ADA070275

NAMRL - 1259 LEVEL

HUMAN PERFORMANCE IN TIME-SHARED VERBAL

AND TRACKING TASKS

Steven D. Harris, Jerry M. Owens, and Robert A. North

See 1473 in back.

APRIL 1979

Approved for public release; distribution unlimited.

HUMAN PERFORMANCE IN TIME-SHARED VERBAL AND TRACKING TASKS

Steven D. Harris, Jerry M. Owens, and Robert A. North

Naval Medical Research and Development Command
ZF51.524.004-9017

Approved by	Released by
Ashton Graybiel, M. D. Assistant for Scientific Programs	Captain R. E. Mitchel, MC, USN Commanding Officer

April 1979

Naval Aerospace Medical Research Laboratory
Naval Air Station
Pensacola, Florida 32508

SUMMARY PAGE

THE PROBLEM

Significant progress has been made in the development of automated speech understanding systems for application to naval aviation systems. One advantage that is anticipated for speech over conventional man-machine interfaces is that speech could function as an independent channel for the control of systems. The experiment reported in this paper represents a preliminary investigation of the assumption that an automatic speech synthesizing and recognition system can provide the human operator an additional and parallel channel for processing information and effecting control responses.

The experiment required human subjects to timeshare a digital information processing task and a continuous compensatory tracking task. Independent variables in the design were task loading (single- vs. dual-task conditions), stimulus presentation modality for the digital task (auditory vs. visual), and response modality for the digital task (voice vs. keyboard). Data from 16 subjects were analyzed.

FINDINGS

The results indicated that the combination of visual stimulus modality and voice response provided optimum joint-task performance. No combination of stimulus and response modalities resulted in equivalent single- and dual-task performance. Future experiments should be designed to investigate the joint-task performance space for tasks that are more representative of the information processing performance requirements of specific systems. However, the interpretability of the results of such research will depend upon the solution of methodological problems, such as how to control or account for subjects' speed-accuracy tradeoff strategies and the priorities they place upon the concurrent tasks.

INTRODUCTION

Recent advances in artificial intelligence technology have resulted in commercially available computer systems that are able to synthesize auditory messages and to recognize spoken words and phrases in near-real-time with a high level of reliability. Several researchers have enumerated the benefits that are expected to accrue from computer recognition of speech (E 9). Beek, Neuberg, and Hodge (1) summarized the possible applications of this new technology to military systems. Curran (4), and Coler, Plummer, Huff, and Hitchcock (3) reported significant progress in the development of automated speech understanding systems for the control of on-board systems in military aircraft.

Among the anticipated advantages for automatic speech recognition over conventional techniques for effecting man-machine communication is that speech should function as an additional, independent channel for the control of systems (12). It is assumed that even though the operator's eyes and hands may be heavily occupied in the performance of a task, in many instances he would have adequate residual processing capacity to perform another task if the information to be processed could be presented aurally and responses could be made vocally. The results of recent research are equivocal on this point (5). The most frequent finding has been that the performance of one or both tasks will be degraded when a visual/manual task is performed concurrently with an auditory/vocal task. The experiment reported in this paper represents an initial effort to investigate the assumption that an automatic speech synthesis and recognition system can provide the human operator an additional, parallel channel for processing information and affecting control responses. A central focus of the present study was to determine not only if performance capabilities will be enhanced, but also the nature and extent of combined-task performance tradeoffs when audition and vocalization are used as alternatives to visual input and manual output modalities.

In two recent theoretical papers Norman and Bobrow (10, 11) introduced the concept of a system of limited processing resources to account for the limits in human information processing. Examples of resources are effort, memory capacity, and information channels. When two concurrent processes require access to the same resource, that resource must be allocated between them. Performance of one or both of the competing processes will deteriorate when the amount of the resource required by both processes exceeds the limit available to the system. To examine the tradeoff that occurs when two tasks are performed concurrently, Norman and Bobrow proposed the use of a performance operating characteristics (POC). The POC is a plot of performance on one task as a function of conjoint performance on another task, and is generated by varying resource allocation between two time-shared tasks. As Norman and Bobrow pointed out, the interpretation of a POC depends upon the assumption of complete complementarity of processing resources required for the competing tasks. Navon and Gopher (7) noted that complementarity is only one of several ways in which resources may be shared between two complex tasks. Navon and Gopher (7) noted that complementarity is only one of several ways in which resources may be shared between two complex tasks. Navon and Gopher showed that interpretation of an empirical POC requires considerable knowledge of the specific resources required for each of the competing tasks.

To summarize the arguments of Norman and Bobrow, Navon and Gopher, and others (see Kantowitz and Knight, 6), two parameters must be considered when tasks are performed concurrently: the relative priorities between the tasks, and the specific resources required by the tasks. In the present experiment relative priorities between two time-shared tasks were held constant, and the input and output (I/O) channels for one of the tasks were varied. The two tasks chosen for the experiment included a continuous compensatory tracking task and a digital information processing task. The independent variables were: task loading (single- vs. dual-task conditions); stimulus presentation mode for the dig task (visual vs. auditory); and response mode for the digit task (vocal or manual).

PROCEDURE

SUBJECTS

Twenty male naval officers and civilian staff members participated as subjects in the experiment. All subjects were right-handed and were between the ages of 22 and 35 years.

EXPERIMENTAL DESIGN

Subjects were tested in single- and dual-task performance of both a one-dimensional compensatory tracking task and a continuous absolute difference digit-processing task. As mentioned above, the three independent variables were task loading and stimulus and response modalities in the absolute difference task. Figure 1 shows the eight experimental conditions in the design. The stimulus presentation modality for the absolute difference task represented a between-subject variable with ten subjects serving in each condition. Task loading and response modality were within-subject variables. The various experimental trials are presented in Table I. The order of trials 1T1, 1T2, 1V, and 1K were counterbalanced across the ten subjects, as were trials 2V and 2K. Subjects were tested in the same conditions on each of the two successive days.

