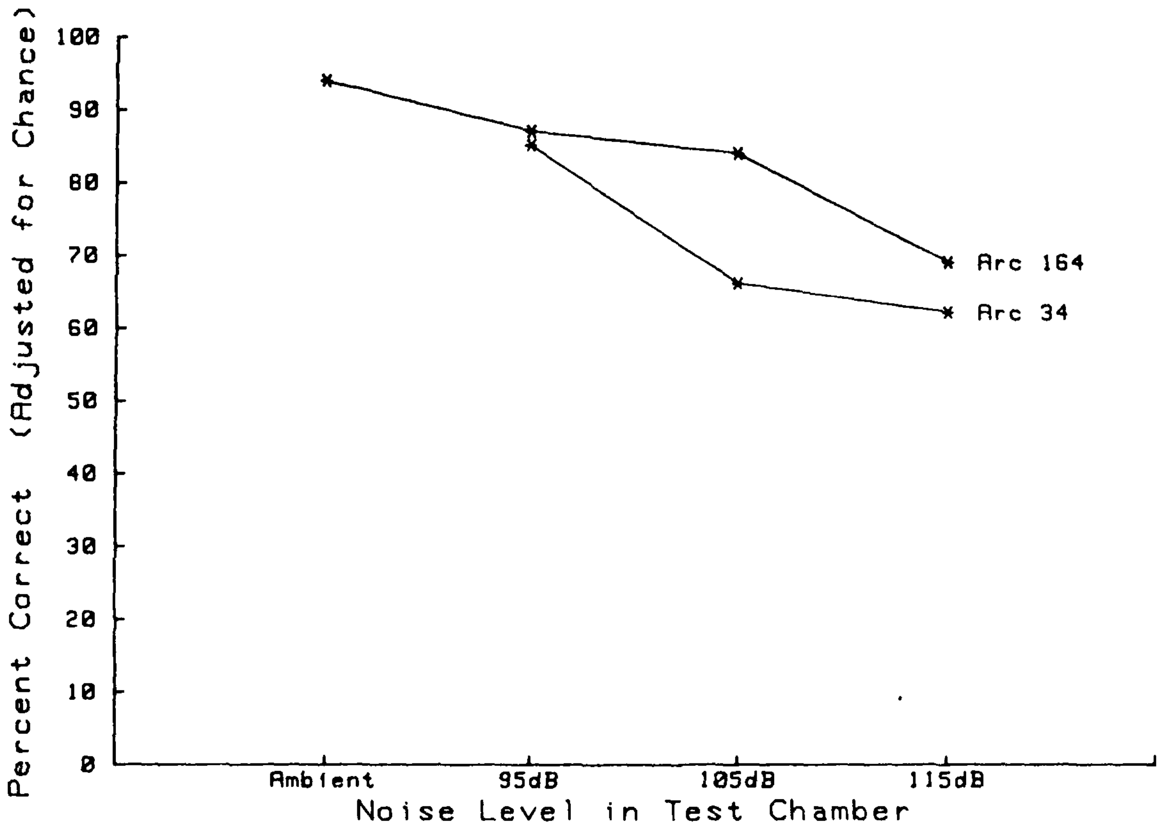

Figure 3. MRT percent correct (adjusted for guessing) for the ARC-164 and ARC-34 radio systems. (Source: Moore, McKinley and Mortimer, 1979)

Nine randomizations of six lists were presented to each subject. Results are shown in Figure 4. As in the earlier study, inspection of the figure shows that as noise increased intelligibility of the systems decreased.

In 1981, McKinley, Nixon and Moore evaluated several types of USAF standard in flight and ground voice communication equipment. A pink noise source generated three levels of aircraft noise (95, 105 and 115 dB OASPL). This study further replicated results of the previous two studies. Again, although results are not shown, as noise was increased intelligibility (MRT performance) decreased.

From these studies, one might conclude that the percent correct on the MRT will decrease with increasing levels of aircraft noise. Generally we would expect this to be true. However, one criticism of the previous studies is that the MRT was accomplished as a single task in noise. The noise-communication literature suggests that performance in the single task paradigm is unrealistic. The MRT performance decrement due to noise in previous research may have been underestimated and might be much larger under dual task conditions. Adding a second task could cause a greater decrement in intelligibility or the decrement could occur at lower intensities of noise.

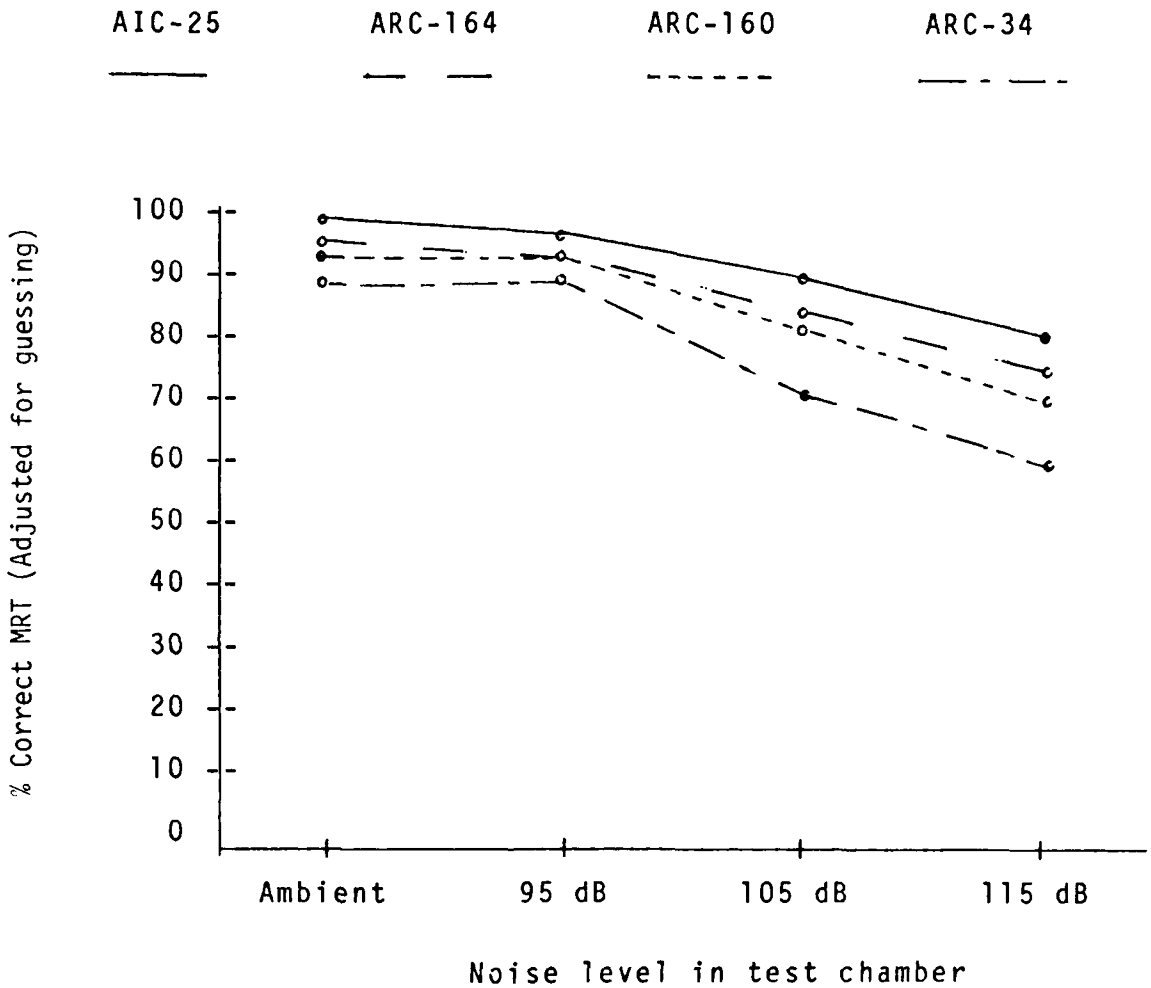

Figure 4. MRT percent correct (adjusted for guessing) for the AIC-25 intercomm and the ARC-164, ARC-150 and ARC-34 radios. (Source: Moore, Nixon and McKinley, 1980)

Dual task paradigm

Knowles (1963) states why performance on a single task does not sufficiently load the operator.

> The results of part-task simulations are often deficient in that the performance appears unduly good because the operator is permitted to focus all his attention on the part-task whereas he must share his attention in the total job situation.

Johnston (1975) adds

> In standard laboratory evaluations of communications systems by intelligibility tests, only the audio-environment is stimulated. No other demands are made on the operator, so high performance scores on intelligibility tests can be achieved. If a system is designed and evaluated against conventional intelligibility scoring methods it may appear adequate for use but it may prove inadequate because of the unlimited attention which users can devote to it.

When subjects are allowed to devote their full time to the listening task, their performance may not be representative of typical listening in real-world situations when there are other tasks and distractions. These statements suggest that communication testing in a dual-task environment would better simulate real world listening situations.

The present study

The primary purpose of the present study was to determine the effect of adding a second task to the current single task noise communication paradigm. The underlying real world question was: If a pilot is performing a communication task in noise, does performance change if a second task is added?

A dual task paradigm was used in this study to "bring pressure" on the primary MRT task. Knowles (1963) explains that the secondary task has been used in two distinct ways in previous studies. First, although not used in the current study, it has been used as a measure of operator workload. In this situation, the operator performs the primary and secondary tasks concurrently. If the operator performs well on the secondary task, it is interpreted as an indication that the primary task imposes low operator workload. If he is unable to perform the secondary task, this indicates that the primary task imposes high operator workload.

Another use of secondary tasks, and the one of primary concern in this study, is as a loading task on a subject's primary task performance. As Knowles (1963) indicates when utilized in this manner "there is little interest in the secondary task performance per se. The secondary task is used simply to bring pressure on the primary task."

Subjects performed two tasks in varying levels of aircraft noise. The primary task was a communication task (Modified Rhyme Test; House et al. 1965) while the secondary task was a short-term memory search task (Sternberg, 1969). Simulated aircraft noise was varied between levels of 95, 105 and 115 dB using a pink noise source. Task loading was varied with Sternberg's memory search (MS) task by requiring subjects to memorize one, four, or six alphanumeric characters. The MS task was chosen due to its close representativeness to an actual pilot's task while in flight (e.g., a pilot will typically be required to remember coordinates displayed on a CRT). The combination of the MS task with the MRT task is representative of a dual-task situation where the pilot monitors a CRT screen while receiving an auditory message.

By introducing a second task, one would expect a subject's MRT performance to decrease while his impression of task loading would be expected to increase. As a check for increase in task loading the Subjective Workload Analysis Technique (SWAT) was employed. SWAT uses the dimensions of Time Load, Mental Effort Load and Stress Load to indicate changes in subject workload. Subjects rate the task on these dimensions for each trial. Scores are developed using a conjoint measurement technique (a more detailed discussion of SWAT and its development is given in Appendix B - see pages 69-72). This technique has been used in a variety of experimental situations (e.g., Eggemeier, McGhee and Reid (1983); Eggemeier, Melville, Crabtree (1984) and Courtright and Kuperman (1984)).

A secondary interest in this study was to focus more attention on the MRT response time measure. In previous AAMRL studies, percent correct and response time data were collected for all noise conditions, where percent correct was typically used as the primary measure of adequate system performance. Only when percent correct was equivalent for the systems being tested was response time examined. Response time will be given a more thorough examination in the present study.

This study was accomplished in three phases. In Phase 1 subjects performed the communication task in simulated aircraft noise to determine the effect of noise on the MRT task. Phase 1 attempts to replicate previous research by the Bioacoustics Branch of AAMRL which show that increased levels of noise cause a decrement in MRT performance. Phase 1 was a baseline to determine the noise). In Phase 2 subjects accomplished the communication and the memory search tasks together in an ambient noise condition to determine the effect of a adding a second task to the MRT task. Phase 2 was a baseline to determine the effect of adding a second task to the MRT (dual task without noise). In Phase 3 subjects completed both tasks in simulated aircraft noise to determine the joint effect of noise and task loading on MRT task performance.

Phase 1 performance (single task in noise) was compared with Phase 3 performance (dual task in noise) to determine the effect of adding a second task to the standard communication paradigm. Phase 2 performance (dual task without noise) was compared to Phase 3 (dual task in noise) performance to determine the effect of adding noise to MRT/MS dual task.

METHOD

Subjects

Ten right-handed male subjects aged 18-25 participated as subjects. All subjects had many hours experience in the VOCRES system performing the MRT listening task for various communication systems. All subjects had normal hearing which was verified through periodic hearing tests during the course of the experiment. Subjects were paid an hourly rate with a built-in bonus paid if they were present for all sessions.

Experimental design

Ten subjects participated in a three phase within-subjects study. In Phase 1 three levels of noise (95, 105 and 115 dB) were factorially combined with three MRT lists in nine trials. Response time and percent correct on the MRT task were recorded for each trial. A SWAT score was also recorded for each trial. In Phase 2 three levels of MS task loading (one, four and six memory set characters) were factorially combined with three MRT lists in nine trials. Response time and percent correct on both the MRT and MS tasks were recorded during each trial. A SWAT score was also recorded for each trial. In Phase 3 three levels of noise (95, 105 and 115 dB), three task loading levels (one, four and six memory set characters) and three MRT lists were factorially combined in 27 trials. Response time and percent correct on both the MRT and MS tasks were recorded during each trial. A SWAT score was also recorded for each trial.

Test facilities

The experiment was accomplished using the Performance and Communications Research and Technology (PACRAT) System (see Figure 5) at the Biological Acoustics Branch of AAMRL at Wright-Patterson Air Force Base, Dayton OH. This system is very similar to the VOCRES system used in the previous AAMRL noise communication studies. In the PACRAT system each subject's console was designed to be more representative of an typical cockpit environment (see Figure 6).

