Technical Memorandum 17-91

THE EFFECTS OF SPEECH INTELLIGIBILITY LEVEL ON CONCURRENT VISUAL TASK PERFORMANCE

David G. Payne
State University of New York at Binghamton

Leslie J. Peters, Deborah P. Birkmire, Georges R. Garinther
U.S. Army Human Engineering Laboratory

September 1991
AMCMS Code 61110274A0011

U.S. ARMY HUMAN ENGINEERING LABORATORY

Aberdeen Proving Ground, Maryland

Realistic® is a registered trademark of Radio Shack Corporation.

PAGES ______

ARE

MISSING

IN

ORIGINAL

DOCUMENT

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Memorandum 17-91	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Human Engineering Laboratory — 6b. OFFICE SYMBOL (If applicable) SLCHE	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Aberdeen Proving Ground, MD 21005-5001	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION — 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State and ZIP Code)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 6.11.02 — PROJECT NO. 1L161102B74A — TASK NO. — WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
The Effects of Speech Intelligibility Level on Concurrent Visual Task Performance

12. PERSONAL AUTHOR(S) Payne, D. G., State University of New York at Binghamton; Peters, L. J., Birkmire, D. P., & Garinther, G. R., U.S. Army Human Engineering Laboratory

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM ______ TO ______	1991, September	35

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP
05	08	
25	05	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
cross-modal interference; speech intelligibility; dual task performance; multiple resource theory

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

Two experiments were performed to determine if changes in speech intelligibility level can impact performance levels in concurrent visual tasks. The auditory task used in both experiments was an auditory memory search task in which subjects memorized a set of words and then decided whether auditorally presented probe items were members of the memorized set. Experiment 1 used an unstable tracking task as the visual task, and Experiment 2 used a spatial decision-making task. Results showed that unstable tracking performance was unaffected by the level of speech intelligibility during the auditory task, whereas accuracy in the spatial decision-making task was significantly worse at low speech intelligibility levels. These results have clear implications for the design of communications systems. The findings are interpreted within the framework of multiple resource theory, and future directions for research are described.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified	
22a. NAME OF RESPONSIBLE INDIVIDUAL Technical Reports Office	22b. TELEPHONE (Include Area Code) (301) 278-4478	22c. OFFICE SYMBOL SLCHE-SS-TSB

AMCMS Code 61110274A0011 Technical Memorandum 17-91

THE EFFECTS OF SPEECH INTELLIGIBILITY LEVEL ON CONCURRENT VISUAL TASK PERFORMANCE

David G. Payne
State University of New York at Binghamton

Leslie J. Peters, Deborah P. Birkmire, & Georges R. Garinther
US Army Human Engineering Laboratory

September 1991

APPROVED ____________
JOHN D. WEISZ
Director
Human Engineering Laboratory

Approved for public release;
distribution unlimited.

U.S. ARMY HUMAN ENGINEERING LABORATORY

Aberdeen Proving Ground, Maryland

CONTENTS

EXECUTIVE SUMMARY

Two experiments were performed to determine if changes in speech intelligibility level can impact performance levels in concurrent visual tasks. The auditory task used in both experiments was an auditory memory search task in which subjects memorized a set of words and then decided whether auditorally presented probe items were members of the memorized set. Experiment 1 used an unstable tracking task as the visual task, and Experiment 2 used a spatial decision-making task. Results showed that unstable tracking performance was unaffected by the level of speech intelligibility during the auditory task, whereas accuracy in the spatial decision-making task was significantly worse at low speech intelligibility levels. These results have clear implications for the design of communications systems. The findings are interpreted within the framework of multiple resource theory, and future directions for research are described.

THE EFFECTS OF SPEECH INTELLIGIBILITY ON
CONCURRENT VISUAL TASK PERFORMANCE

INTRODUCTION

One important factor that contributes to the performance of human-machine systems is the ability of the human operators to communicate with one another. If operators cannot clearly and effectively communicate with one another, system performance will be impacted because of increased errors, longer latency to understand a spoken message, and so forth. These effects of poor speech intelligibility are of obvious concern to system designers, especially when the systems of interest are inherently noisy (e.g., tanks, helicopters). In addition to the demonstrated impact on performance in verbal communication tasks, changes in speech intelligibility have the potential to affect performance in other nonauditory tasks. The present research was designed to investigate the hypothesis that changes in speech intelligibility may affect performance levels in some nonauditory tasks.