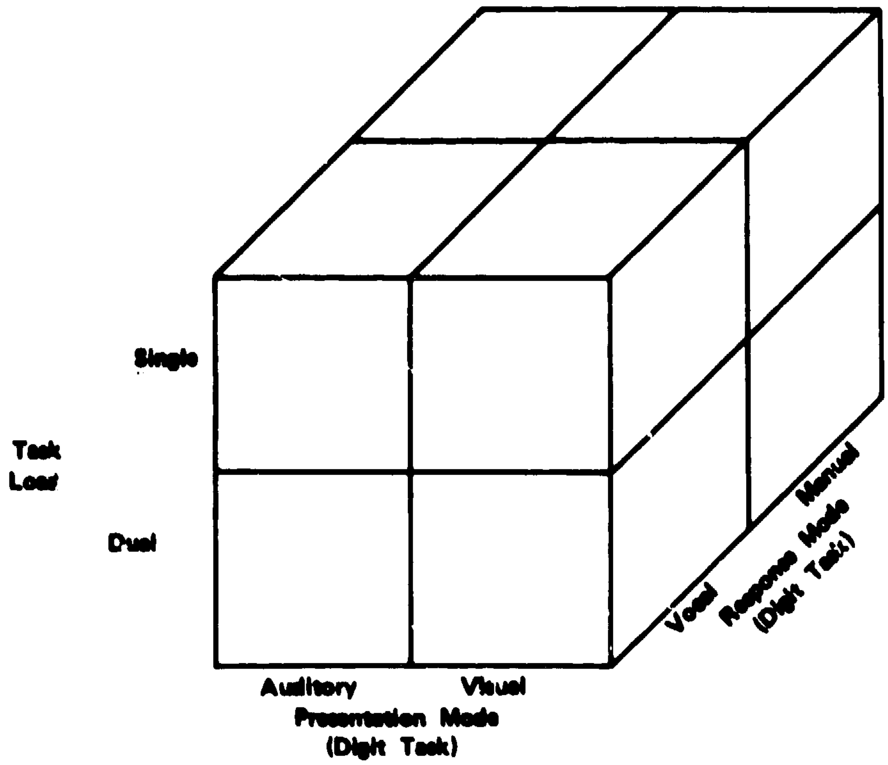

Figure 1. Experimental Design

Table I

Experimental Trial Sequence

TRIAL	DESCRIPTION
1T1	Single-Task Tracking (3 minutes)
1T2	Single-Task Tracking (3 minutes)
1V	Single-Task Subtraction, Vocal Responses (50 trials)
1K	Single-Task Subtraction, Keyboard Response (50 trials)
2V	Dual-Task with Vocal Responses (3 minutes)
2K	Dual-Task with Keyboard Responses (3 minutes)

APPARATUS

The experiment employed one subject booth of the Multipurpose Automated Research Test Station (MARTS) system illustrated in Figure 2. Stimulus sequences were controlled by a Data General Corporation NOVA 800 minicomputer with 32K x 16 core memory. The computer console was used by the experimenter for input of experimental conditions and for display of performance statistics at the end of each trial. The line printer, a Versatek Matrix electrostatic printer-plotter, provided output of more complete tables and graphs of subject performance at the end of each test session. On-line storage of data was accomplished by means of the magnetic disk. The analog-to-digital (A/D) converter and standard multi-line asynchronous data multiplexor (MPX) converted voltage signals from the joystick, and accepted codes from the keyboard, respectively. A custom-built interface (the MCDS in Figure 2) received and decoded switch closures from the keyboard and transmitted codes to the NOVA 800 MPZ device.

The Megatek Corporation Megagraphics 6000 system used to display tracking and digit-processing tasks to the subjects is a random stroke-drawn cathode-ray tube (CRT) display system capable of presenting alphanumerics and other line-drawn shapes. Stimuli were presented on a Hewlett-Packard model 1310A CRT oscilloscope.

The keyboard, configured with microswitches, was positioned on the left side of the testing booth and arranged in two rows of four buttons each: 0 - 1 - 2 - 3 (bottom row), and 4 - 5 - 6 - 7 (top row). Switch travel was approximately 1 mm before contact. The two-row arrangement was selected to provide rapid learning of the keyboard and to decrease the requirements to shift visual attention from the CRT.

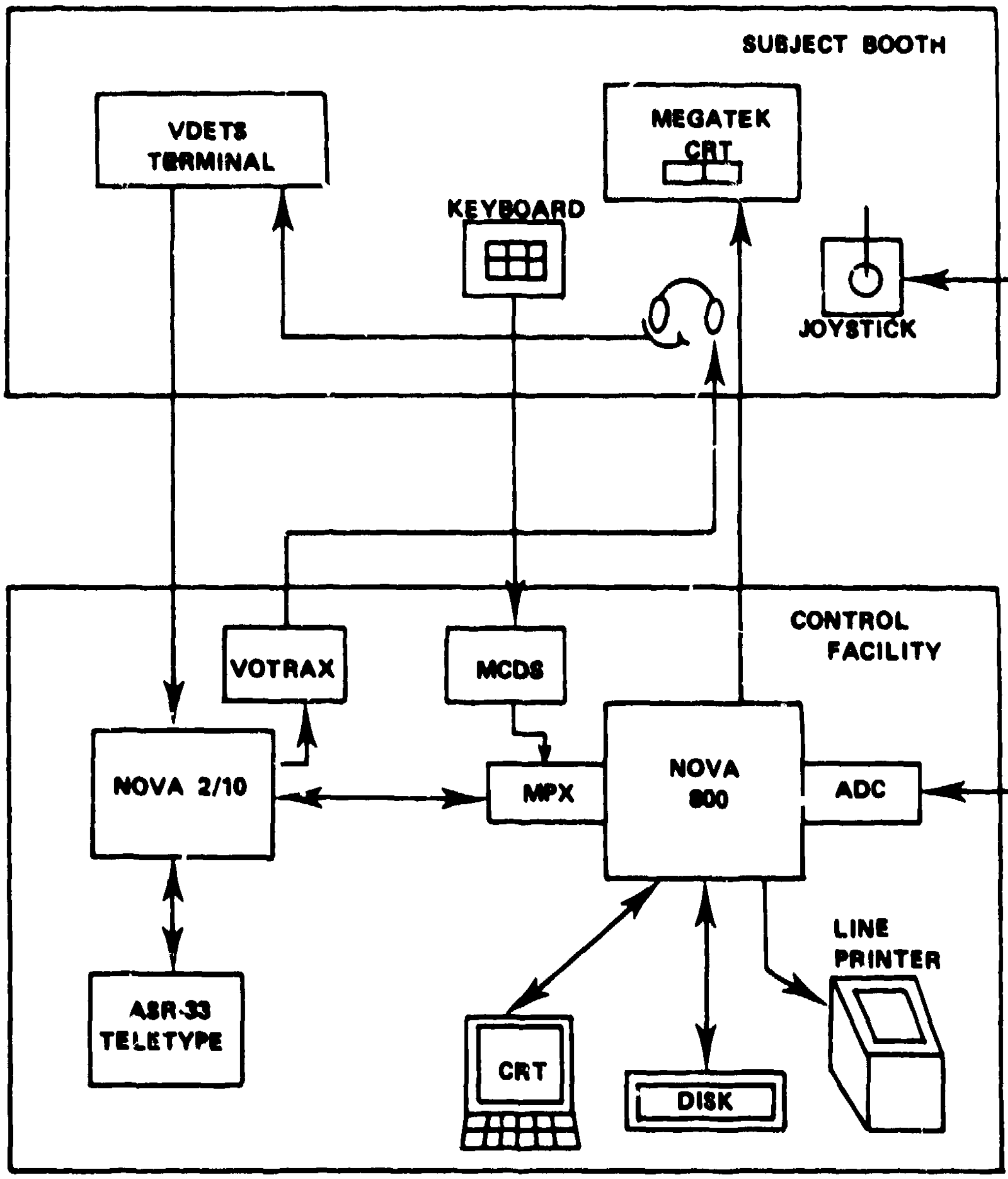

Figure 2. Computer, hardware, and display interfaces for performance of digit and tracking tasks.

Voice recognition and synthesis functions were performed by a Scope Electronics Voice Data Entry Terminal System (VDETS)*, consisting of a Data General Corporation NOVA 2/10 with 16K of 16 core memory, a Scope user's station, a voice synthesizer, and an ASR-33 Teletype. The NOVA 2/10 was linked to the NOVA 800 host computer via a duplex MPX channel. The Scope user's station converts voice analog signals from a microphone mounted on the subject's headset to digital format for entry into the NOVA 2/10. A Vocal Interface Division model VS-6 Votrax (VTX) voice synthesis unit provided auditory output signals to the subject's headset in the testing booth. The Teletype was used by the experimenter to control and monitor the VDETS utterance recognition performance.