The MRT communications task was generated on the PACRAT system using a Perkin Elmer computer while the Sternberg MS task was generated by a Commodore 64 computer (For more information on the hardware and software used to generate the MS task, see Appendix C - pages 73-76). Two adjacent five-inch diagonal cathode ray tube (CRT) screens in the PACRAT console (see Figure 7) were used to present the stimulli to subjects. The MRT response sets always appeared on the right CRT screen.

Figure 5. PACRAT System

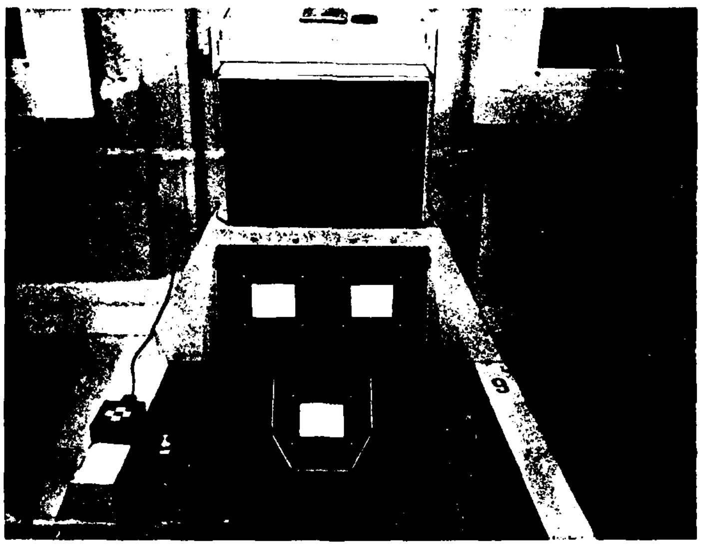

Figure 6. PACRAT Console

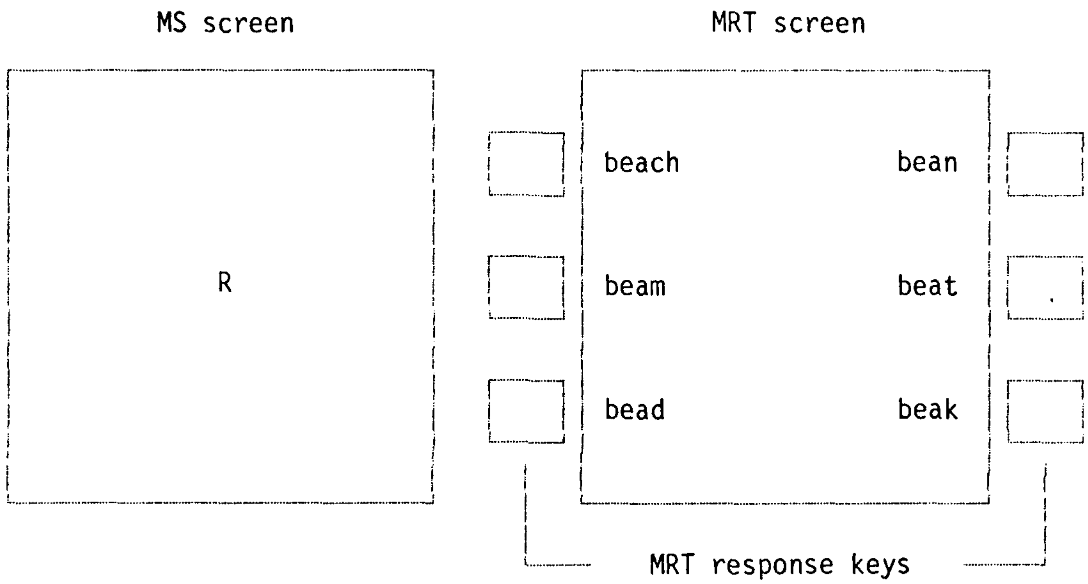

Figure 7. CRT's used for single and dual tasks. The left CRT was used for the MS task while the right CRT was used for the MRT.

One MRT response button was adjacent to each of the six response word locations on the right CRT. The MS stimulli always appeared on the left CRT screen. The MS response keypad was placed on the left side of the PACRAT console.

The acoustical environment was simulated using the large reverberation chamber in the PACRAT facility. The PACRAT system is capable of operation in a high power modes (14,000 watts) and a low power mode (12,000 watts). The power amplifiers drive eight banks of loudspeakers containing a total of 96 Altec 15" low frequency speakers, eight Altec horn loaded compression drivers and 384 Stromberg Carlson high frequency speakers. A pink noise generator produced the three levels of aircraft noise (95, 105, 115 dB overall sound pressure level (OASPL)). For greater generalization, the spectrum of sound used in the present study was only shaped by the loud speakers in the acoustical chamber.

During training and the experimental phases the levels of simulated aircraft noise were measured with a Hewlett Packard 9845A spectrum analyzer by placing a microphone parallel to both ears at a distance of six inches from a subjects helmet. The frequency spectra for all levels of aircraft noise are shown in Appendix D - see pages 77-85. The frequencies generated by the noise did not exceed the normally accepted hearing range for humans (20 to 20,000 Hz). Exposure to noise levels did not exceed the safety and time limits given in Air Force Regulation 161-35, Hazardous Noise Exposure (1982).

Materials

MRT lists. The six lists developed by House et al, 1965 (see Table 1) were presented to the subjects. A live male speaker presented all the MRT lists in all conditions for all phases. The speaker had many hours of previous experience speaking in the VOCRES system. Each list was composed of 50 response words. These response words were presented in a carrier phrase (e.g., "Number one, you will mark beat, please." - where beat was the response word). Six response words (e.g., a response set) were presented for each trial. Each of the lists had previously been shown to be equivalent in intelligibility at noise levels of 95, 105, and 115 dB.

MS stimuli. The MS stimuli developed by Shingledecker (1984) were presented to the subjects. Stimulus items in the MS task were visually presented alphabetic characters. Due to the acoustic confusability of certain letters, only 15 of the 26 letters of the alphabet were used in the task (ABCEFHIJLORSXYZ). Memory set items were randomly selected from the letter population, and the remaining items were used in the negative set. Test items were also randomly generated with the restriction that positive and negative set items were drawn with equal probability.

Tasks

MRT. MRT lists were presented to the subjects using the PACRAT intercom system and subjects responded on the right CRT in the PACRAT console (see Figure 7). The trained male speaker presented one of three MRT lists in one of three presentation sequences (i.e., three presentation sequences were randomly generated for each MRT list) in noise levels of 95, 105 and 115 dB. For each trial a carrier phrase (e.g., "Number one, you will mark beat, please.") was presented on the speaker's console. The speaker read the carrier phrase to the subjects over the PACRAT intercom system.

The subjects task was to listen for response words (e.g., for the example above beat would be the response word) presented in the carrier phrase. When the carrier phrase was spoken, subjects responded by choosing one of six rhyming words presented on the right CRT. Response word location on the CRT was randomly varied between six standard positions (see Figure 7) by the PACRAT system.

A trial consisted of 50 carrier phrases presented to each subject and was five minutes and 33 seconds in duration. Consequently, a carrier phrase was presented to each subject once every 6.62 seconds. Each subject was instructed to identify and select the appropriate response word as rapidly and as accurately as possible. Immediate verification of a subject's response was given via a red light emitting diode (LED) on the display panel.

Subjects could change their response but were required to respond within the 6.62 second window or a miss was recorded. Subjects were instructed to guess if unsure of the response word. Since the number of possible responses was greater for the MRT (six possible responses) than for the MS task (two possible responses) all subjects were required to use their preferred hand (right-hand) for the MRT task. Subjects were given feedback about their performance (number of correct responses) at the end of each trial.

MS. Task loading was varied via the MS task. MS stimulli were presented to the subjects on the left CRT and subjects responded on the left CRT in the PACRAT console. The subjects task was to memorize character sets containing one, four or six alphabetic characters. Once memorized, single alphabetic characters were successively presented to subjects. Subjects performed a recognition task using the response keypad previously shown in Figure 6. "Yes" and "No" responses were made using the left and right keys respectively. The top and bottom keys were not used. Subjects were to respond "yes" if the character presented was in the memory set and "no" if the character was not in the memory set.

A trial consisted of continuous stimuli presented at a rate controlled by the subject and lasted 5 minutes and 33 seconds to coincide with the MRT task. A new character, randomly chosen out of the 15 possible characters, was presented on the left CRT immediately after each response. Subjects who responded more quickly received more stimuli. Subjects could not change their response and were required to respond to each character within a 6.62 second window (the same window given for the MRT task) or a miss was recorded. The MS task (2 possible responses) was performed with subjects non-preferred hand (left-hand). Subjects were given feedback about their performance (response time, number of correct stimuli, percent correct) at the end of each trial.

<u>SWAT card sort</u>

All subjects were required to complete the SWAT card sort by rank ordering 27 cards representing unique combinations of the three parameters of SWAT (Time Stress, Mental Stress and Psychological Stress). Each card represented a unique level of time, mental or psychological stress. Instructions included a brief paragraph introducing time mental effort and stress load. Independence of these dimensions was stressed. Each subject took between 30 minutes and one hour to complete the SWAT card sort (SWAT card sort instructions to subjects are provided in Appendix E - see pages 86-90).

Procedure

Tra· ing. Subjects were trained in the PACRAT console previously shown in Figure 6 (see page 22). A live male talker with many hours of experience presented the MRT lists using a standard AIC-25 intercom system. Four subjects were tested in the morning (morning subjects) while six were tested in the afternoon (afternoon subjects). In training, morning subjects 1-2 received MRT lists 1-3 and morning subjects 3-4 received MRT lists 4-6 while afternoon subjects 5-7 received MRT lists 1-3 and afternoon subjects 8-10 received MRT lists 4-6.

Subjects wore the AIC-25 compatible terminal headgear with oxygen masks and used an air breathing system controlled with a standard Air Force A-19 regulator. The signal-to-noise (S/N) ratio for each headset was adjusted to a comfortable level by each subject. MS stimuli were presented on the PACRAT system using a Commodore-64 computer. Subjects were trained in the MS single task and MRT and MS dual task conditions. Subjects were given practice in rating task difficulty using the SWAT technique.

Training on the MRT task alone in PACRAT was not given since all subjects had many hours of experience performing this task in the VOCRES system and were adequately trained (MRT instructions to subjects are provided in Appendix F as a source of reference for future research - see page 91). However, subjects were given training on the MS task alone. One, four or six memory set characters were randomly presented to each subject. Subjects responded using the MS response keypad.

Shingledecker (1984) recommended a minimum of 21 minutes of training at each level to eliminate the effects of learning. Subjects actually received 33 minutes of training at each level to ensure that their performance had reached an asymptotic level (MS task instructions to the subjects are provided in Appendix G - see page 92). Subjects provided SWAT ratings immediately following each trial (SWAT task instructions to the subjects are provided in Appendix H - see page 93).

Subjects completed the MRT/MS dual task both without noise and with noise. For each condition subjects were presented with 24 trials with three randomizations of three MRT lists. MS loading levels were also randomly presented to subjects. Subjects provided SWAT ratings immediately following each trial (MRT/MS dual task instructions to subjects are provided in Appendix I - see page 94).

Experimental conditions. Subjects were tested in the PACRAT facility with the same procedures used in training. Four subjects were tested in the morning (morning subjects) while six were tested in the afternoon (afternoon subjects). During experimental conditions (Phases 1-3), morning subjects 1-2 received MRT lists 4-6 and morning subjects 3-4 received MRT lists 1-3 while afternoon subjects 5-7 received MRT lists 4-6 and afternoon subjects 8-10 received MRT lists 1-3.

Since subjects performed two tasks in Phase 2 and Phase 3, performance on the MRT was stressed through instructions to the subjects. Subjects were instructed to perform both tasks if possible. However, if they were unable to respond to both, they were to ensure that they responded to the MRT.

Phase 1 was a baseline which replicated previous research and determined the effect of noise on the MRT performance (response time, percent correct). Subjects performed the MRT task alone under three levels of noise (95, 105 and 115 dB) factorially combined with three MRT lists in nine trials. Subjects were given three blocks of three trials each. Each noise level and list occurred once within a block. The particular list which occurred with a given noise level and the order of presentation within a block was determined using a Greco Latin Square (see Table J-1 in Appendix J - page 95). Response time and percent correct on the MRT task were recorded for each trial. Subjects provided SWAT ratings immediately following each trial.

Phase 2 was a baseline which determined if adding a second task affected MRT performance (response time, percent correct). Subjects performed the MRT/MS dual task combination without noise for three levels of MS task loading (1, 4 and 6 memory set characters) factorially combined with three MRT lists in nine trials. Subjects were given three blocks of three trials each. Each task loading level and MRT list occurred once within a block. The particular list occurring with a given task loading level and the order of presentation within a block was determined using a Greco Latin Square (see Table J-2 in Appendix J - page 95). Response time and percent correct on both the MRT and MS tasks were recorded for each trial. Subjects provided SWAT ratings immediately following each trial.

Phase 3 determined the combined effect of noise and the task loading on MRT performance (response time, percent correct). Subjects performed the MRT/MS dual task combination in noise for three levels of noise (95, 105 and 115 dB), three levels of MS task loading (1, 4 and 6 memory set characters) and three MRT lists. These conditions were factorially combined in twenty seven trials. Subjects were given three blocks of nine trials each. Each noise level, task loading level and MRT list occurred once within a block. The particular list which occurred with a given noise and task loading level and the order of presentation within a block was determined using a Greco Latin Square (see Table J-3 in Appendix J - page 96). Response time and percent correct on both the MRT and MS tasks were recorded for each trial. Subjects provided SWAT ratings immediately following each trial.

RESULTS

Results are reported according to phases. Phase 1 replicated previous research to determine the effect of noise on the MRT performance (response time, percent correct). Phases 2 and 3 represent dual task situations where the MRT and MS were performed without noise (Phase 2) and with noise (Phase 3). Performance on the MRT was emphasized throughout all phases. In each phase the data are analyzed according to task (i.e., MRT, MS or SWAT).