Before describing this research in detail, it is appropriate to comment briefly about three aspects of the approaches that have been taken to handle or address problems associated with changes in speech intelligibility. First, basic researchers have produced a wealth of information concerning the factors that affect speech intelligibility levels (e.g., signal-to-noise ratios). This information provides the human factors specialist with a variety of ways to try to improve performance of communications tasks. Second, from a systems design perspective, human factors specialists have developed valuable standardized guidelines for speech intelligibility, such as those presented in MIL-STD-1472D (Department of Defense, 1989). These guidelines specify the desired levels of speech intelligibility for different tasks and environments and are therefore quite useful when trying to identify satisfactory levels of performance of communication tasks.

In contrast to our knowledge of the factors that affect speech intelligibility and the specification of satisfactory levels of speech intelligibility (e.g., communication tasks), there has been very little research to date concerning the impact of changes in speech intelligibility on the performance of other nonauditory tasks. This latter fact is regrettable, given that most real world communications take place in the context of other tasks (e.g., driving while listening to radio announcements). This state of affairs suggests that it is of considerable importance to develop an understanding of the manner in which changes in speech intelligibility will affect performance in multi-task environments. The present research represents an initial effort toward this goal.

To put the present research in perspective, consider how system designers have typically handled the issue of speech intelligibility. In the extant speech intelligibility guidelines (e.g., DoD, 1989), speech intelligibility is measured in terms of performance of measures such as the Modified Rhyme Test (MRT) developed by House, Williams, Hecker, and Kryter (1965). In the MRT, a listener is presented with an auditory target item (e.g., rang), and then a set of six rhyming choices (e.g., rang, fang, gang, sang, hang, bang) is visually presented. The subjects' task is to determine which item was presented auditorally; performance is measured in terms of recognition accuracy.

An important feature of the MRT metric is that this test is typically administered in isolation (i.e., the listener's sole task is to try to identify which target item has been presented). It has been well documented in laboratory studies that as task difficulty increases, subjects try to increase the amount of attention (or capacity) allocated to performing the task. If this phenomenon of changes in task difficulty affecting the amount of attention paid to the task holds for real world situations as well as laboratory studies, this suggests that one consequence of a degradation in speech intelligibility will be an increase in the amount of attention allocated to the auditory task. If the total amount of attention that can be paid to all tasks is limited, then by inference, a change in speech intelligibility will have a "ripple effect" in that it will decrease the amount of attention that an operator can allocate to other tasks. If other factors such as stress and fatigue also come into play, then during these conditions, performance of tasks performed concurrently with the auditory task is likely to deteriorate even more precipitously.

The overall question of interest in the present research was what effect, if any, changes in speech intelligibility will have on operators' performance. There were two general objectives of these experiments. First, the authors sought to identify experimental procedures and methods appropriate for investigating the human information processing resources used when simultaneously performing auditory and visual tasks. Second, assuming that the methodology proved to be sensitive to changes in speech intelligibility, the authors sought to identify what types of nonauditory tasks are likely to be affected by changes in speech intelligibility. Identifying tasks that load on changes in speech intelligibility is crucial for any subsequent efforts to develop tests of individual differences in changes in speech intelligibility.

Following Broadbent (1958), Kahneman (1973) and others, the authors assume that humans are limited in the amount of information they can process per unit of time. It is further assumed that as the amount of attention devoted to performing communications tasks decreases, overall performance levels will also decrease, following the principle of graceful degradation outlined by Norman and Bobrow (1975). That is, performance levels tend to decrease slowly as the amount of attention paid to a task decreases. The interest in the present study is not so much in the fact that performance falters when the operator is overloaded, but rather the specific manner in which performance is affected.