SINGLE-TASK PROCEDURES, TRACKING

The subjects performed a one-dimension compensatory tracking task requiring appropriate left-right movements of a joystick control to maintain the position of a diamond-shaped cursor in the center of the 9 cm-long horizontal track (see Figure 3). The disturbance forcing function input consisted of the sum of three nonharmonically related sinusoidal waveforms. The joystick was a Measurement Systems, Inc., model 526 spring-centered finger control with lateral deflection range of $\pm$ 30 degrees, a break-out force of 170 gm, and a full-deflection actuating force of 283.5 gm.

Subjects tracked for two 3-minute trials with a 2-minute rest period intervening. The joystick initially acted as a pure velocity controller. Task difficulty was adaptively increased by adjusting the ratio of acceleration-to-velocity components in the stick control dynamics. When the subject maintained less than 20 percent of scale error, the percentage of acceleration gradually increased in 0.05-percent steps every 50 msec. Acceleration was decreased in the same manner whenever the subject was outside the adaptive criterion. The difficulty of the tracking task was manipulated in this manner in an attempt to reduce the effects of individual differences in tracking skill on the dependent measures of tracking performance. The percent acceleration variable was successful in manipulating tracking difficulty in previous studies (2).

The task remained adaptive for the first four minutes of performance (the entire first trial plus 1 minute of the second trial) and remained at the attained percent acceleration for the final 2 minutes of the second trial. a digital approximation to Root Mean Square Error (RMSE) was computed over 10-second intervals for the final 2 minutes of single-task performance, and the mean and standard deviations of these values were computed to represent the subject's single-task tracking performance. Time on target (TOT) was also computed for this interval.

---

*VDETS is now marketed by Interstate Electronics.

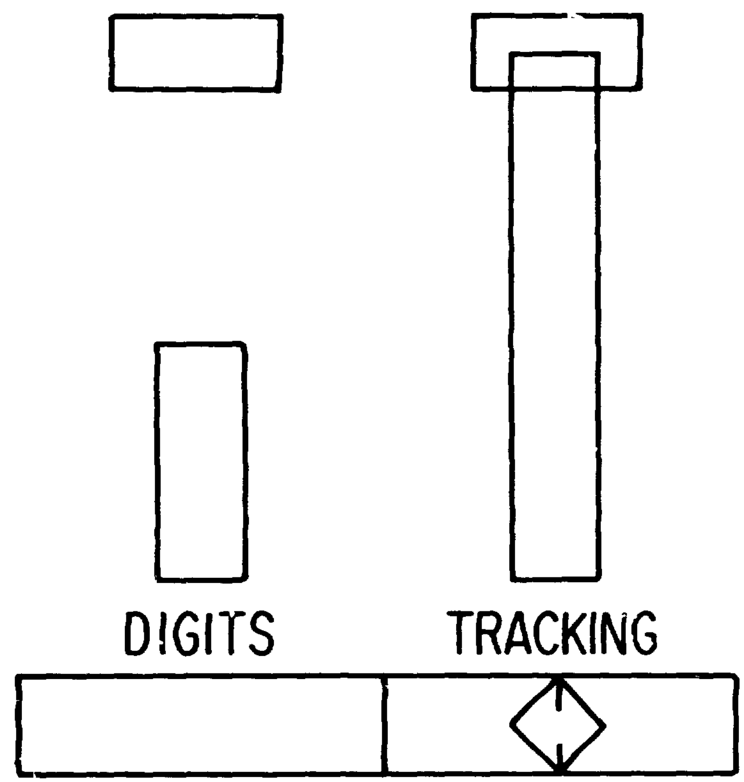

Figure 3. CRT Display Format

A continuous visual performance feedback indicator was presented to the subject throughout single-task performance in the form of a vertically moving bar graph (see Figure 3). The momentary height of the graph, updated each second, corresponded to tracking TOT computed over the immediately preceding 10-second interval; the higher the indicator, the better the performance. A small rectangular box indicated a desired performance level which the subject was instructed to reach or exceed. This level represented 50 percent TOT, which corresponded to the adaptive criterion (20 percent of scale). The maximum height of the performance indicator corresponded to a 100 percent TOT score. In single-task conditions the underlying scale of the feedback graph was linear.

SINGLE-TASK PROCEDURES, DIGIT PROCESSING

The digit-processing task required subjects to compute the absolute difference between two successive digits in a pseudo-random sequence. Stimulus digits varied between 0 and 7. Responses fell within the same range. The task was subject-paced. As soon as the subject responded with the absolute value of the difference between the current digit and the previous digit in the sequence, a new digit was presented. An example of a typical presentation sequence and associated responses is given below:

Stimulus sequence: 7 - 4 - 2 - 7 - 3 - 5 - 2 - 0 . . .

Subject responses: 3 - 2 - 5 - 4 - 2 - 3 - 2 . . .

Subjects were tested in two response conditions: 1) vocal (VDETS); 2) manual (keyboard). The order of response conditions was counterbalanced across subjects. In both conditions a new stimulus digit was presented only after a correct response. A single-task session consisted of 50 trials.

In the event the subject forgot the previous stimulus digit, he could request that it be repeated by either pressing a designated key in the keyboard condition, or by saying "again" in the vocal response condition. Also, in the vocal response condition if the recognition system failed to understand the subject's response, he was notified through the Votrac unit, which responded with the phrase "Say again." In this instance, subjects were instructed to repeat their response. Average response time on correct trials, average response time for trials containing errors, the number of errors, the number of trials with requests for repeated stimuli, and the number of trials with recognition failures were recorded for each session.

A vertically moving feedback bar graph was used during digit-task performance. The momentary height corresponded to the average time between correct responses for the preceding ten trials. The desired level box initially represented a 3.0 second average correct response interval and the full range of the bar graph extended from 4.5 to 1.5 seconds from bottom to top, respectively. When performance improved beyond 3.0 seconds, the desired level became the current best average. Thus, the criterion for good performance was continually changed to represent maximum performance during the 50-trial sequence. The underlying scale for feedback in the digit-processing task was linear in the single-task condition.

DUAL-TASK PROCEDURES

After performing the tasks individually, subjects performed both tasks together for two 3-minute trials. The order of response modality conditions for the digit task was again counterbalanced. The tracking task difficulty was fixed at the attained level of acceleration control achieved in the adaptive portion of single-task sessions. Performance feedback indicators were again used for each task; however, the desired performance region represented the mean single-task performances of the two tasks (see Figure 3). For tracking this was the mean RMSE percent of scale from the final 2 minutes of single-task performance. For digit-processing the goal represented the mean correct response latency for the final 30 trials in single-task performance. Thus the subject was given continuous momentary performance indications, representing the difference between current dual-task performance and the mean of his single-task performance. Subjects were instructed to attempt to reach or exceed these goal lines during the session, and that the tasks were of equal priority. The actual levels that the goals represented were not revealed to the subjects in the instructions. The first minute of dual-task performance was excluded from computation of performance measures to reduce warm-up effects.