Repeated Measures Multivariate Analysis of Variance (MANOVA) techniques were used to analyze the data from all three experimental phases. Since subjects were instructed to guess on the MRT, scores were adjusted for guessing using a standard correction factor cited by Brown (1983): # correct = # right - (# wrong/ # of choices available - 1). Percent correct scores were initially transformed using a standard arcsin transformation (i.e., $X_T = 2*\arcsin X$). This is recommended (Tukey, 1977) for scores such as percent correct. Statistical analysis for the untransformed and transformed data yielded the same results (i.e., the same effects were significant for the transformed and untransformed data). Therefore, only analysis of untransformed data are reported here.

When the multivariate F statistic (Pillai's trace) was significant, univariate Analysis of Variance (ANOVA) and post-hoc comparisons were performed on the MRT and MS data to more precisely determine the locus of effects. Post-hoc comparisons between pairs of means were done using the Tukey (1977) critical difference test. In addition, ANOVA was performed on the SWAT scores as a validation check on variation of loading levels (memory set size).

Phase 1

Phase 1 MRT data was analyzed to determine the effect of noise on response time and percent correct. Phase 1 provided a baseline for the single MRT task in noise. Results for Phase 1 are shown in Table 2.

MRT. Inspection of Table 2 shows that as noise increased from 95 to 115 dB response time changed very little while percent correct decreased. MANOVA performed on this data indicated that the effect of noise on the combination of dependent variables (DV's) was significant $F(4,36) = 4.24$, $p = 0.007$. Individual ANOVAs indicated that the effect of noise on response time was non-significant $F(2,18) = 0.68$, $p = 0.521$, while it's effect on percent correct was significant $F(2,18) = 13.53$, $p < 0.001$.

Table 2

Means and standard deviations for Phase 1

Modified Rhyme Test (MRT)

Response Time (RT) in seconds

95 dB		105 dB		115 dB	
Mean	S.D.	Mean	S.D.	Mean	S.D.
2.32	0.21	2.35	0.20	2.35	0.22

Percent Correct (PC)

95 dB		105 dB		115 dB	
Mean	S.D.	Mean	S.D.	Mean	S.D.
98.88	0.94	98.44	1.13	95.92	1.83

Subjective Workload Analysis Technique (SWAT)

95 dB		105 dB		115 dB	
Mean	S.D.	Mean	S.D.	Mean	S.D.
1.69	5.81	2.31	6.55	6.55	8.80

Tukey analysis showed that percent correct for the 115 dB noise condition was significantly lower than both the 95 and 105 dB noise conditions while these conditions did not differ from each other.

SWAT. As indicated in Table 2 SWAT ratings increased as noise increased. ANOVA analysis indicated a significant effect of noise, $F(2,18) = 6.46$, $p = 0.008$. Further Tukey analysis showed that SWAT ratings for the 95 and 105 dB noise conditions did not significantly differ while both were significantly less than ratings for the 115 dB condition. While these results are significant, the SWAT scores show a large amount of variability (i.e., standard deviation for each noise level is greater than the mean) for each level of noise while showing a low overall level of task loading (i.e., average score for 115 dB = 6.55).

Phase 2

The effect of performing both the MRT and MS, without adding noise, was evaluated and results are given in Table 3.

MRT. MRT performance (see Table 3) changes very little as memory set size increased from one to six characters. Memory set size had a non-significant effect on the combination of response time and percent correct $F(4,36) = 1.02$, $p = 0.411$.

Table 3

Means and standard deviations for Phase 2

Modified Rhyme Test (MRT)

Response Time (RT) in seconds

	Mean	S.D.
1 char	2.61	0.26
4 char	2.59	0.22
6 char	2.57	0.22

Percent Correct (PC)

	Mean	S.D.
1 char	98.08	1.61
4 char	98.00	1.66
6 char	98.64	1.20

Memory Search (MS) Task

Response Time (RT) in seconds

	Mean	S.D.
1 char	0.79	0.18
4 char	1.08	0.22
6 char	1.26	0.26

Percent Correct (PC)

	Mean	S.D.
1 char	99.46	0.57
4 char	99.18	0.77
6 char	97.92	2.25

Subjective Workload Analysis Technique (SWAT)

	Mean	S.D.
1 char	19.71	15.47
4 char	27.07	16.15
6 char	39.07	16.51

(task loading) on response time and percent correct. Phase 3 MS data was analyzed to determine if the combination of noise and loading levels caused a decrement in performance on the MS task. Tables 4 and 5 show the means and standard deviations for response time and percent correct for both tasks and SWAT in Phase 3.

MRT. The noise by memory set size interaction and the main effect of memory set size were non-significant $F(8,72) = 0.42$, $p = 0.908$ and $F(4,36) = 1.01$, $p = 0.417$. while the main effect of noise was significant $F(4,36) = 10.18$, $p < 0.001$. Inspection of Table 4 shows that both response time and percent correct decreased as a function of increase in noise. ANOVA confirmed that both response time and percent correct were significant $F(2,18) = 11.81$, $p = 0.001$ and $F(2,18) = 18.80$, $p < 0.001$. Tukey analysis indicated significantly slower response time for the 95 dB noise condition while showing no differences between the 105 and 115 dB condition. Tukey analysis for percent correct showed that the 95 and 105 dB noise conditions did not differ from each other while the 115 dB condition was significantly lower than both.

MS. The noise by memory set size interaction and the main effect of noise was non-significant $F(8,72) = 0.86$, $p = 0.558$ and $F(4,36) = 1.47$, $p = 0.233$ while the main

Table 4

Means and standard deviations for Phase 3

Modified Rhyme Test (MRT)

Response Time (RT) in seconds

	95 dB Mean	95 dB S.D.	105 dB Mean	105 dB S.D.	115 dB Mean	115 dB S.D.	AVE
1 char	2.66	0.25	2.61	0.22	2.59	0.20	2.62
4 char	2.68	0.23	2.63	0.24	2.58	0.17	2.63
6 char	2.71	0.19	2.59	0.26	2.59	0.18	2.63
AVE	2.68		2.61		2.59		

Percent Correct (PC)

	95 dB Mean	95 dB S.D.	105 dB Mean	105 dB S.D.	115 dB Mean	115 dB S.D.	AVE
1 char	98.40	0.75	98.16	1.73	94.56	4.06	97.04
4 char	98.32	1.66	98.00	1.32	95.28	2.05	97.20
6 char	97.60	1.25	97.28	2.17	94.16	2.42	96.35
AVE	98.11		97.81		94.66		

Memory Search (MS)

Response Time (RT) in seconds

	95 dB Mean	95 dB S.D.	105 dB Mean	105 dB S.D.	115 dB Mean	115 dB S.D.	AVE
1 char	0.89	0.24	0.81	0.17	0.93	0.32	0.88
4 char	1.07	0.15	1.11	0.19	1.12	0.24	1.10
6 char	1.45	0.53	1.34	0.31	1.32	0.30	1.37
AVE	1.14		1.09		1.12		

Percent Correct (PC)

	95 dB Mean	95 dB S.D.	105 dB Mean	105 dB S.D.	115 dB Mean	115 dB S.D.	AVE
1 char	99.56	0.41	99.60	0.31	99.55	0.39	99.57
4 char	99.11	0.73	99.12	0.75	99.05	0.82	99.09
6 char	98.36	1.39	98.50	0.95	98.69	1.01	98.52
AVE	99.01		99.09		99.10		

Table 5

Means and standard deviations for Phase 3 SWAT scores

Subjective Workload Analysis Technique (SWAT)

	95 dB		105 dB		115 dB		
	Mean	S.D.	Mean	S.D.	Mean	S.D.	AVE
1 char	20.50	14.71	19.95	13.85	26.86	23.99	22.44
4 char	28.10	17.29	30.22	16.28	31.35	15.14	29.89
6 char	41.13	20.45	39.32	15.53	42.74	18.71	41.06
AVE	29.91		29.83		33.65		

effect of memory set size was significant $F(4,36) = 6.58$, $p < 0.001$. ANOVAs for the effect of memory set size disclosed that both response time and percent correct was significant $F(2,18) = 48.15$, $p < 0.001$ and $F(2,18) = 12.78$, $p < 0.001$. Investigation with the Tukey procedure determined that response times increased as memory set size increased from one to six characters while for percent correct the one and four character conditions did not differ from each other but percent correct was significantly lower for the six character condition.

SWAT. The noise by memory set size interaction and the main effect of noise was non-significant $F(4,36) = 0.84$, $p = 0.512$ and $F(2,18) = 1.99$, $p = 0.165$ while the main effect of memory set size was significant $F(2,18) = 15.51$, $p < 0.001$. Tukey analysis for memory set size showed no difference between the one and four character conditions while the SWAT ratings were significantly higher for the six character condition.

Phase 1 vs Phase 3

The major purpose of this comparison was to determine the effect of adding a second task to the current single task communication paradigm. Inspection of Figure 8 shows that adding a second task to the MRT communication task paradigm increases MRT response time, decreases MRT

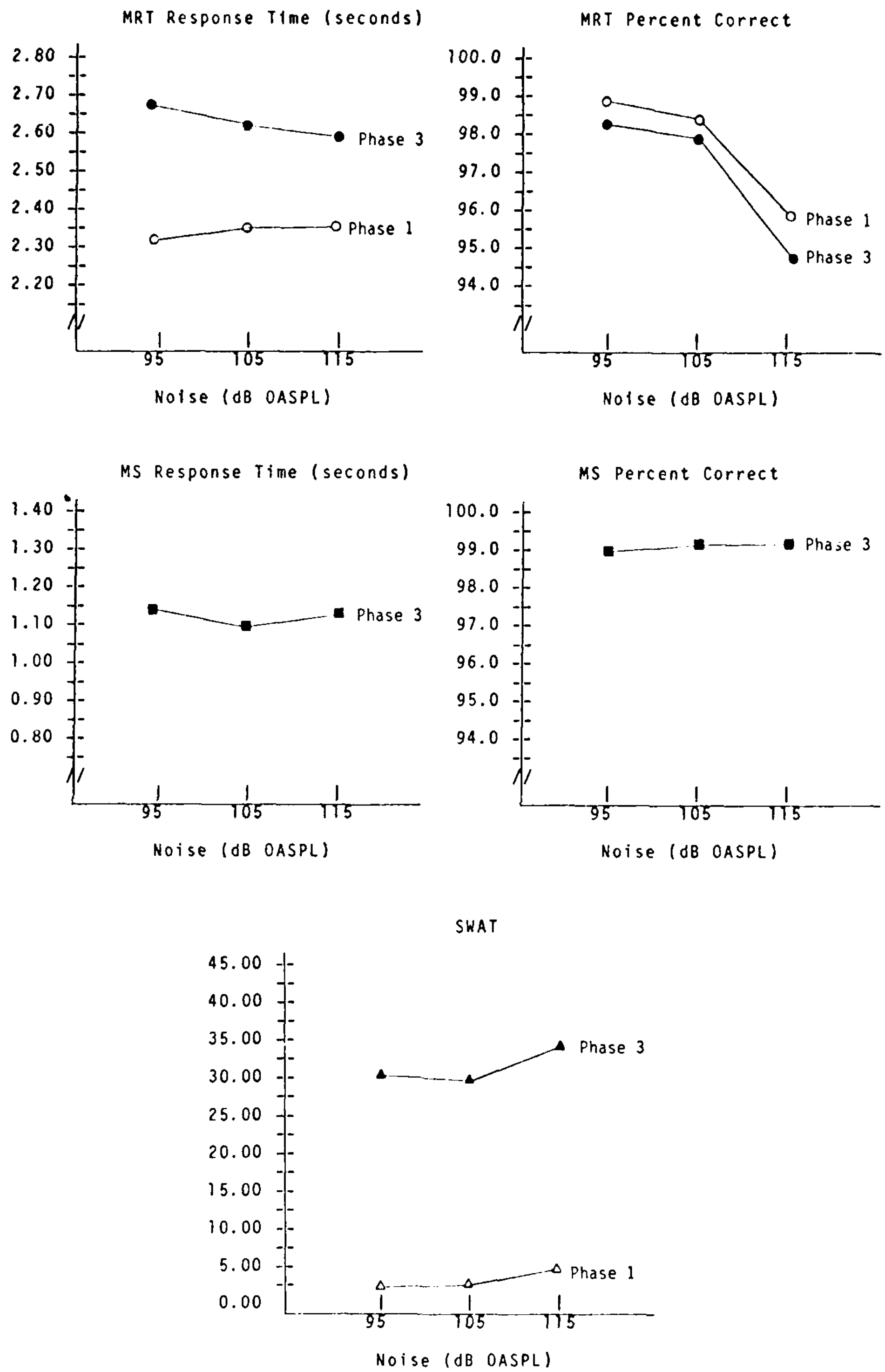

Figure 8. Dependent measures as a function of noise for Phase 1 (single task in noise) vs Phase 3 (dual task in noise).

percent correct and increases the perceived workload of the task. Multivariate planned comparisons performed on the MRT data revealed that the main effect of phase (i.e., single task in noise vs dual task in noise) was significant $F(2,8) = 15.32$, $p = 0.002$. Individual ANOVAs indicated that response time and SWAT scores were significantly greater for the dual task condition, $F(1,9) = 30.26$, $p < 0.001$ and $F(1,9) = 33.45$, $p < 0.001$ respectively, while showing no difference for percent correct, $F(1,9) = 5.03$, $p = 0.052$.