The present research was conducted within the theoretical framework provided by multiple resource theory (e.g., Navon & Gopher, 1979; Wickens, 1980; 1984). Navon and Gopher (1979) proposed that the human cognitive system could be viewed as being composed of a limited number of processing "resources." These processing resources are hypothetical constructs that refer to some underlying commodity that enables a person to perform some task(s). According to this framework, resources are limited in the sense that specific resources may only be allocated to specified processes or subprocesses. In simplest terms, if two tasks require the same resource, these tasks can be efficiently performed together only if the total demand for resources does not exceed the available supply. If demand exceeds supply, performance levels will decline. On the other hand, if the two tasks use different resources, they may be performed together with no change in performance levels relative to single task control conditions. There is ample evidence in the experimental literature to support the general claim that some tasks may be performed simultaneously with little change in performance levels (e.g., Allport, Antonis, & Reynolds, 1972; Shaffer, 1975), whereas other tasks will greatly interfere with each other (e.g., Brooks, 1968). Multiple resource theory also assumes that different resources are differentially efficient when applied to processes or subprocesses. Efficiency here is used in the econometric sense of marginal efficiency (i.e., the change in performance level observed when one unit of a resource is added to or removed from a process). Finally, tasks are assumed to have different resource compositions, which means that different tasks require different resources for the processing involved in performing the task to be completed. Given this overview, multiple resource theories assume that the following factors will affect performance in single and multiple task situations: (a) the resource composition(s) of the task(s) being investigated; (b) the amount of each resource type available to be allocated to the task(s); and (c) the relative efficiency of the resources allocated to the task(s).

One obvious difficulty with an unconstrained multiple resource model is how one determines what the available resources are. Fortunately, Wickens (1980, 1984) has identified the following as reasonable candidates for processing resources: (a) the type of input and output modality (e.g., visual versus auditory stimuli; manual versus vocal responses); (b) the code or representational format used by the subject (e.g., a linguistic code versus a spatial code), and (c) the stage of processing (e.g., encoding, central processing, and response selection and execution). The present research uses Wickens' multiple resource framework by varying the task demands of an auditory and a visual task.

The overall predictions derived from Wickens' multiple resource theory are straightforward. If a visual task and an auditory task employ the same resources, performing these tasks together should be quite difficult, and performance decrements should be seen, especially as the difficulty levels or resource demands of the tasks of interest increase. If, however, the two tasks tap into different resources, the subject should be able to perform the two tasks together quite efficiently, and there should be no effect on concurrent task performance when one task's difficulty level is increased. For present purposes, changes in speech intelligibility levels are assumed to increase the difficulty levels of the auditory tasks by increasing the amount of central processing resources that are required to analyze the auditory stimulus.

To test the foregoing hypotheses, this study employed a dual task methodology (cf., Ogden, Levine, & Eisner, 1979) that has proved useful to researchers investigating memorial and attentional processes in both basic (e.g., Kerr, 1973; Proctor & Proctor, 1979; Roediger, Knight, & Kantowitz, 1977; Tyler, Hertel, McCallum, & Ellis, 1979) and applied (e.g., Carswell & Wickens, 1985; Damos, 1985; Hanson, Payne, Shively, & Kantowitz, 1981) research settings. The logic behind this method is as follows. An operator is required to perform two tasks, both singly and simultaneously, with performance being measured in both single and dual task conditions. The single task conditions provide base line performance levels, and the dual task conditions allow us to determine whether the two tasks selectively interfere with one another. Finally, task difficulty level is manipulated in both tasks so as to vary the amount of resources allocated to the tasks, thereby allowing us to look for the presence or absence of selective interference effects (e.g., does changing speech intelligibility interfere with performance of the visual task?).

To summarize, then, the present research was conducted within the framework provided by Wickens' (1980, 1984) multiple resource theory. One major goal of these experiments was to determine if the multiple resource theory provides a viable approach for understanding how changes in speech intelligibility impact performance in nonauditory tasks. A second goal was to determine if the dual task methodology could be used effectively to examine the impact(s) of degrading speech intelligibility. A third goal was to try to identify tasks and processes that are impacted by changes in speech intelligibility. If it is possible to identify such tasks and processes, this will be very useful in subsequent efforts toward developing effective methods for identifying those individuals who are likely to be severely affected by changes in speech intelligibility.

To test the effects of changes in speech intelligibility on ongoing visual task performance within the context of multiple resource theory, it is necessary to (a) have a reliable method for establishing the desired intelligibility levels, and (b) select visual tasks that tap into different underlying resources. Fortunately, both these needs can be met by using procedures that have been validated in previous research.

Peters and Garinther (1990) (see also Whitaker, Peters, & Garinther, 1990) have employed a chopping circuit designed by the U.S. Army Human Engineering Laboratory (HEL) to vary the intelligibility levels obtained when communications take place between two member teams. (This chopping circuit is described in detail in the Method section of this report.) The Whitaker et al. (1990) research is important to this study because their results indicated that the chopping circuit parameters can be easily adjusted to produce the desired level of speech intelligibility. These studies have also demonstrated that changes in speech intelligibility appear to interfere selectively with performance of concurrent nonauditory tasks.