The movement of the performance feedback indicators in dual-task conditions was individualized for each subject, based on his mean and standard deviation from single-task performance. The height of the indicator represented the difference between single-task performance and the current momentary dual-task performance measures in standard score units. The formula for this calculation was:

$$\text{Standard Score} = \frac{\bar{X}_{st} - \bar{X}_{dt}}{st}$$

where X_{st} represented the single-task means; X_{dt} represented the momentary dual-task performance computed over the previous 10 seconds of tracking, or ten digit responses; and $\underline{sd}_{st}$ was the standard deviation of the performance distribution in single-task performance. This standard score was then displayed to the subject as the momentary height of the graph. The range of height covered 1.5 standard units above and below the mean. For tracking the bar height was updated every second, and for digit processing, after every response.

RESULTS

A separate analysis of variance (ANOVA) was conducted for each dependent measure discussed below. The first day of testing was treated as a learning session. Data from the second day of testing are analyzed below. Presentation and response modality for the absolute difference task and task loading (single- vs. dual-task conditions) were the independent variables in this analysis. Tracking data from two subjects in the auditory stimulus presentation (VTX) condition were lost due to experimenter error. Data from one subject in the CRT presentation condition indicated that he failed to learn the digit task, and his data were excluded from the analysis. One additional subject was randomly discarded from the CRT presentation group. Data from a total of 16 subjects were analyzed, eight subjects in each group.

TRACKING PERFORMANCE

A digital estimate of root mean square tracking error (RMSE) was computed, based on the immediately preceding 10-second of absolute error measured every 500 msec. Values of RMSE were recorded every 5 seconds during the 3-minute trials. The mean of these values for the final two minutes of each trial represented tracking performance.

A graph of RMSE as a function of the experimental conditions is presented in Figure 4. The ANOVA summary for these scores is shown in Table II. A Tukey post hoc analysis revealed that tracking performance was reliably superior in the vocal condition compared to the manual keyboard response condition. Single-task tracking was superior to both dual-task conditions. There was no reliable effect of presentation modality or interaction between presentation and response modes.

Table II

Analysis of Variance for RMS Tracking Error

SOURCE	df	MS	F	P
Between Subjects	15			
A - Stimulus Mode	1	137.025	.843	.623
Sub. W. Groups	14	162.613		
Within Subjects	32			
B - Task Load	2	2465.832	50.305	.001
A x B	2	7.295	.149	.863
B x Subj. in Groups	28	49.018		

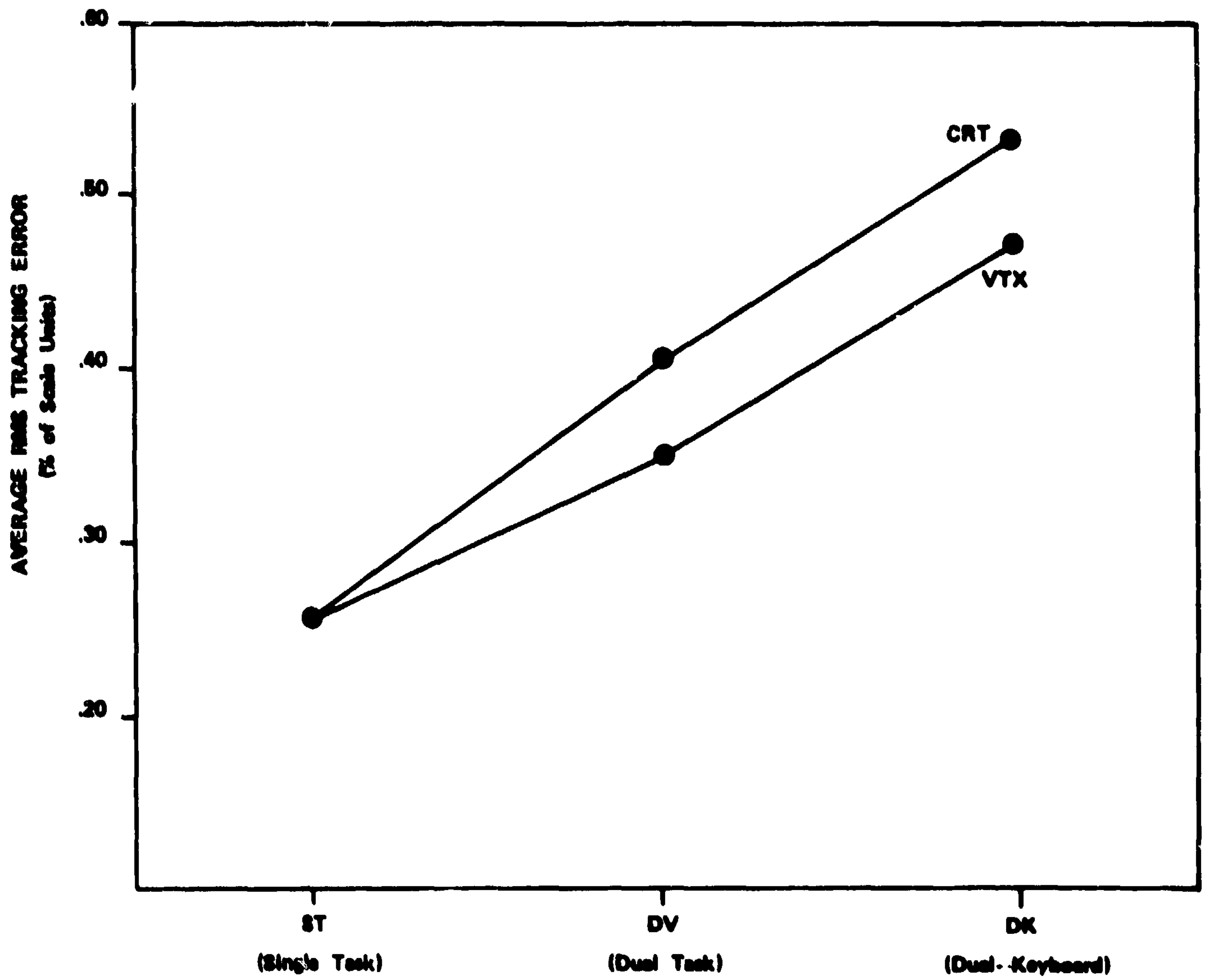

Figure 4. RMS Tracking error as a function of task load and stimulus and response modes for the digit task.

DIGIT-PROCESSING PERFORMANCE

Two dependent measures represented absolute difference task performance: 1) Average Correct Response Latency (ACRL); and 2) Percent Correct Trials (PCT). ACRL was computed by averaging latencies of all responses from trials in which no error or system recognition failure occurred. PCT was computed by dividing the total number of errorless trials by the total number of trials, excluding trials containing recognition failures in the voice condition. Three types of errors could occur in the digit-processing data: errors made by the subject, misrecognition of a correct response by the VDETS, and failure of the VDETS to recognize a correct response as a member of the task vocabulary. Therefore, the denominator in the PCT score could contain an unknown number of misrecognitions.

The means of the two dependent mesures describing digit-processing performance are summarized in Figures 5 and 6. The analysis of variance summaries for these data are presented in Tables III and IV. The visual presentation conditions produced reliably superior performance over the auditory conditions for ACRL. PCT was not reliably affected by this factor. Response modality did not affect either of the two measures. However, the interaction of presentation and response modes was reliable for both scores.