The noise by phase interaction was examined to determine if there was a differential effect of noise across the two paradigms. MANOVA revealed a significant noise by phase interaction $F(4,36) = 3.60$, $p = 0.014$. However, ANOVA indicated the interaction was only due to response time $F(2,18) = 9.69$, $p = 0.001$ while percent correct and SWAT scores were non-significant $F(2,18) = 0.59$, $p = 0.566$ and $F(2,18) = 0.10$, $p = 0.905$. The noise by phase interaction reflects the fact that response time decreased in the dual task paradigm (Phase 3) while it remained unchanged for the single task paradigm (see previously reported Tukey analyses in Phase 1 and Phase 3).

ANOVA revealed a significant main effect of noise for response time $F(2,18) = 5.65$, $p = 0.012$, percent correct

$F(2,18) = 21.66$, $p < 0.001$ and SWAT $F(2,18) = 3.58$, $p = 0.049$. As noise increased response time and percent correct decreased whereas SWAT scores increased. However, MS performance did not vary as a function of noise (as shown by previous Tukey analysis in Phase 3).

Phase 2 vs Phase 3

The effect of adding noise to the dual task paradigm was determined by comparing the results of the dual task without noise (Phase 2) with results for the dual task in noise (Phase 3). As Figure 9 shows, the addition of noise to the dual task paradigm appears to increase MRT response time, decrease MRT percent correct and increase subjective impressions of workload. MANOVA for the MRT data revealed that the main effect of noise (i.e., dual task without noise vs dual task in noise) was significant $F(2,8) = 7.50$, $p = 0.015$. However, separate ANOVAs showed that the effect of noise on both response time and SWAT scores was non-significant $F(1,9) = 1.21$, $p = 0.301$ and $F(1,9) = 2.00$, $p = 0.191$ respectively while the effect of noise on percent correct was significant $F(1,9) = 13.56$, $p = 0.005$. MANOVA performed on the MS data showed that the main effect of noise was non-significant $F(2,8) = 4.42$, $p = 0.051$. Thus adding noise to the dual task condition only significantly affected MRT percent correct.

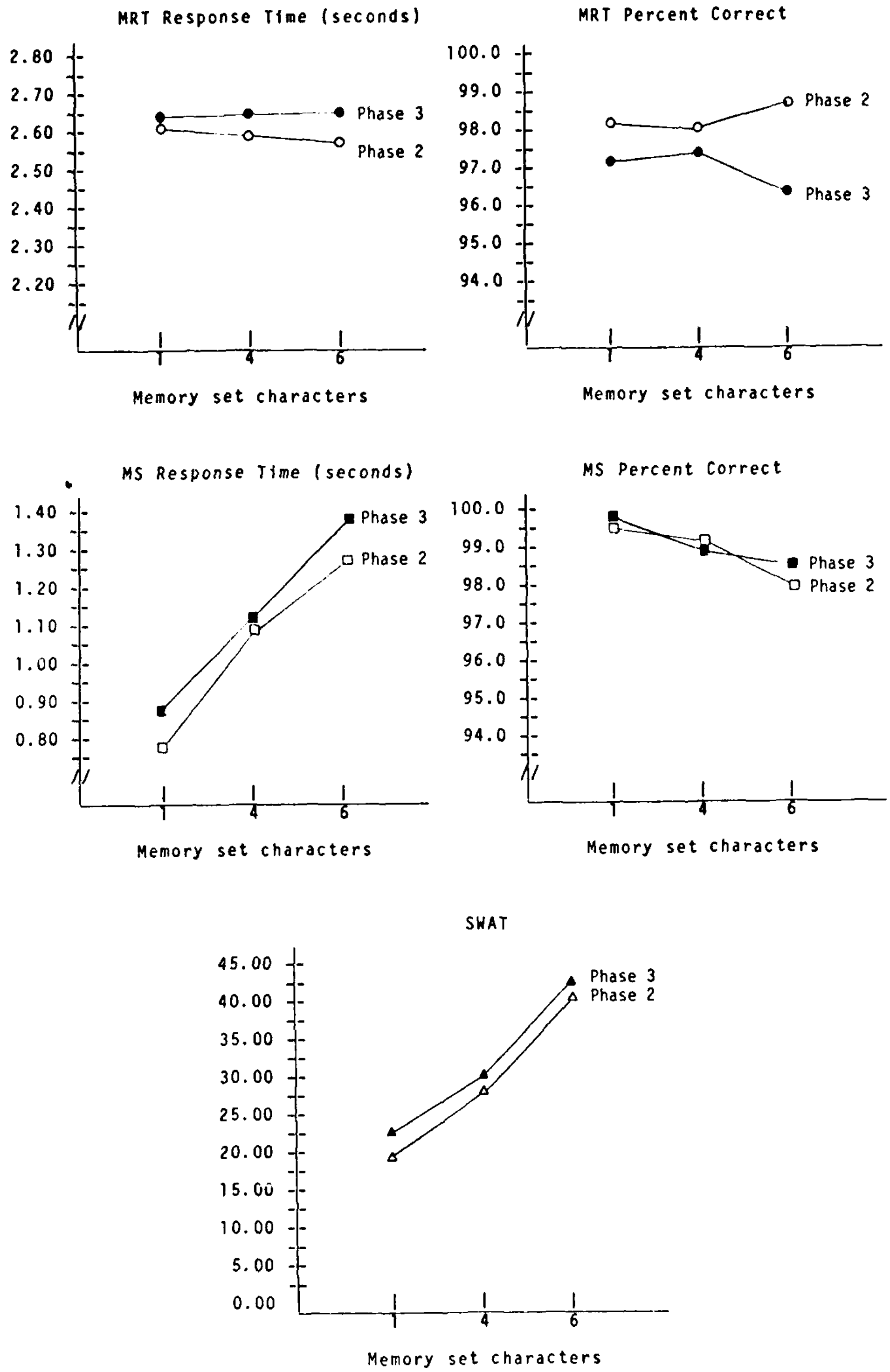

Figure 9. Dependent measures as a function of memory set size for Phase 2 (dual task without noise) vs Phase 3 (dual task in noise).

The memory set size by noise interaction was examined to determine if there was any differential effect of noise as a function of memory set size. Results indicated that the noise by memory set size interaction was non-significant for the MRT, MS and SWAT $F(4,36) = 1.62$, $p = 0.190$, $F(4,36) = 2.58$, $p = 0.054$ and $F(2,18) = 0.05$, $p = 0.955$ respectively.

Memory set size affected MS performance and SWAT scores but not MRT performance. ANOVA indicated that the main effect of memory set size was non-significant for MRT response time $F(2,18) = 0.11$, $p = 0.899$ and MRT percent correct $F(2,18) = 0.94$, $p = 0.408$ but was significant for MS response time $F(2,18) = 63.92$, $p < 0.001$, MS percent correct $F(2,18) = 10.54$, $p = 0.001$ and SWAT $F(2,18) = 16.39$, $p < 0.001$. As memory set size increases, MS response time increases while MS percent correct decreases. SWAT scores also increase as memory set size increases.

DISCUSSION

The major objective of this study was to simulate a dual task environment to determine the effect of the noise and task loading stressors on the MRT task. Previous research has extensively used the single communication task paradigm. It was anticipated that requiring subjects to perform a second task in noise would affect intelligibility on the MRT. There was a trend towards decreased overall intelligibility when the second task was added, although it was non-significant (i.e., $p = 0.052$). For both single and dual task conditions, intelligibility decreased as noise increased.

The lack of significant noise by task interaction for MRT percent correct indicated that the functional relationship between noise and intelligibility did not vary with the addition of the second task. Additionally, intelligibility did not change as a function of memory set size. Although performance on the MS task remained unchanged, it appeared to sufficiently increase overall task loading as shown by the marked increase in SWAT scores when a second task was added. Thus it may be concluded that a dual task situation will have a minimal effect on the intelligibility of a pilot's message in noise given the MS task.

It is possible to interpret this data in light of Wickens (1984) multiple resource model. The basic premise of Wickens model is that interference occurs when tasks compete for the same processing resource. This interference can occur on input, central processing or output. Information presented in the same input modality has a greater tendency to cause interference in the cognitive stages of information processing than information presented in different input modalities. Since, in the present study, subjects received the MRT via the auditory modality while the MS task was visually presented, one would not necessarily expect task interference.

It seems that subjects were able to split attention between the two tasks and were not truly time-sharing their resources. In the dual task condition, successive MS characters were presented on the left CRT. When subjects heard the last part of the MRT carrier phrase (e.g., "Number one, you will mark _____, please") subjects would shift their attention, look to the right CRT and select the appropriate response word. After selecting this word subjects would shift their attention back to the MS task (i.e., continue to work on the same MS letter that was on the screen before the MRT was presented). Thus, "time-swapping" between the MRT and MS was probably occuring in Phases 2 and 3.

Since true time-sharing between tasks probably did not occur, loading on the MS task would not be expected to carry over to the MRT. A significant decrement in intelligibility might have occurred if the stimuli had been presented using the same input modality (e.g., if an auditory version of the MS was presented to subjects). A main effect of memory set size on intelligibility was anticipated but not found. If true time sharing was not occuring, manipulation of memory set size in the MS task would have a minimal effect on MRT performance.

Subjective workload (i.e., SWAT scores) did increase but this increase was not reflected in primary task performance. O'Donnell and Eggemeier (1986) state that performance and workload are not always correlated. This is often the case if the primary task is not sensitive to changes caused by the secondary task.

The present study replicates previous AAMRL studies in that intelligibility decreases with increasing noise. The single task MRT percent correct scores are in general agreement with previous AAMRL research cited earlier in this paper. In previous results with the AIC-25 intercom system, percent correct decreased from approximately 98 to

80 percent as noise increased from 95 to 115 dB. In the present study, the decrement was much less, from only 99 to 96 percent correct.

The smaller decrement in percent correct in the present study is probably due to the number of talkers (i.e., one) and the sex of the talker (i.e., male). Kirk et al. 1972 recommend five talkers when evaluating communication systems. In the previous AAMRL studies five talkers (three male and two female) presented the MRT to subjects. Also, intelligibility of a male talker is better than a female talker especially at higher levels of noise (Moore et al. 1980). This is thought to be due to the higher frequencies of female speech being masked by noise.

Of secondary interest was the effect that adding a second task had on MRT response time. Response time has typically been used as a alternate measure of system performance in the previous AAMRL studies when differences in intelligibility were not found. In the present study, MRT response time significantly increased by half a second when a second task was added. This increase in response time would be expected when a second task is added. But unlike intelligibility results, when a second task was added, the level of noise differentially affected response time.

This noise by task interaction reflects the fact that MRT response time significantly decreased as a function of noise in Phase 3 while it remained unchanged in Phase 1. Phase 3 response time for the 95 dB noise condition was slower than the 105 and 115 dB noise conditions. Phase 3 data showed a decrease in response time coupled with an increase in errors while in Phase 1 there was no change in response time with an increase in errors. The lack of significant decrease in response time in Phase 1 may have been the result of a floor effect. The subjects had many hours of prior practice on the MRT and therefore had learned to respond very rapidly to this test. It seems that prior training helped to establish a baseline response time so that subjects could not respond faster as noise increased. In light of the noise literature, the interpretation of these results is unclear.

Hockey (1986), in his review of environmental stressors on performance, concludes that noise can sometimes result in faster response time, cause an increase in errors or both. Gawron (1980), in her review of the noise literature, argues that these variations among performance measures are largely a function of differences in experimental method. She cites variations in levels of noise intensity, noise characteristics, length of noise exposure, type of task and choice of dependent variable as possible reasons for the differing results typically found in the noise literature. Thus, the present results are consistent with previous observations but their interpretation is still uncertain.

Although differences in response time for Phase 3 were significant, they were very small (i.e., a 0.09 second difference exists between the the 95 dB and 115 dB noise conditions). From an application point of view, this differential effect is very small and has little practical implication for the pilot in the cockpit. It can be concluded that only when response time is a critical factor in the cockpit would the half second increase for a second task or the differential effect of noise across tasks have practical significance.

The second objective of this study was to examine the effect of adding noise to the dual task communication paradigm. Adding noise caused a decrement in intelligibility but did not significantly increase response time or perceived workload. In light of previous AAMRL research, this decrease in intelligibility would be expected. These results show that response time remained unchanged which is consistent with Hockey's (1986) previous comment that effect of noise can be seen in speed, accuracy or both.

It appears that for well-trained subjects noise has a minimal effect on subjective ratings of workload. SWAT scores for the single task in noise show that the absolute level was very low and that scores increased very little as a function of noise. This same effect is found when noise is added to the dual task. One explanation would be that this is a function of training. Subjects were well practiced and had many hours of training in noise. One might see different ratings if subjects had minimal training in noise.

MS performance was unchanged as a function of noise. This again can be explained using Wickens (1984) resource model. As discussed earlier, performance decrements are more likely if tasks compete for the same resources. In this case, there is minimal competition for resources due to differing input modalities (i.e., noise is an auditory input while the MS is a visual input). In light of this, one would not expect noise to interfere with MS performance, particularly for subjects accustomed to performing in noisy environments.

The increase in MS response time and decrease in MS percent correct as a function of memory set size was expected and is consistent with previous research

(Shingledecker, 1984). The increase in SWAT scores as a function of memory set size was also expected since more effort was required to search and compare six characters in short term memory than was required for one character.

To conclude, the results show that adding a second task (i.e., Sternberg's (1969) memory task) to the standard noise-communication paradigm will not affect a pilot's ability to communicate in the cockpit. Further, for most cockpit scenarios, the half-second increase in response time for the communication task should not interfere with a pilot's overall performance in the cockpit. Thus the current AAMRL research strategy, utilizing a single communication task in noise, appears to be an adequate paradigm for intelligibility testing.

APPENDICES

APPENDIX A

Voice Communication Research and Evaluation System (VOCRES)

(Adapted from McKinley, 1980)

Introduction

Air and ground crew voice communication may be degraded by a variety of system and environmental factors that include electrical or acoustical noise or both, radio interference, jamming, communication signal processing and various other factors that prohibit effective communication. The Voice Communication Research and Evaluation System (VOCRES) (see Figure A-1) has been developed to provide the capability to research, test and evaluate voice communications effectiveness.