In the Peters and Garinther (1990) study, an armor simulator was employed and two-person crews completed a sequence of mission scenarios at different levels of speech intelligibility. A variety of performance measures relating to mission time, mission completion, mission errors, and gunner accuracy were recorded. One very interesting finding that emerged from this study was that while a number of performance measures were greatly affected by altering the speech intelligibility levels within the crew communication system (e.g., time to identify target, number of enemy targets killed), some measures (e.g., gunner accuracy) were relatively unaffected by changes in speech intelligibility.

The Peters and Garinther (1990) study was not designed to be analytical with respect to the reasons underlying these different patterns of results (i.e., effects versus no effects of changes in speech intelligibility), and hence, it is impossible to isolate the factor(s) critical to producing these differences. However, it is possible that the multiple resource theory framework employed here can provide a tentative explanation. If the auditory and nonauditory tasks tap into the same resource, an overall performance decrement would be expected, along with decreases in speech intelligibility, and such a decrement would not be expected when the tasks do not tap into the same resources. To test these notions, the authors needed a set of visual tasks that tap different resources.

Shingledecker, Crabtree, and Acton (1982) developed the criterion task set (CTS), which is a series of nine visual tasks designed to differentially load on various perceptual motor and cognitive processes. The CTS was developed within the framework of multiple resource theory and has undergone extensive validation (Schlegel & Shingledecker, 1985).

Each of the nine CTS tasks has three different levels of difficulty that can be used to systematically load different resources. The present research used two of the CTS tasks, one that relies primarily upon visual encoding and manual responding (unstable tracking; Experiment 1) and one that requires the central processes of working memory and decision making (spatial processing; Experiment 2). Note that these two tasks are analogous to real world perceptual motor tracking tasks (e.g., sighting on a moving target, driving) and cognitive, decision-making tasks (e.g., locating an enemy target).

Finally, the auditory task used in the present experiments was an auditory analog of the memory search task developed by Sternberg (1969). In this task, subjects memorize a small set of target items (in this case, spoken words) and are then presented with a series of probe items, some of which came from the memorized target set. The subjects' task is to decide as quickly as possible if each probe item is a member of the target set. The primary performance measure is the subjects' response latency or reaction time (RT), and task difficulty is manipulated by varying the number of items in the target set. There is empirical evidence indicating that increasing the number of items in the target set produces significant increases in RT.

EXPERIMENT 1

Experiment 1 was designed to test whether visual task performance levels would be affected by changes in speech intelligibility in a concurrent auditory task when the visual and auditory tasks tapped into largely different mental resources. The visual task was the CTS unstable tracking task, which is presumed to load heavily on visual encoding and manual responding. The auditory task was the auditory Sternberg task performed at four different levels of speech intelligibility. Task difficulty was manipulated in the unstable tracking task by varying the parameter that controls the degree of system instability; this parameter has been shown by Shingledecker et al. (1982) to produce three distinct single task performance levels. Task difficulty in the auditory Sternberg task was manipulated by varying the number of items (two versus four) in the target set.

All subjects performed the visual and auditory tasks in both single task and dual task trials. Half the subjects were given an easy version of the unstable tracking task, and the remaining subjects performed a difficult version. All subjects performed four blocks of trials, with speech intelligibility varied across blocks. If unstable tracking relies predominantly on visual encoding and manual responding with relatively little central processing required, according to Wickens' (1984) multiple resource framework, there should be no effect of speech intelligibility on tracking task performance. There should, however, be effects of the two difficulty manipulations on the corresponding tasks, and speech intelligibility should affect performance in the auditory Sternberg task.

Method

Subjects

Twenty-eight students, who enrolled in Introductory Psychology at the State University of New York at Binghamton (SUNY-Binghamton), participated in partial fulfillment of a course requirement for research experience or library research. Subjects were tested individually in a small (1.8 m by 3.0 m) sound-attenuated room. Each experimental session lasted approximately 2 hours.