Table III

Analysis of Variance for Average Correct Response Latency

SOURCE	df	MS	F	P
Between Subjects	15			
A - Stimulus Mode	1	12.520	34.136	.001
Subj. w. groups	14	.367		
Within Subjects				
B - Response Mode	1	0.000	0.000	.976
C - Task Load	1	.223	6.055	.026
A x B	1	1.545	26.226	.001
A x C	1	.057	1.543	.233
B x C	1	.019	.833	.620
A x B x C	1	0.000	.011	.913
B x Subj. w. groups	14	.059		
C x Subj. w. groups	14	.037		
BC x Subj. w. groups	14	.023		

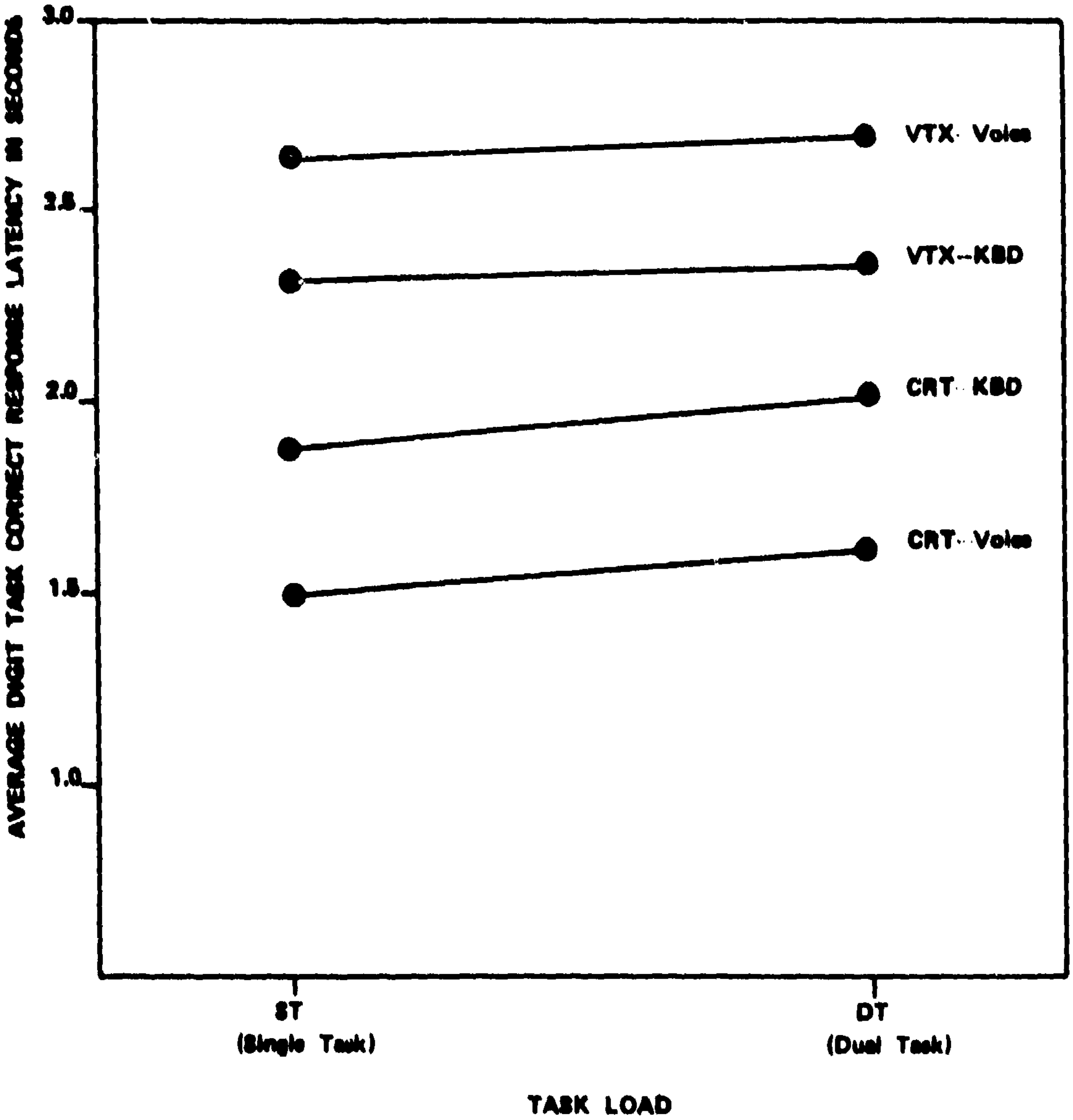

Figure 5. Average correct response latency for the digit task as a function of task load and stimulus and response modes for the digit task.