The general approach when using VOCRES involves the participation of volunteers who communicate as talkers and listeners under controlled conditions that replicate the specific communication environments being evaluated. Subjects are stationed at custom-designed consoles and communicate with standardized or special purpose (speech) vocabulary materials. Various system and environmental characteristics or equipment are varied and the resulting communication effect. eness is quantified. Data derived

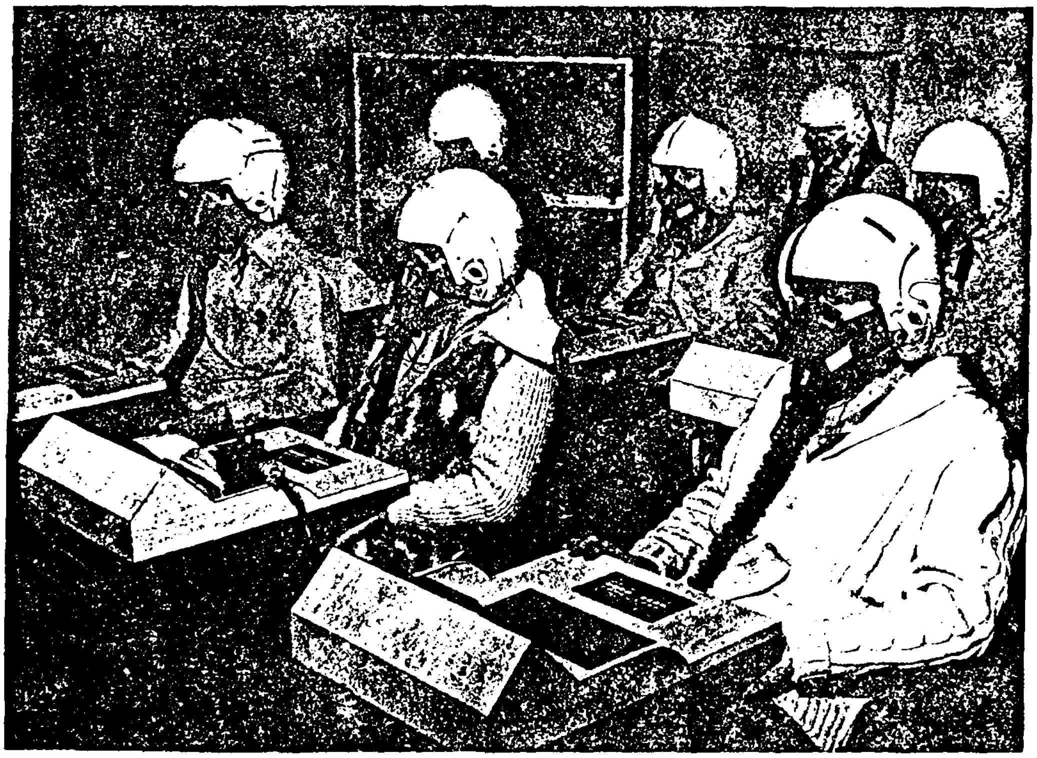

Figure A-1. The VOCRES system. (Source: McKinley, 1980) from these efforts may be used to establish baseline communication system performance profiles. These profiles are then used for comparative testing of specific communication system components, such as radios, intercoms, microphones, earphones and voice processors.

VOCRES: General System

The VOCRES system is an aggregate of four different subsystems integrated into a voice communication network that includes ten individual communication stations and one control station. The individual communication consoles are located in a large reverberation chamber and the master console is located in a control room adjacent to the chamber.

The subsystems include (1) an AIC-25 aircraft intercommunication system (10 stations), (2) an air respiration system with A-19 diluter-demand regulators for use with standard oxygen masks, (3) a high intensity sound source for duplicating operational acoustical environments occupied by crew members and (4) a central processing unit that controls all stations and conducts the individual testing sessions and conditions, i.e., presents materials, monitors participant activity, records, stores and analyzes responses, and provides analyzed data in tabular or graphic form or both. The overall system is adaptable to the incorporation of various aircraft radios, communication jammers, and the like, that are not integral components of VOCRES. Each of the ten communication consoles or stations is equipped with an AIC-25 intercommunication terminal, an A-19 respiration terminal, an A-19 respiration terminal, a display/subject reponse unit, a keypad for subjects communication task response and a large volume unit (VU) meter that indicates the voice level of communications generated at that station (see Figure A-3). The system can be operated with any number of one to ten subjects. The experimental design used most often is a "round robin" procedure where each subject, in turn, performs as talker while the remaining subjects respond as listeners.

Subjects repond using one of two reponse systems shown in Figure A-2. The first system consists of six push buttons, three on either side of the CRT display each with a red light emitting diode (LED) mounted in the bezel. These LEDs provide feedback to the subject indicating their chosen response. Pressing one button illuminates the adjacent LED indicating to the subject that he has selected that button. If a subject changes his resonse (i.e., presses another button) the adjacent LED for the new button is illuminated while the first LED

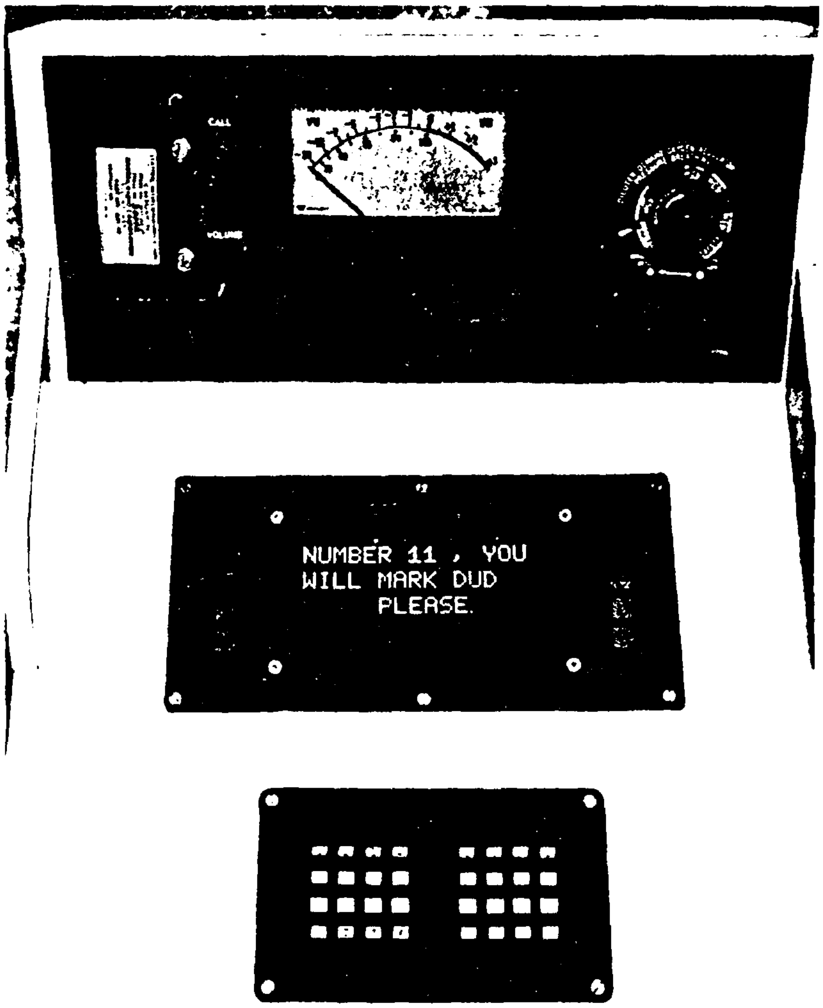

Figure A-2. Individual VOCRES station (Source: McKinley, 1980) is turned off. The second response system consists of two 4x4 calculator type keypads. Only one of the 32 buttons can be chosen at any one time. Operation is similar to the six LED pushbuttons except that pressing one of the 4x4 keys can illuminate one to five of the six LEDs which forms a specific pattern for that key.

Communication materials

Communication materials consist of the standardized Modified Rhyme Test (MRT) developed by House, et al. 1965. Other test materials such as the Diagnostic Rhyme Test (DRT) developed by Voiers, 1967 are used from time to time for special purpose applications.

Communcation link capabilities

The communications assemblage diagram shown in Figure A-3 demonstrates the high flexibilty of VOCRES that allows a variety of different communication links to be examined either individually or in combination with one another. The range of communication links can be varied from a simple face-to-face communications situation (i.e., direct talker to listener) to a complex configuration using encoders, encrypters and the like by varying appopriate subunit controls. Any of the alternate pathways shown in Figure A-3 can be used to complete the talker to listener link. The direct talker to listener path theoretically

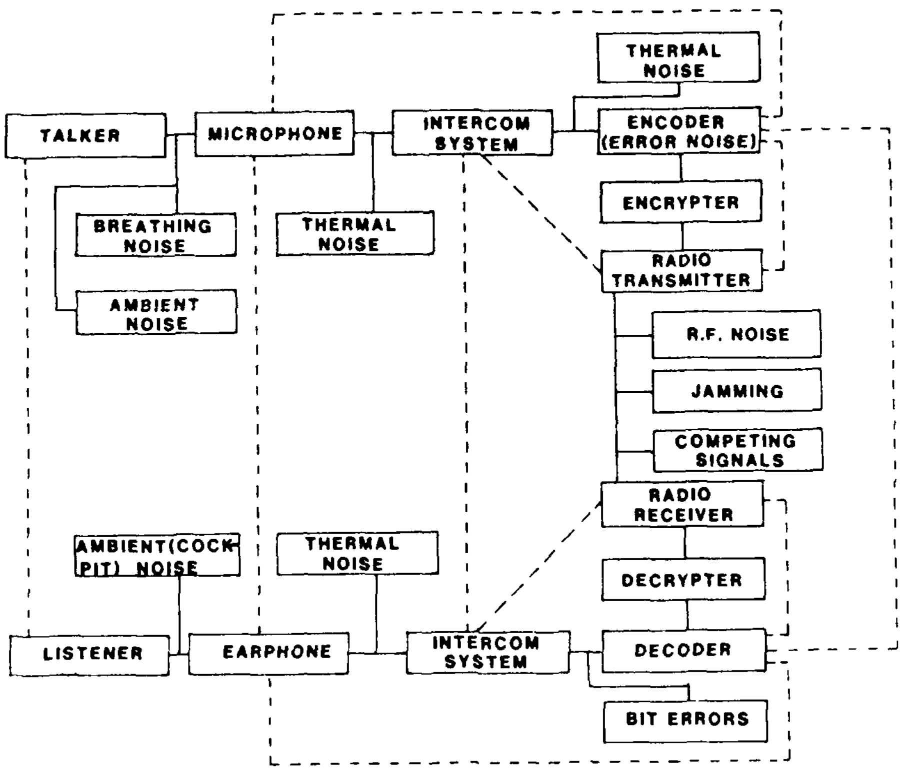

Figure A-3. Block diagram of the VOCRES communication link (Source: McKinley, 1980) provides a data baseline free from environmental and component effects.

High intensity sound system

The high intensity sound system is shown in Figure A-4. The system is capable of operation in a high power mode (14,000 watts) and a low power mode (12,000 watts). The power ampifiers drive eight banks of loudspeakers containing a total of 96 Altec 15" low frequency speakers, eight Altec horn loaded compression drivers and 384 Stromberg Carlson high frequency speakers. The noise generator and the spectrum shaper allow almost any desired noise environment (spectrum) within the human audio-frequency range to be generated inside the test chamber. This permits the accurate reproduction of ambient and environmental noise conditions of specific operational situations within the laboratory, which is a vital aspect of the validity of the communcation testing.

The room in which the loudspeaker banks are located is a specially designed and constructed acoustic reverberation chamber. The room is designed for maximum reverberation time and has a volume of approximately 8,000 cubic feet. The irregular wall surfaces are designed to disrupt the formation of standing waves and maximize the uniformity of the level of a noise distributed throughout the room.

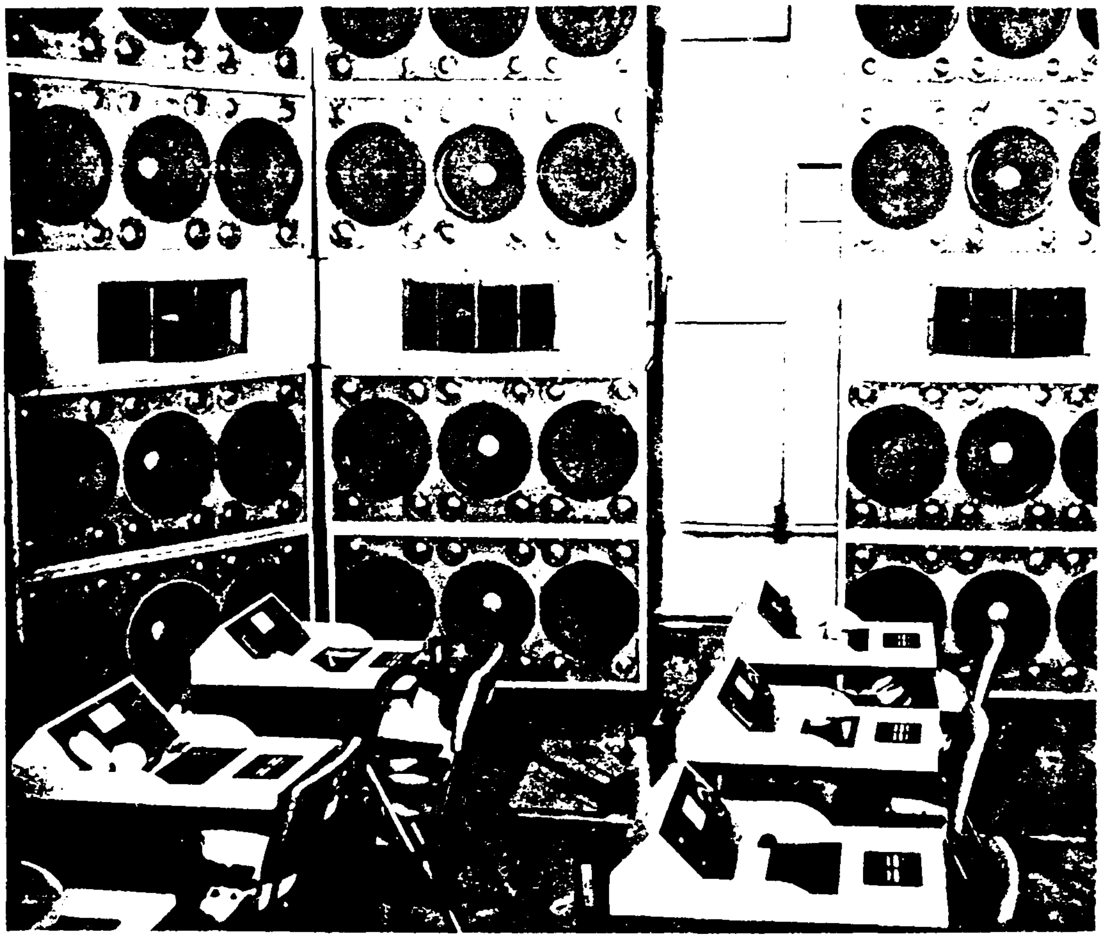

Figure A-4. High intensity sound system (Source: McKinley, 1980)

AIC-25 Intercommunication system

A standard AIC-25 standard aircraft intercommunication (intercom) system has been installed in each individual VOCRES station. A talker's intercom can be connected to the audio input of any transmitter using the control console switching circuit. The audio output is then routed to the other nine listeners. Some of the terminal equipment used with the AIC-25 intercom system include standard H-157A headsets, H-133 headsets, MBU-5/P oxygen masks and HGU-26/P flight helmets which are shown in Figure A-5.