Design

The experiment involved one between-subjects variable, level of difficulty (easy versus difficult) on the unstable tracking task. The three within-subjects variables were (a) level of speech intelligibility (20%, 40%, 60%, 80%), (b) number of target items (two or four) in the memory set on the auditory Sternberg task, and (c) trial type, either single task (unstable tracking or auditory Sternberg) or dual task (unstable tracking performed concurrently with the auditory Sternberg task). All subjects completed four blocks of trials, one at each level of speech intelligibility. The order of speech intelligibility was varied using a balanced Latin square design. Within each block of trials, subjects completed seven tasks in the following order: (a) initial MRT trial, (b) single task auditory Sternberg with two targets, (c) single task unstable tracking (at the level of difficulty appropriate for that subject), (d) dual task trial (auditory Sternberg plus unstable tracking) with two target items in the Sternberg task, (e) single task auditory Sternberg with four target items, (f) dual task trial with four target items in the Sternberg task, and (g) final MRT trial.

Apparatus

The unstable tracking task was controlled by a Commodore 64 microcomputer interfaced with a Commodore Model 1702 color monitor and a 1.25-inch (3.5-cm) diameter rotary knob mounted in a 4-inch x 2-inch (10- by 5-cm) response box. The response box was fastened on the right-hand side of the desk at the subject's test station. Subjects used their right hands when performing the unstable tracking task.

The auditory stimuli were recorded using a high quality microphone and a Data Translation analog-to-digital interface card (Model DT2801) along with a 286-based IBM-compatible microcomputer operating at 12.7 mHz. The stimuli were presented to subjects using the Data Translation card's digital-to-analog capabilities. These items were processed by the chopping circuit described by Peters and Garinther (1990), amplified using a Radio Shack model SA-150 amplifier and presented to subjects over Realistic® stereo headphones (model Nova 65).

Speech intelligibility levels were varied using a chopping circuit designed by HEL and described in detail in Peters and Garinther (1990). This chopping circuit removes portions of the speech signal by chopping the signal for varying durations. The chopping circuit gated the speech signal at 60 Hz with a duty cycle variable from 0% to 95%. The circuit also adds a speech-shaped masking noise (i.e., pink noise) passed through a first order low pass filter of 250 Hz followed by a first order high pass filter of 350 Hz. This masking noise was used in previous studies (Peters & Garinther, 1990; Whitaker et al., 1990) to prevent subjects' (who were tested in pairs) shouted speech from being heard directly. To maintain comparability across the studies, masking noise was employed in the present study as well.

Materials

The target and nontarget items for the auditory Sternberg task and the to-be-identified items for the MRT trials consisted of the 300 stimulus words developed by House et al. (1965) for use in the MRT. These 300 items were categorized into six lists of 50 items each as specified by House et al. Items from List A were used as target and nontarget items in the auditory Sternberg task. The remaining five lists (Lists B through F) were used in the

MRT trials. All stimuli were spoken by a male native English speaker and were digitally recorded (12 bit, 10 kHz sample rate).

Modified Rhyme Test

During the MRT trials, subjects were presented with 50 words, one at a time. Each target word was preceded by a carrier phrase ("The next word is..."). After each word was auditorally presented six alternate targets were presented on a 14-inch (35.6-cm) cathode ray tube (CRT), and subjects indicated which word they thought had just been presented auditorally by pressing a number (1 through 6) on a numeric key pad. After each response was made, the CRT screen was cleared, and 2 seconds later, the carrier phrase and the next target item were presented. The speech intelligibility level obtained during each trial was defined as the percentage of correct responses on that list.

Auditory Sternberg Task

During each auditory Sternberg trial, subjects were presented with two or four items to memorize. The items selected as target (and nontarget) items during each trial were randomly determined and held constant for all subjects. After the target items were presented, subjects could elect to review the items if they wished; otherwise, subjects signaled the experimenter that they had memorized the items, and the trial began.

Each trial consisted of 24 probe items, 12 target and 12 nontarget items, presented in a random order. Subjects used their left hands for responding to the Sternberg items. Subjects were instructed to press the 1 key on a numeric keypad if the item came from the set of memorized target items for that trial (positive probes) and to press the 3 key when the item was not a member of the memorized set. For both the two- and four-item trials, each target item appeared equally often, as did the negative probe items. Subjects were instructed to respond as quickly and accurately as possible in classifying the probe items. The reaction time for each item was defined as the interval from the onset of the probe to the initiation of the key press. The next item in the list was presented immediately after the subject's response. Total trial duration thus somewhat depended on the subjects' reaction times, but trials lasted an average of 3 minutes.