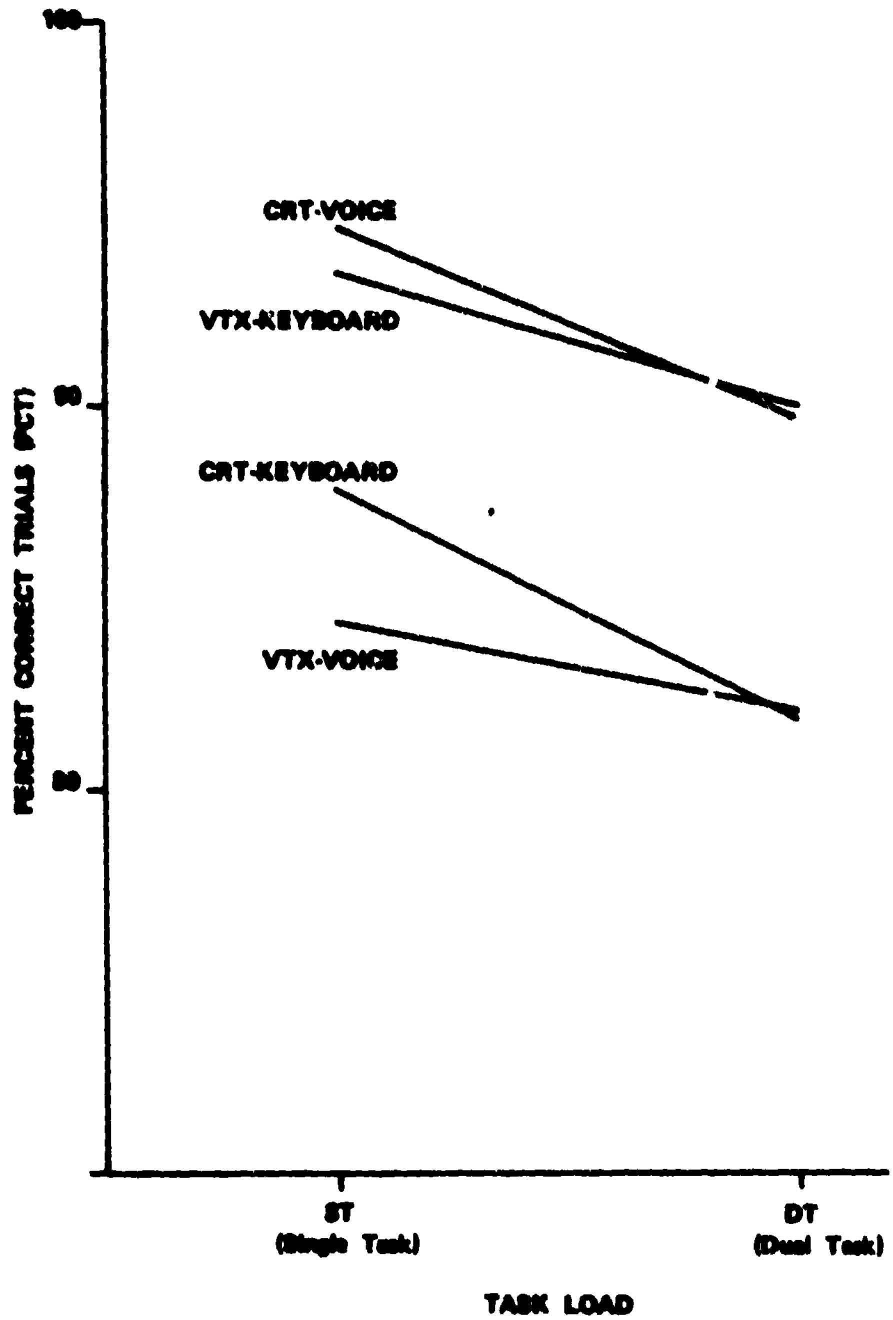

Figure 6. Percent correct trials for the digit task as a function of task load and stimulus and response modes for the digit task.

Table IV

Analysis of Variance for Percent Correct Trials

SOURCE	df	MS	F	P
Between Subjects	15			
A - Stimulus Mode	1	.001	.093	.761
Subj. w. groups	14	.010		
Within Subjects	48			
B - Response Mode	1	0.000	.040	.837
C - Task Load	1	.026	9.612	.008
A x B	1	.096	11.145	.005
A x C	1	.004	1.387	.258
B x C	1	0.000	.249	.630
A x B x C	1	0.000	.038	.842
B x Subj. w. groups	14	.008		
C x Subj. w. groups	14	.003		
BC x Subj. w. groups	14	.003		

The main effect of task load was reliable for both performance scores. Dual-task conditions produced longer response times and higher error rates than single-task trials.

DISCUSSION

The purpose of this experiment was to determine whether a speech understanding system would provide a parallel channel for the performance of an information processing task concurrently with a continuous visual/manual control task. The question can be restated in two parts: 1) What combination of input and output (I/O) channels for the discrete information processing task provides optimum information transmission for both tasks? 2) Is this optimum equivalent to single-task performance?

The results indicate that both tracking and digit-processing performance deteriorated in dual-task conditions. In answer to the second question above, no combination of I/O channels resulted in dual-task performance equivalent to single-task performance. The assumption that a speech understanding system provides a completely parallel channel is apparently unwarranted in this case.

To answer the first question, the data presented in Figures 4, 5, and 6 were redrawn in Figures 7 and 8. Because the priorities between the tasks were held constant and equal, it seems reasonable to interpret each of the points

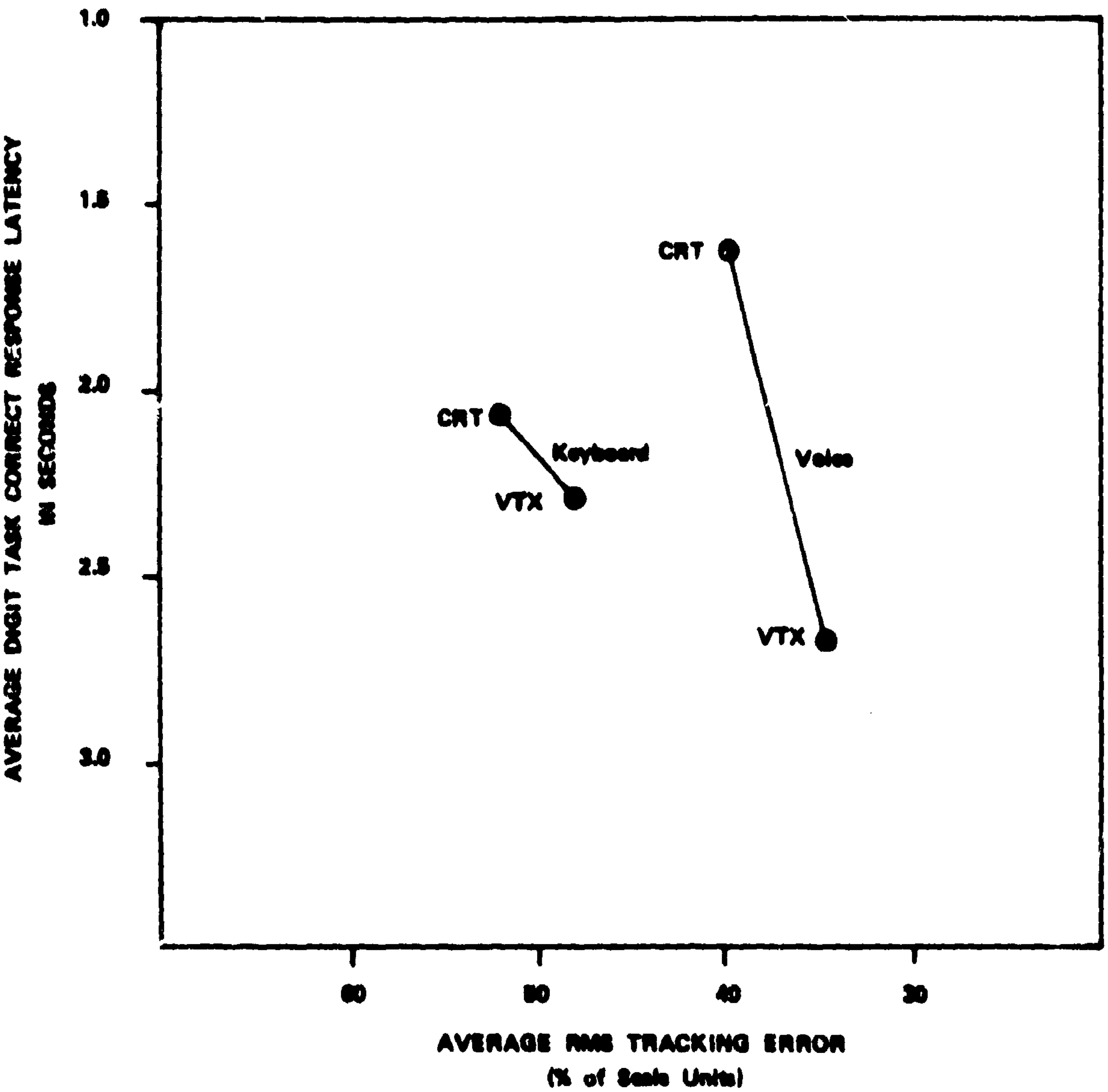

Figure 7. Joint performance space for tracking error and digit-processing correct response latency.