Air Respiration System

The air breathing system uses the standard Air Force A-19 diluter demand regulator as its primary component. Each individual VOCRES station has its own A-19 regulator which is supplied through feeder lines by a semiautomatic regulator manifold. The manifold connects six standard size breathing air bottles to the system through two regulators (i.e. each regulator controls three bottles). When the supply of the first three bottles is exhausted the system auotmatically switches to the second set of bottles. The normal operating pressure in the system is 150 pounds per square inch (psi).

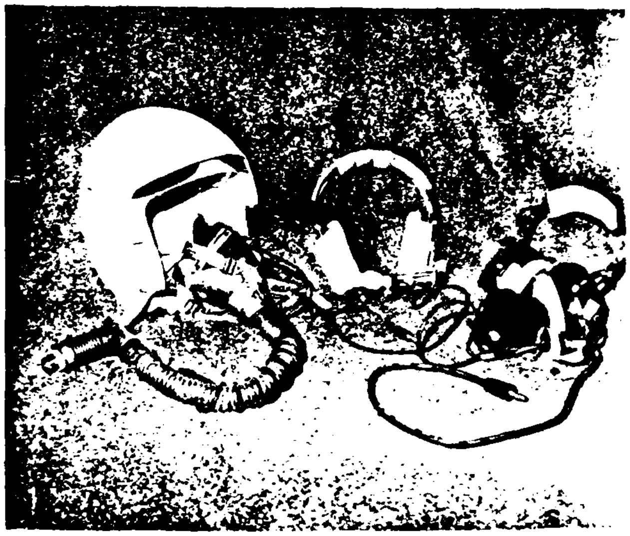

Figure A-5. AIC-25 Compatible terminal headgear (Source: McKinley, 1980)

APPENDIX B

Subjective Workload Assessment Technique (SWAT)

SWAT was generated by Reid, Shingledecker, and Eggemeier (1981) using a psychometric technique known as conjoint measurement. Reid et al. (1981) give a succinct description of the conjoint measurement technique:

> In conjoint measurement the joint effects of several factors are investigated and the rule or composition principle that relates the factors to one another is extracted from the data. One major advantage of conjoint measurement and other related scaling procedures is that only the ordinal aspects of the data are required for the production of interval level data. Other advantages include ease of administration and unobtrusiveness.

Reid et al. have defined the following three major dimensions of workload in their creation of SWAT.

1. Time Load - the function of the total time completing a task that you consider yourself busy.

2. Mental Effort Load - an index of the amount of attention or mental effort required by a task regardless of the number of tasks to be performed or any time limitation.

3. Stress Load - the contribution to total workload of any conditions that produce anxiety, frustration or confusion while performing a task or tasks.

Each of these three dimensions can be further subdivided into three levels as shown in Table B-1. Reid et al. remind us that "The primary assumption of SWAT is that subjective workload can be adequately represented by the combination of these three dimensions." Figure B-1 is a three-dimensional respresentation of SWAT.

Use of the SWAT rating scales begins with the Scale Development phase. In it subjects are required to rank order the 27 possible levels of workload (presented on indexed cards) on the "basis of their general experience and not on the basis of any particular task." Subjects are given as much time as they need to rank order 27 index cards. Once this is accomplished, each individual's rank order is used in conjunction with the conjoint measurement technique described earlier to develop a workload scale. Finally when data are collected, they are compared to this workload scale to determine the subjects perception of workload.

Table B-1

Dimensions of SWAT (Source: Reid et al. 1981)

TIME LOAD

1. Often have spare time, interruptions or overlap among activities occur infrequently or not at all.

2. Occasionally have spare time. Interruptions or overlap among activities occur frequently.

3. Almost never have free time. Interruptions or overlap among activities are very frequent, or occur all the time.

MENTAL EFFORT LOAD

1. Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

2. Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

3. Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

STRESS LOAD

1. Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

2. Moderate stress due to confusion, frustration, or anxiety noticeably adds to workload. Significant compensation is required to maintain adequate performance.

3. High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

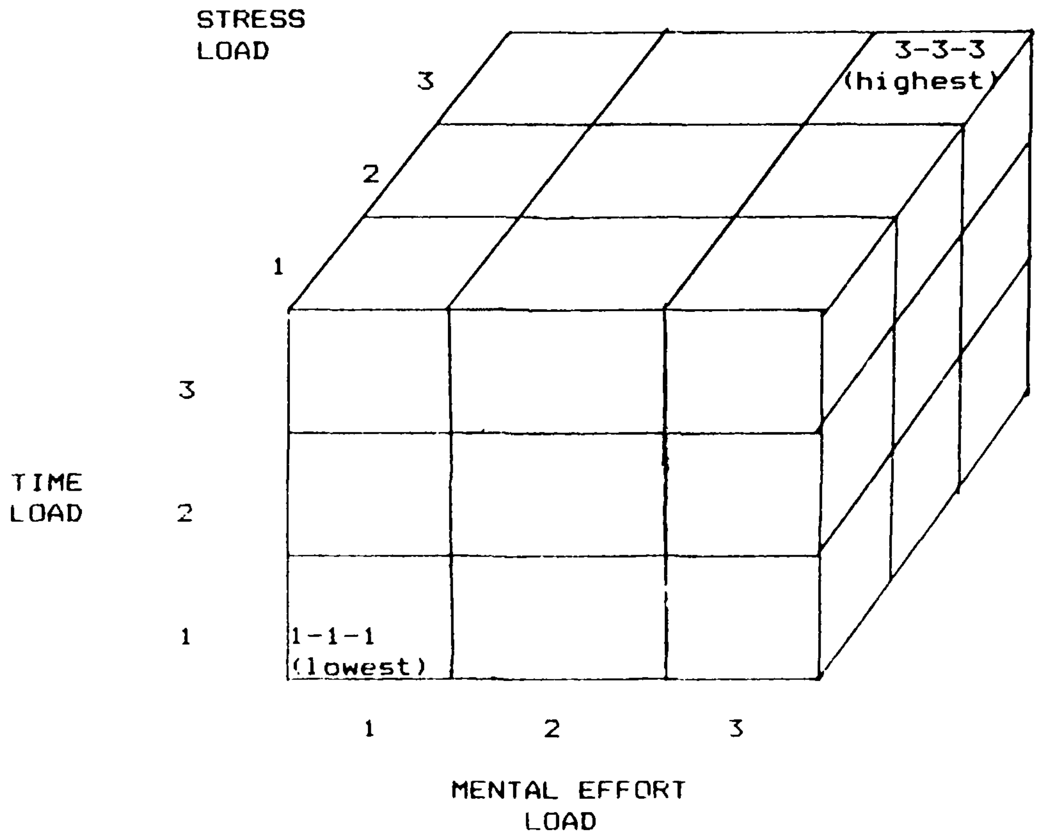

Figure B-1. SWAT workload dimensions. (Source: Reid et al. 1981)

APPENDIX C

Memory Search (MS) Software and Hardware

(Adapted from Acton and Crabtree, 1984)

The Sternberg memory search (MS) task can be implemented on a commercially available microcomputer system with a minimum of additional custom-built hardware. The following equipment are required for system operation: 1) a Commodore 64 microcomputer; 2) a Commodore 1541 disk drive; 3) a monochrome monitor or equivalent (with 75 ohm loop-through and female BNC video input connector) and 4) a custom-built response keypad and cable.

In the present study, this equipment was configured as part of the PACRAT system and the MS task was displayed on one of the five inch (5") diagonal cathode ray tubes (CRT's). The keyboard on the microcomputer was used by the experimenter for data input while the disk drive was used to load the MS software into the microcomputer and to store subjects data.

Custom Hardware - Response Kepad Desrcription

A four button keypad was designed to be compatible with the spatial layout of the criterion task set (CTS, Shingledecker, 1984) Probability Monitoring task and isused for the MS task. Binary choice responses are made using two of the four keys (left and right keys) on the four button keypad. The keypad contains four single pull-single throw (SPST) push button switches which are normally open. The switches should have low activation force requirements, provide at least a small amount of tactile and auditory feedback upon closure, require little depth for mounting, have minimal travel (i.e., a short "throw") and be highly reliable.

The principal of operation is that each switch is connected to ground (through a 1000 ohm current limiting resistor) and a bit line at the Commodore 64's user I/O port. Depressing the key causes the corresponding bit line to go "low" (i.e., the line is switched from a nominal 5 volts to ground). The four switches are numbered one through four and are connected to bit lines PB1, PB2, PB3, and PB4, respectively. The CTS task software is written to sense changes in bit values caused by key presses.

Software description

The software for the MS is written primarily in BASIC to run on the Commodore 64 computer. The program was compiled to improve execution speed and efficiency. The reaction time measure is recorded in milliseconds with a resolution of + 1.5 milliseconds.

The MS software is structured to minimize experimenter familiarization and training requirements. Standardized, self-explanatory menus are used for all tasks to simplify trial preparation and data handling activities. Once task software is loaded into the computer, initial menus permit the experimenter to select training or test conditions and specific loading levels on the task.

Options are also provided to test the response device for the task, to analyze previously stored raw data, and to display correct responses along with each stimulus presentation when required for training. Furthermore, explicit prompts are given to sequence the user through the menus and to ensure accurate insertion of subject and test condition identifiers. Following data collection, additional menus allow the experimenter to examine the new data in a "quick look" mode; calculate summary statistics; and store or print a detailed, time-based record of all stimuli presented and subject responses.

Tests of the CTS hardware and software under actual experimental data collection conditions have indicated that the combined system is highly reliable. No hardware failures were experienced during approximately 2500 three minute test trials run over a two week period. In addition, experimenter error was minimized by the user-friendly software design which limited cases of irretrievable data loss to 0.2 percent of all test trials.

APPENDIX D

Frequency Spectra at Aircraft Noise Levels

The frequency spectra for noise levels of 95, 105 and 115 dB OASPL were presented to all three listening stations in all training sessions and all three experimental phases. These spectra were recorded using a Hewlett Packard spectrum analyzer (HP-9845A). Each noise spectra is represented by a unique Figure and Table combination. Figure D-1 shows the ambient noise spectra while Table D-1 shows the individual sound pressure levels for ambient noise. Although ambient noise conditions were not used in this study, this ambient data provides a baseline for comparison with the other noise conditions. Figure D-2 and Table D-2 show the noise spectra and sound pressure levels at 95 dB overall sound pressure level (OASPL), Figure D-3 and Table D-3 show the noise spectra and sound pressure levels at 105 dB OASPL and Figure D-4 and Table D-4 show the noise spectra and sound pressure levels at 115 dB OASPL.

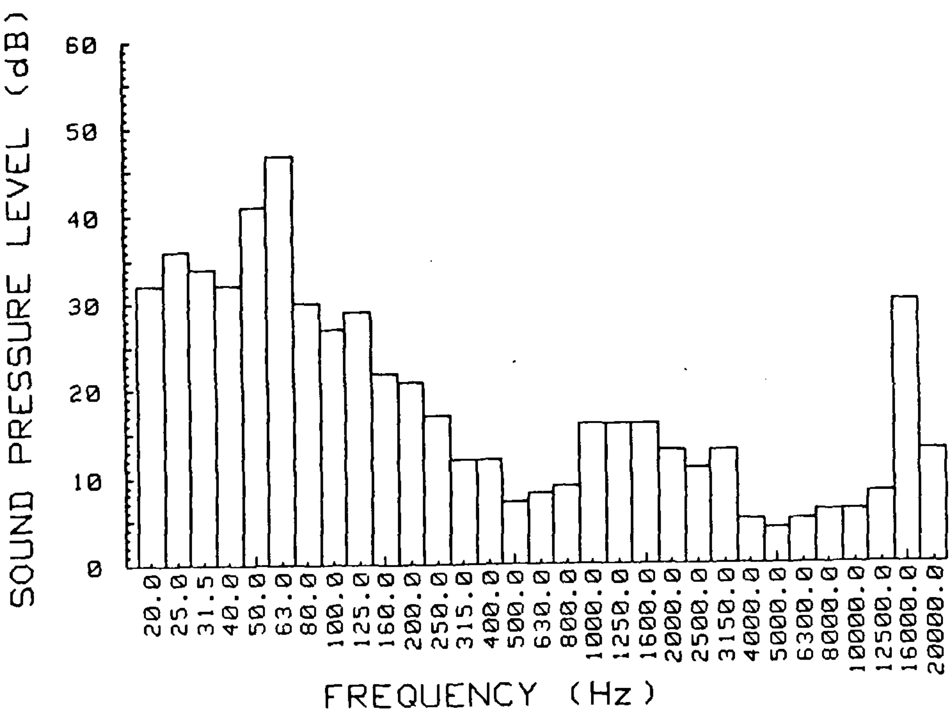

Figure D-1. Frequency spectra for ambient noise at 75 dB overall sound pressure level (OASPL).

Table D-1

Individual sound pressure levels at 75 dB OASPL

Frequency (Hz)	Level (dB)
20.0	032*4
25.0	036*0
31.5	034*7
40.0	032*6
50.0	041*6
63.0	047*2
80.0	030*6
100.0	027*3
125.0	029*0
160.0	022*0
200.0	021*5
250.0	017*3
315.0	012*7
400.0	012*7
500.0	007*9
630.0	008*1
800.0	009*3
1000.0	016*8
1250.0	016*9
1600.0	016*9
2000.0	013*9
2500.0	011*6
3150.0	013*0
4000.0	005*1
5000.0	004*8
6300.0	005*4
8000.0	006*7
10000.0	006*0
12500.0	008*6
16000.0	030*5
20000.0	013*5

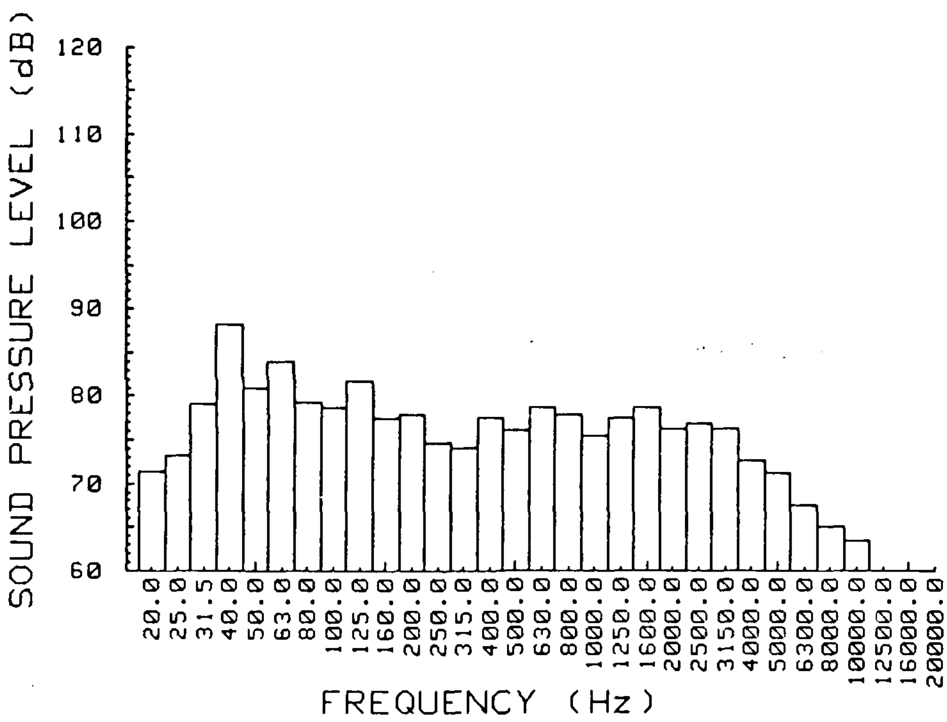

Figure D-2. Frequency spectra for unweighted pink noise at 95 dB overall sound pressure level (OASPL).

Table D-2

Individual sound pressure levels at 95 0ASPL

Frequency (Hz)	Level (dB)
20.0	071.3
25.0	073.2
31.5	079.0
40.0	088.2
50.0	080.9
63.0	083.9
80.0	079.3
100.0	078.7
125.0	081.7
160.0	077.4
200.0	077.9
250.0	074.7
315.0	073.9
400.0	077.5
500.0	076.0
630.0	078.7
800.0	077.8
1000.0	075.4
1250.0	077.4
1600.0	078.6
2000.0	076.3
2500.0	076.8
3150.0	076.3
4000.0	072.6
5000.0	071.1
6300.0	067.4
8000.0	065.1
10000.0	063.5
12500.0	060.0
16000.0	060<0
20000.0	060<0

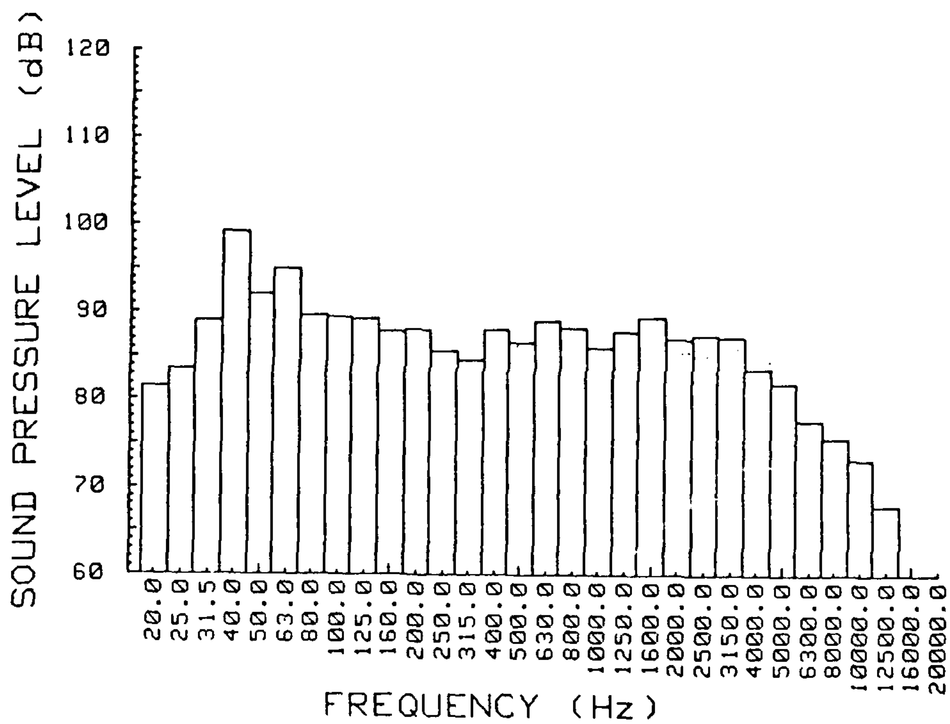

Figure D-3. Frquency spectra for unweighted pink noise at 105 dB overall sound pressure level (OASPL).

Table D-3

Individual sound pressure levels at 105 dB OASPL

Frequency (Hz)	Level (dB)
20.0	081.5
25.0	083.6
31.5	089.1
40.0	099.1
50.0	092.0
63.0	094.9
80.0	089.7
100.0	089.4
125.0	089.2
160.0	087.9
200.0	088.0
250.0	085.5
315.0	084.5
400.0	088.1
500.0	086.6
630.0	089.1
800.0	088.2
1000.0	086.0
1250.0	087.9
1600.0	089.5
2000.0	087.0
2500.0	087.5
3150.0	087.3
4000.0	083.5
5000.0	082.0
6300.0	077.7
8000.0	075.6
10000.0	073.2
12500.0	067.9
16000.0	060<0
20000.0	060<0

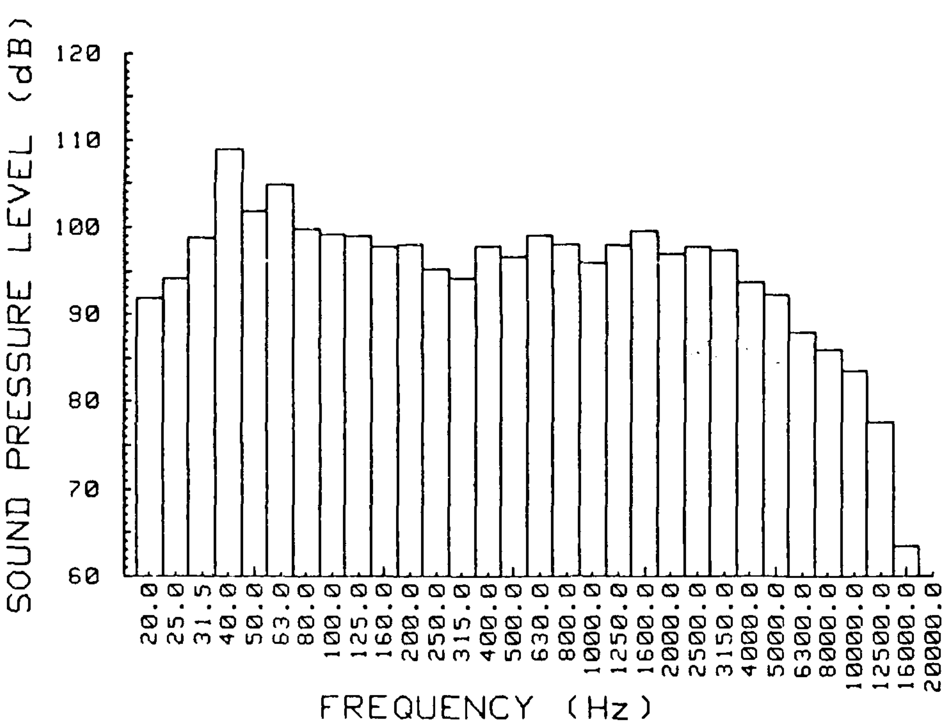

Figure D-4. Frquency spectra for unweighted pink noise at 115 dB overall sound pressure level (OASPL).

Table D-4

Individual sound pressure levels at 115 dB OASPL

Frequency (Hz)	Level (dB)
20.0	091.9
25.0	094.2
31.5	098.7
40.0	109.0
50.0	101.8
63.0	104.9
80.0	099.9
100.0	099.3
125.0	099.0
160.0	097.7
200.0	098.0
250.0	095.1
315.0	094.1
400.0	097.8
500.0	096.6
630.0	098.9
800.0	098.0
1000.0	095.9
1250.0	097.9
1600.0	099.6
2000.0	097.0
2500.0	097.7
3150.0	097.4
4000.0	093.7
5000.0	092.2
6300.0	088.0
8000.0	085.9
10000.0	083.6
12500.0	077.7
16000.0	063.5
20000.0	060<0

APPENDIX E

SWAT Card Sort Instructions to Subjects

(Source: Reid, 1985)

The workload imposed on a person at any one time can be thought of as the combination of the particular level of Time Load, Mental Effort Load and Stress Load present in a task or combination of tasks. The various possible combinations of the three dimensions can be represented by the cube in Figure E-1. Each cell represents one possible combination of Time Load, Mental Effort Load and Stress Load.

In order to use a new subjective scale to rate the workload associated with a particular job or task, we need information from you regarding the amount of workload imposed by various combinations of the dimensions illustrated in the cube shown in Figure E-1. We can get the necessary information on workload by having you rank order the workload associated with each possible combination.

In order to rank order the workload associated with each of the combinations, a set of 27 cards with the combination from each of the dimensions is provided. Each card contains a different combination of possible levels of Time Load, Mental Effort Load and Stress Load. Your job is to sort the cards so that they are rank ordered according to the level of workload represented on each.

In completing your card sort, please consider the workload imposed on a person by the combination represented on each card, and arrange the cards from the lowest workload condition through the highest workload condition. You may use any strategy that you choose in rank ordering the cards. One strategy that has proven useful to others is to first arrange the cards into a preliminary number of stacks representing "Low", "Moderate" and "High" workload. Individual cards can be exchanged between stacks, if necessary, and then rank ordered within stacks. Stacks can then be recombined and checked to be sure that they represent your ranking of lowest to highest workload. However, the choice of

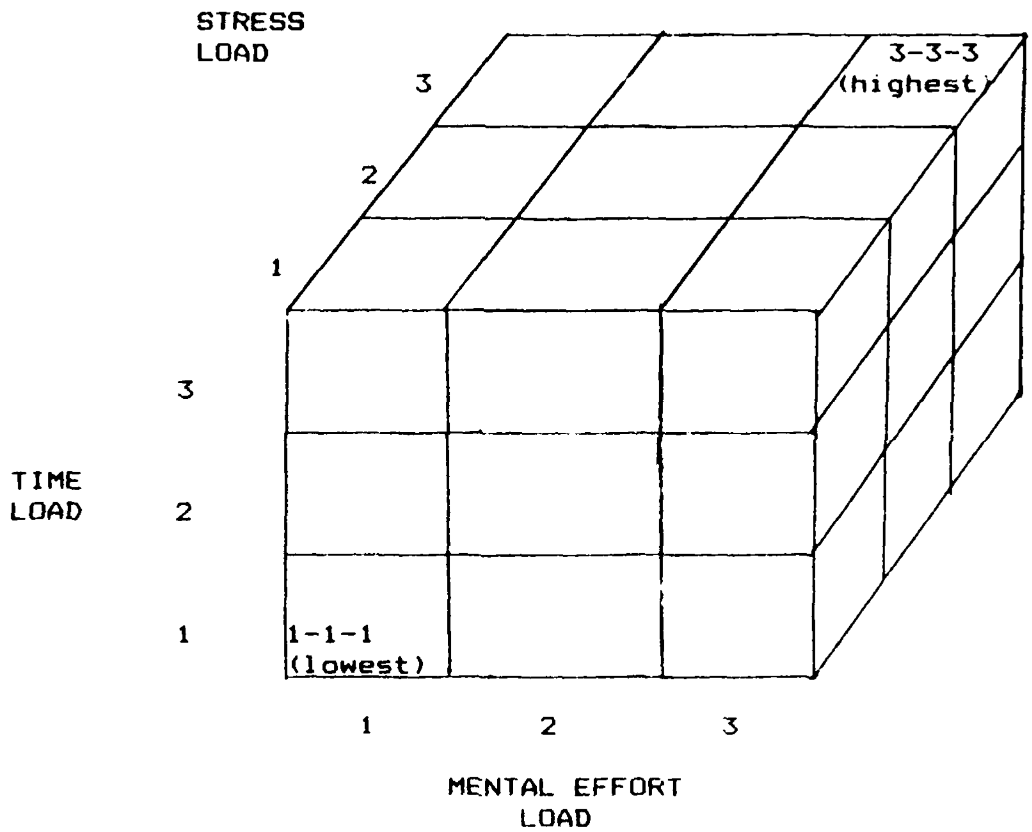

Figure E-1. SWAT workload dimensions. (Source: Reid et al. 1981) strategy is up to you and you should choose the one that works best for you. Please feel free to ask questions at any time. Thank you for your cooperation.