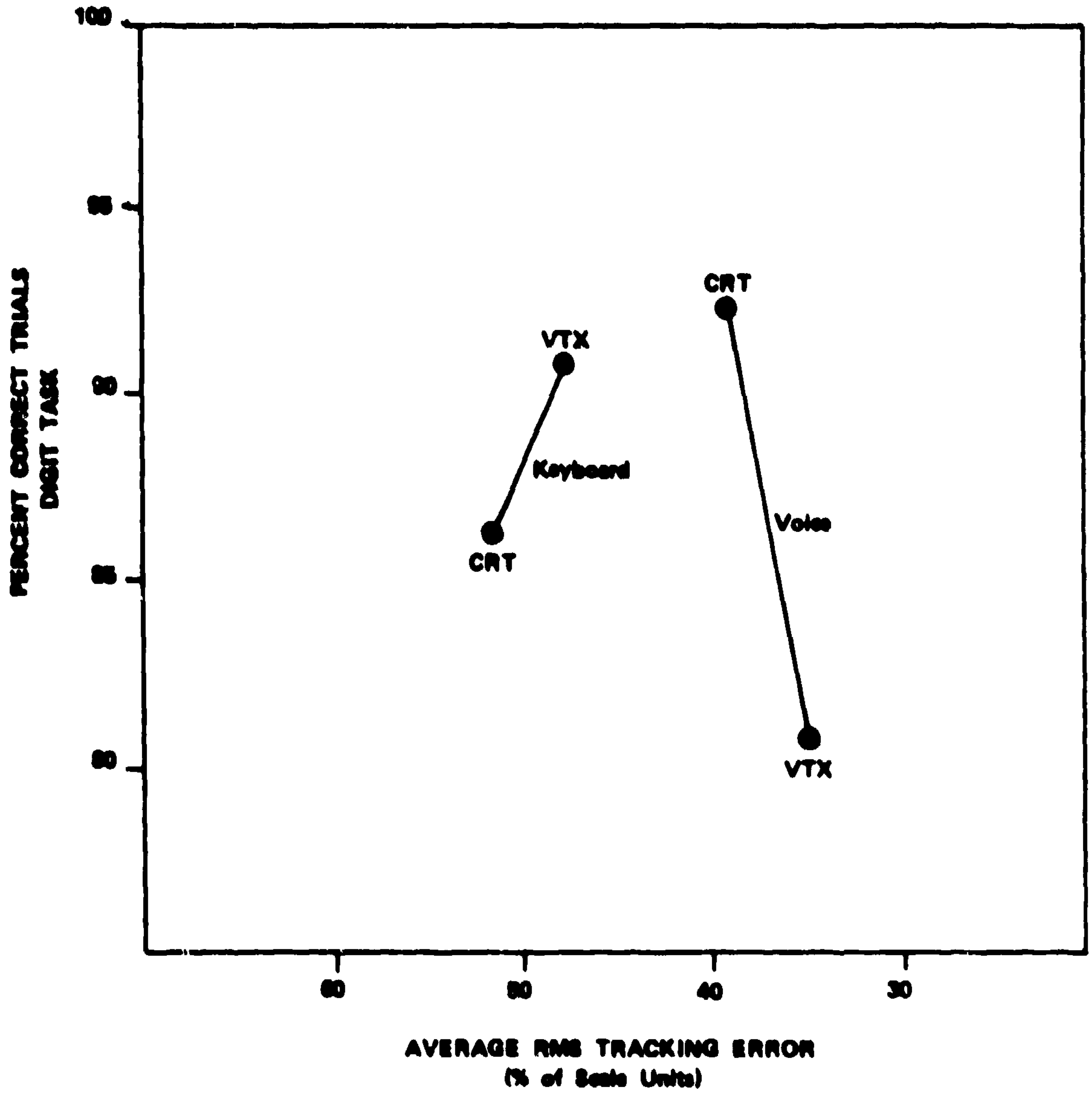

Figure 8. Joint performance space for tracking error and digit-processing accuracy.

plotted in Figures 7 and 8 as an estimate of the center point of each of four performance operating characteristics. The remaining points for each curve could theoretically be generated by repeating the experiment with priorities set for different tradeoffs. Until the shape of the POC for each combination of tasks is known, the conclusions discussed below concerning the relative merits of various combinations of I/O channels for the digit task must be regarded as tentative. The data in Figures 7 and 8 suggest that the conjoint performance of the tasks is maximized when CRT presentation and voice response are employed in the digit task. Other combinations of I/O channels for the digit task resulted in different points in the joint performance tradeoff space.

In practical military aviation systems the placement and organization of keyboards are usually not optimal, and visual displays are cluttered, while voice I/O channels are relatively unused. The format of the stimulus display and the layout of the keyboard and joystick in this experiment were selected to provide high S-R compatibility and to minimize structural interference between the time-shared tasks. Any alterations in the stimulus format or keyboard/joystick relationships would probably have either no effect, or would be detrimental to performance of the tasks. A significant improvement in performance seems unlikely. Therefore, the relative advantage of the voice responce channel over the keyboard channel that is apparent in Figures 7 and 8 may be expected to increase in a practical system. However, the location of points in the joint performance space should be explored for tasks that are more representative of the physical and information processing characteristics of a specific system.

The latency data (ACRL) for the digit-processing task represented the elapsed time from the onset of the stimulus to the termination of a correct response. In the voice response mode this time included an average 527 msec required by the VDETS to accomplish utterance recognition. It is expected that recognition latency will be reduced to very nearly zero in speech understanding systems that are currently under development. An estimate of the improvement in man-machine system performance that such a development would afford can be obtained by subtracting 0.527 sec from the ACRL data for the voice conditions. This estimate considerably enhances the apparent advantage for the voice channel.

Because the voice recognition system could misinterpret subject responses, perhaps confusing the digits five and nine as might a human listener, the error score (PCT) for the digit-processing task reflected both subject performance and system performance. The effect of the task loading variable on error data may have been due to unknown variations in the speech signal spectrum as a function of processing load imposed on the subject, resulting in increased misrecognitions by the system. However, the effect of processing load was also observed in the latency data. System recognition latency is a function of vocabulary size, syntax structure, and length of the utterance. Vocabulary size and syntax for the digit task were fixed; any variations in utterance duration were due to the subject. The effect of processing load on the latency data supports the conclusion that the error data reflect errors due to the subject rather than system recognition failures. Future experiments should include procedures for confirming off-line that the system's recognition accuracy is constant across experimental conditions.

A short-coming in the design of the present experiment was that feedback to the subject was a function of correct response latency only. Subjects could maximize the height of the performance feedback indicator in the digit task by generating very fast, but frequently inaccurate, responses. The problem of interpreting performance data when the subjects' speed-accuracy tradeoff strategy has not been controlled is evident from the data for the keyboard response mode in Figures 7 and 8. As shown in the Figures, the CRT-keyboard combination of I/O channels resulted in faster average correct response latencies than the VTX-keyboard condition, but accuracy of responses in the CRT-keyboard condition was lower than in the VTX-keyboard condition. Conjoint tracking performance for the two conditions was not reliably different. It appears that subjects in the CRT group assumed a response strategy that emphasized speed, whereas the VTX group emphasized accuracy. To investigate this question an analysis of variance was performed, comparing the average latency of correct and incorrect keyboard responses for the two groups. The results of the ANOVA are summarized in Table V.

Table V

Analysis of Variance for Response Latencies

SOURCE	df	MS	F	P
Between Subjects				
A - Stimulus Mode	1	.155	.431	.528
Subj. w. groups	14	.380		
Within Subjects				
B - Response Type*	1	1.044	8.241	.012
A x B	1	1.034	8.155	.012
Subj. w. groups	14	.127		

* Correct vs. Incorrect Responses.