TIME LOAD

Time load refers to the fraction of the total time that you are busy. When time load is low, sufficient time is available to complete all of your mental work with some time to spare. As time load increases, spare time drops out and some aspects of performance overlap and interrupt one another. This overlap and interruption can come from performing more than one task or from different aspects of performing the same task. At higher levels of time load, several aspects of performance often occur simultaneously, you are busy and interruptions are very frequent.

Time Load may be judged on the three-point scale shown below.

I. Time Load

1. Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

2. Occasionally have spare time. Interruptions or overlap among activities occur frequently.

3. Almost never have spare time. Interruptions or overlap among activities are very frequent or occur all the time.

MENTAL EFFORT LOAD

Time load refers to the amount of time one has available to perform a task or tasks. In contrast, mental effort load is an index of the amount of attention or mental effort required by a task regardless of the number of tasks to be performed or any time limitation. When mental effort load is low, the concentration and attention required by a task is minimal and performance is nearly automatic. As the demand for mental effort increases, the degree of concentration and attention required to perform increases, due to task complexity or the amount of information which must be dealt with in order to perform adequately. High mental effort load demands total attention or concentration due to task complexity or the amount of information to be dealt with.

Mental Effort Load may be judged on the three point scale shown below.

II. Mental Effort Load

1. Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

2. Moderate conscious mental effort or concentration required. Complexity of activity is moderately high in uncertainty, unpredictability or unfamiliarity. Considerable attention is required.

3. Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

STRESS LOAD

Stress load refers to the contribution to the total workload of any conditions that produce anxiety, frustration and confusion while performing a task or tasks. At low levels of psychological stress, one feels relatively relaxed. As stress increases, confusion, anxiety or frustration increase and greater concentration and determination are required to maintain control of the situation.

Stress Load may be judged on the three point scale shown below.

III. Stress Load

1. Little confusion, risk, frustration or anxiety exists and can be easily accommodated.

2. Moderate stress due to confusion, frustration or anxiety noticeably adds to workload. Significant compensation is required to maintain adequate performance.

3. High to very intense stress due to confusion, frustration or anxiety. High to extreme determination and self control required.

Each of the three dimensions just described contribute to workload during the performance of a task or group of tasks. Note that although all three factors may be correlated, they need not be. For example, one can have many tasks to perform in the time available (high time load) but the tasks may require little concentration (low mental effort load). Likewise, one can be anxious and frustrated (high stress load) and have plenty of spare time (low time load) between relatively simple tasks. Since the three dimensions contributing to workload are not necessarily correlated, please treat each dimension individually and give independent assessments of the time load, mental effort load and stress load that you experience when you perform the tasks in this study.

APPENDIX F

Modified Rhyme Test Instructions to Subjects

(adapted from House, et al. 1965)

You are going to hear some one syllable words presented in different loudness levels of aircraft noise. A carrier phrase which identifies the current item number and the target word will be presented in your headset. For example:

Number one, you will mark tree, please.

Number two, you will mark mile, please.

The word presented will be one of the six words appearing on the right CRT screen. Your task is to identify the word presented by pressing the button next to the word on the CRT screen. For example:

Number three, You will mark beat, please.

CRT Screen

	meat	seat	
	feat	beat	x
	heat	neat	

Some words will be easier to hear than others. If you are not sure what the word is -- guess. Always press one of the buttons for each item number presented.

Are there any questions?

APPENDIX H

SWAT Task Instructions to Subjects

(Source: Reid, 1985)

You will now rate workload based on your perceptions of the amount of workload imposed on you by the task(s). As you know from your card sort task, workload is composed of three dimensions: time load, mental effort load, and stress load. Each of the three workload dimensions contains three corresponding workload statements. These should look familiar to you as they are the same workload dimensions and statements that you became familiar with in the card sort.

Your responsibility will be to place a checkmark next to the statement, under each of the three major dimensions, which most accurately reflects your feelings about the amount of workload imposed on you during the task. Remember to consider each of the three dimensions separately when making your ratings. As soon as you finish the task, pick up the pencil, turn the rating sheet over, and rate the workload for that task. Please do not compare the workload for the current task with the workload for any of your previous tasks.

APPENDIX I

Dual Task Instructions to Subjects

You will now perform both the MRT and the MS task in the PACRAT console. As in previous training sessions please use your preferred hand to respond to the MRT and your non-preferred hand to respond to the MS task. Your performance on the MRT is of upmost importance. Please insure that you respond to this task. If it is difficult to respond to both tasks, please respond to the MRT first and then to the MS task second. If you are unable to respond to both tasks, please insure that you respond to the MRT task.

APPENDIX J

Order of Presentation Using a Greco Latin Square Design

Table J-1

Presentation order for Phase 1 (single task in noise)

S1-S2, S5-S7 - - - - - - - - - - - - - - - - - - -

N1 L4(A)	N2 L5(B)	N3 L6(C)
N2 L6(B)	N3 L4(C)	N1 L5(A)
N3 L5(C)	N1 L6(A)	N2 L4(B)

S3-S4, S8-S10 - - - - - - - - - - - - - - - - - -

N1 L1(A)	N2 L2(B)	N3 L3(C)
N2 L3(B)	N3 L1(C)	N1 L2(A)
N3 L2(C)	N1 L3(A)	N2 L1(B)

Table J-2

Presentation order for Phase 2 (dual task without noise)

S1-S2, S5-S7 - - - - - - - - - - - - - - - - - - -

W3 L6(C)	W2 L4(A)	W1 L5(B)
W2 L5(A)	W1 L6(B)	W3 L4(C)
W1 L4(B)	W3 L5(C)	W2 L6(A)

S3-S4, S8-S10 - - - - - - - - - - - - - - - - - -

W3 L3(C)	W2 L1(A)	W1 L2(B)
W2 L2(A)	W1 L3(B)	W3 L1(C)
W1 L1(B)	W3 L2(C)	W2 L3(A)

N - Noise level
W - Task loading level
L - List - where L4(A) = list 4 randomization A

Table J-3

Presentation order for Phase 3 (dual task in noise)

S1-S2 and S5-S7 -

W1 N1 L4(A)	W2 N2 L5(B)	W3 N3 L6(C)
W2 N3 L5(C)	W3 N1 L6(A)	W1 N2 L4(B)
W3 N2 L6(B)	W1 N3 L4(C)	W2 N1 L5(A)
W3 N3 L4(C)	W1 N1 L5(A)	W2 N2 L6(B)
W1 N2 L5(B)	W2 N3 L6(C)	W3 N1 L4(A)
W2 N1 L6(A)	W3 N2 L4(B)	W1 N3 L5(C)
W2 N2 L4(B)	W3 N3 L5(C)	W1 N1 L6(A)
W3 N1 L5(A)	W1 N2 L6(B)	W2 N3 L4(C)
W1 N3 L6(C)	W2 N1 L4(A)	W3 N2 L5(B)

S3-S4 and S8-S10 - - - - - - - - - - - - - - - - - - -

W1 N1 L1(A)	W2 N2 L2(B)	W3 N3 L3(C)
W2 N3 L2(C)	W3 N1 L3(A)	W1 N2 L1(B)
W3 N2 L3(B)	W1 N3 L1(C)	W2 N1 L2(A)
W3 N3 L1(C)	W1 N1 L2(A)	W2 N2 L3(B)
W1 N2 L2(B)	W2 N3 L3(C)	W3 N1 L1(A)
W2 N1 L3(A)	W3 N2 L1(B)	W1 N3 L2(C)
W2 N2 L1(B)	W3 N3 L2(C)	W1 N1 L3(A)
W3 N1 L2(A)	W1 N2 L3(B)	W2 N3 L1(C)
W1 N3 L3(C)	W2 N1 L1(A)	W3 N2 L2(B)

N - Noise level

W - Task loading level

L - List - where L4(A) = list 4 randomization A

References

Acoustical Society of America (1983). Proposed AMERICAN NATIONAL STANDARD method for measuring the intelligibility of speech communication systems (ANSI S3.2-198X). New York, NY.

Acton, W. H., & Crabtree, M. S. (1984). User's guide for the criterion task set (Draft Copy). Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory.

Air Force Regulation 161-35: Hazardous noise exposure. (1982). Washington, DC: Department of the Air Force.

Brown, F. G. (1983). Principles of education and psychological testing (3rd ed.). New York: Holt, Rinehart & Winston.

Courtright, J. F., & Kuperman, G. (1984). Use of SWAT in USAF system T&E. Proceedings of the Human Factors Society - 28th Annual Meeting, 700-703.

Eggemeier, F. T., McGhee, J. Z., & Reid, G. B. (1983). The effects of variations in task loading on subjective workload rating scales. Proceedings of the IEEE 1983 National Aerospace and Electronics Conference, 1099-1106.

Eggemeier, F. T., Melville, B. E., & Crabtree, M. S. (1984). The effect of intervening task performance on subjective workload ratings. Proceedings of the Human Factors Society - 28th Annual Meeting, 954-958.

Gawron, V. J. (1980). The effects of noise: A review. Proceedings of the Human Factors Society - 24th Annual Meeting, 233-237.

Hockey, G. R. (1986). Changes in operator efficiency as a function of environmental stress, fatigue, and circadian rhythms. In K. R. Boff, L. Kaufman & J. P. Thomas (Eds.), Handbook of perception and human performance: Vol. I. Cognitive processes and performance (pp. 44-1 - 44-49). New York: John Wiley and Sons.

House, A. S., Williams, C. E., Hecker, M. H. L., & Kryter, K. D. (1965). Articulation-testing methods: Consonantal differentiation with a closed-response set. The Journal of the Acoustical Society of America, 37(1), 158-166.

Johnston, M. E. (1975). The effect of a tracking task on speech intelligibility in noise. Farnborough, Hants, UK: Royal Aircraft Establishment. (DTIC No. AD-A014-516).

Kirk, B., Marks, S., & Sergeant, R. L. (1972). Variability among ten talkers of word intelligibility in noise (NSMRL-93). Groton, CT: Naval Submarine Medical Research Laboratory. (DTIC No. AD-46-104).

Knowles, W. B. (1963). Operator loading tasks. Human Factors, 5, 155-161.

McKinley, R. L. (1980). Voice communication research and evaluation system (AFAMRL-TR-0-25). Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory.

McKinley, R. L., & Carr, S. (1984). (Best fit - averaged mean square error for pink and white noise). Unpublished raw data.

McKinley, R. L., Nixon, C. W., & Moore, T. J. (1981). Voice communication capability of selected in flight headgear devices. Advisory Group for Aerospace Research and Development (AGARD) - Conference Proceedings No. 311: Aural Communication in Aviation, Seine, France.

Moore, T. J., McKinley, R. L., & Mortimer, V. D. (1979). Comparative intelligibility of standardized test materials processed by the ARC-164 and ARC-34 radio systems in the presence of simulated cockpit noise (AMRL-TR-9-108). Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory.

Moore, T. J., Nixon, C. W., & McKinley, R. L. (1980). Comparative intelligibility of speech materials processed by standard Air Force voice communication systems in the presence of simulated cockpit noise. Advisory Group for Aerospace Research & Development (AGARD) - Conference Proceedings No. 311: Aural Communication in Aviation, Seine, France, pp 14-1 - 14-5.

O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology. In K. R. Boff, L. Kaufman & J. P. Thomas (Eds.), Handbook of perception and human performance: Vol. 1. Cognitive processes and performance (pp. 42-1 - 42-49). New York: John Wiley and Sons.

Reid, G. B., Shingledecker, C. A., & Eggemeier, F. T. (1981). Application of conjoint measurement to workload scale development. Proceedings of the Human Factors Society - 25th Annual Meeting, 522-526.

Reid, G. B. (1985). (SWAT card sort instructions). Unpublished raw data.

Shingledecker, C. A. (1984). A task battery for applied human performance assessment research (AFAMRL-TR-4-071). Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory.

Sternberg, S. (1969). Memory scanning: Mental processes revealed by reaction time experiments. American Scientist, 57, 421-457.

Tukey, J. W. (1977). Exploratory data analysis. Reading, MA: Addison-Wesley.

Voiers, W. D. (1967). Performance evaluation of speech processing devices III: Diagnostic evaluation of speech intelligibility (AFCRL-7-0101). L. G. Hanscom AFB, MA: Air Force Cambridge Research Laboratories.

Wickens, C. (1984). Engineering psychology and human performance. Columbus, OH: Charles E. Merrill.

Webster, J. C. (1979). Effects of noise on speech. In C. M. Harris (Ed.), Handbook of noise control (2nd ed.) (pp. 14-1 - 14-9). New York: McGraw Hill.

Webster, J. C., & Allen, C. R. (1972). Speech intelligibility in naval aircraft radios (ONR Task Number 213-089). San Diego, CA: Naval Electronics Laboratory Center.