If subjects were trading speed for accuracy, incorrect responses should, on the average, have been faster than correct responses. The results indicated that incorrect responses were slower than corrects for the CRT group. There was no difference between corrects and incorrects for the VTX group. If the assumption is made that subjects attempted to maintain a consistent speed-accuracy strategy in all experimental conditions, the results suggest that the two measures of performance in the digit task (ACRL and PCT) reflect processes that are differentially affected by stimulus presentation and response modes. In future experiments an attempt should be made to control the subjects' speed-accuracy tradeoff strategy, perhaps by computing feedback as a joint function of speed and accuracy, and varying the contribution of each to the level of a performance feedback indicator. A related experimental question concerns whether the results reported in this paper are representative of human performance in tasks that lack graphical indicators of performance. Future experiments should also include conditions in which graphical performance indicators are not available to the subject, in order to assess the effects of the feedback display on the shapes of the performance operating characteristics for the tasks.

The interaction between stimulus presentation mode and response mode for both accuracy and latency scores in the digit-processing task indicated that for auditory inputs, the keyboard response mode was superior, and for visual inputs, the voice response mode resulted in better performance. In the auditory input, vocal output condition, the acoustical attributes of both stimulus and response may have been a unique source of interference that contributed to reduced performance efficiency. It seems especially likely that rehearsal and retrieval processes active during digit-task performance were more susceptible to disruption by intervening vocal responses than by manual responses. The results demonstrated that the peculiar information processing requirements of a task must be taken into account when specifying the I/O channel structure for optimum task performance.

CONCLUSIONS

The results indicated that the combination of CRT stimulus mode and voice response mode provided optimum joint-task performance. No combination of I/O channels resulted in equivalent single- and dual-task performance. Future experiments should be designed to investigate the joint-performance space for tasks that are more representative of the information processing performance requirements of specific systems. However, the interpretability of the results of such research will depend upon the solution of methodological problems, such as how to control or account for the subjects' speed-accuracy tradeoff strategies and the priorities they place upon the concurrent tasks.

REFERENCES

1. Beeks, B., Neuberg, E. P., and Hodge, D. D. An assessment of the technology of speech recognition for military applications. IEEE Transactions on Acoustics, Speech, and Signal Processing, VOL. ASSP-25: No. 4, August, 1977.

2. Crooks, W. H., and Roscoe, S. N. Varied and fixed error limits on automated adaptive skill training. In: Ranc, M. P. Jr., and Malone, T. B. (Eds.), Proceedings of the Seventeenth Annual meeting of the Human Factors Society. Santa Monica, Calif.: Human Factors Society, 1973.

3. Coler, C., Plummer, R., Huff, E., and Hitchcock, M. Automatic speech recognition research at NASA-Ames Research Center. In: Curran, M., Breaux, R., and Huff, E. (Eds.), Voice Technology for Real-time Command/Control Systems Application. Proceedings of a symposium. Moffett Field, Calif.; Ames Research Center, 1977.

4. Curran, M. Voice integrated systems. In: Curran, M., Breaux, R., and Huff, E. (Eds.), Voice Technology for Interactive Real-time Command/Control Systems Application. Proceedings of a symposium. Moffet Field, Calif.: Ames Research Center, 1977.

5. Harris, S. Human performance in concurrent verbal and tracking tasks: A review of the literature. Special Report 78-2, Pensacola, Fla.: Naval Aerospace Medical Research Laboratory, 1978.

6. Harris, S. D., North, F. A., and Owens, J. M. A system for the assessment of human performance in concurrent verbal and manual control tasks. Behav. Res. Meth. & Instrumentation, 1978, Vol. 10 (2), 329-333.

7. Kantowitz, B., and Knight, J. On experimenter-limited processes. Psych. Rev., 83(6): 502-507, 1976.

8. Navon, D., and Gopher, D. On the economy of the human processing system: A model of multiple capacity. Technical Report Number AFOSR-77-1. Haifa, Israel: Israel Institute of Technology, 1977.

9. Newell, A., Barnett, J., Forgie, J., Green, C., Klatt, D., Licklider, J. C. R., Munson, J., and Reddy, R. Speech-understanding systems: Final report of a study group. Pittsburgh: Carnegie-Mellon University, 1971.

10. Newell, A., Cooper, F., Forgie, J., Green, C., Klatt, D., Medress, M., Neuburg, E., O'Malley, M., Reddy, R., Ritea, B., Shoup, J., Walker, D., and Woods, W. Considerations for a follow-on ARPA research program for speech understanding systems. Pittsburgh: Carnegie-Mellon University, 1975.

11. Norman, D., and Bobrow, D. On data-limited and resource-limited processes. Cog. Psychol., 1: 44-64, 1975.

12. Norman, D., and Bobrow, D. On the analysis of performance operating characteristics. Psych. Rev., 83(6): 508-510, 1976.

13. Reddy, D. R. Speech recognition by machine: A review. Proceedings of the IEEE, 64(4), 1976.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	(14) NAMRL-1259
2. GOVT ACCESSION NO.	
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle)	(6) Human Performance in Time-Shared Verbal and Tracking Tasks
5. TYPE OF REPORT & PERIOD COVERED	(9) INTERIM rept.
6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s)	(10) Steven D. Harris, ~~LT MSC USN~~ Jerry M. Owens ~~LT MSC USN~~ ~~and~~ Robert A. North ~~Ph.D~~
8. CONTRACT OR GRANT NUMBER(s)	
9. PERFORMING ORGANIZATION NAME AND ADDRESS	Naval Aerospace Medical Research Laboratory, NAS Pensacola, Fl 32508
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS	ZF51.524.004-9017
11. CONTROLLING OFFICE NAME AND ADDRESS	Naval Medical Research and Development Command, National Naval Medical Center, Bethesda, Maryland 20014
12. REPORT DATE	(11) Apr 1979
13. NUMBER OF PAGES	24
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	(12) 27 p.
15. SECURITY CLASS. (of this report)	UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	

406 061

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

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

(16) F51524 (17) ZF51524004

18. SUPPLEMENTARY NOTES

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

- Automated Speech
- Information Processing Task
- Visual/Manual Control Task
- Time-Share
- Voice Response
- Speed-Accuracy Tradeoff
- Human Performance
- Verbal and Tracking Tasks
- Information Processing
- Controlling Systems

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

Significant progress has been made in the development of automated speech understanding systems for application to naval aviation systems. One advantage that is anticipated for speech over conventional man-machine interfaces is that speech could function as an independent channel for the control of systems. The experiment reported in this paper represents an initial effort to investigate the assumption that an automatic speech understanding system will provide a parallel channel for the performance of an information processing task currently with a visual/manual control task.

The experiment required human subjects to time-share a digital information processing task and a continuous compensatory tracking task. Independent variables in the design were task loading (single- vs. dual- task

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

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

406 061

20.
conditions), stimulus presentation modality for the digital task (auditory vs. visual), and response modality for the digital task (voice vs. keyboard). Data from 16 subjects were analyzed.

The results indicated that the combination of visual stimulus modality and voice response provided optimum joint-task performance. No combination of stimulus and response modalities resulted in equivalent single- and dual-task performance. Future experiments should be designed to investigate the joint-task performance space for tasks that are more representative of the information processing performance requirements of specific systems. However, the interpretability of the results of such research will depend upon the solution of methodological problems, such as how to control or account for subjects' speed-accuracy tradeoff strategies and the priorities they place upon the concurrent tasks.