Visual Task

The visual task used in Experiment 1 was the CTS unstable tracking task from Shingledecker et al. (1982). The CTS unstable tracking task is similar to the critically unstable tracking task developed by Jex, McDonnell, and Phatak (1966). The CTS unstable tracking task is designed to place variable demands on human information processing resources involving the execution of rapid and accurate manual responses. In the task, subjects viewed a video screen displaying a fixed target area centered on the screen. A cursor moved vertically from the center of the screen, and the operator tried to keep the cursor centered over the target area by rotary movements of a control knob. The system represented by the task is an inherently unstable one. The operator's input introduced error which was magnified by the system with the result that it became increasingly necessary to respond to the velocity of the cursor movement as well as cursor position. No external forcing function was applied to the tracking loop. The unstable dynamics were simply excited by human tracking remnant and by noise in the controller digitization process. If the subject lost control and the cursor reached the edge of the display, it was automatically reset to display center and the subject continued tracking. The active area of the display was ±9.5 cm.

Subjects performed the tracking task using their right hands. The unstable systems dynamics of the CTS tracking task were a first order divergent element of the form:

$$P_{(s)} = \frac{\lambda}{s - \lambda} e^{-ts}$$

in which λ was selected by the experimenter to vary the manual control work load. In the CTS version of the unstable tracking task, the magnitude of time delay term (t) was determined by Shingledecker et al. (1982) to be no greater than 49 msec. This delay period includes the 21-msec frame time (1000 msec/47 Hz), an 11-msec sample and hold (0.5 by frame time) associated with display generation, and a 17-msec sample and hold associated with the television frame time.

Task difficulty was manipulated by varying the weighting factor (λ) used in the CTS unstable track. In the present experiment, the easy tracking task employed a λ value of 1, and the difficult task used a λ value of 3. The tracking task was performed continuously for 3 minutes.

Procedure

At the beginning of the experimental session, subjects were given three practice trials designed to familiarize them with the tasks. First, a 10-item MRT trial was given. Second, a 12-item (six targets, six nontargets) auditory Sternberg task was presented using a memory set of two items. For the practice MRT and the practice auditory Sternberg task, the speech signals were presented with no chopping and no masking noise added. In all regards other than the trial length and no degradation of the speech signals, these practice trials were identical to the critical trial tasks. Finally, subjects performed a 30-second dual task trial with the appropriate level of difficulty on the tracking task and a two-item memory set with 12 probe items (six targets, six nontargets) for the auditory Sternberg task. After these practice trials were completed, the experimenter answered any questions the subjects had, and the four blocks of critical trials were begun.

Each block of critical trials was identical except that (a) a different level of speech intelligibility was used for each block, (b) different lists were used for each MRT trial, and (c) different targets and nontargets were used for each Sternberg task trial. The order of presentation of the MRT lists was counterbalanced across subjects and across levels of speech intelligibility. In the auditory Sternberg task, assignment of items to the positive set (i.e., memory set) and negative set (i.e., nontargets) was also counterbalanced across subjects and speech intelligibility levels. This latter point is important in that reaction time during the Sternberg task was measured from the onset of the stimulus item. Counterbalancing items across levels of speech intelligibility ensures that any differences in reaction times for the four levels of speech intelligibility cannot be attributed to different stimulus durations.

The first task in each block of trials was a MRT task run with the chopping circuit set the appropriate settings for the desired target level of speech intelligibility. Based on previous research by Peters and Garinther

(1990) and pilot research in SUNY-Binghamton, the duty cycle setting on the chopping circuit for the 80%, 60%, 40%, and 20% speech intelligibility conditions resulted in the speech signals being chopped to 83%, 62%, 31%, and 10.3% of the original signal. Following this initial MRT, subjects performed the remaining six tasks in order as described above. Subjects completed the four blocks of trials in order, with a brief rest being given between blocks 2 and 3. Subjects were thoroughly debriefed at the conclusion of block 4.

Results

Several dependent variables were used to assess performance levels in the single and dual task conditions. Nominal speech intelligibility levels were measured by computing the mean percentage of correct responses on the initial and final MRT given within each block of trials. For the auditory Sternberg task, accuracy and reaction time were the dependent variables of interest while for the unstable tracking task, the authors used the average absolute tracking error (in pixels). This value is computed by the CTS software using Equation 1.

$$1.\quad \text{Average Absolute Tracking Error} = \sum_{i=1}^{n} | e_i | /n$$

in which n = the number of 21 msec time frames in the 180-second trial and e_i = the absolute error of the cursor during time frame i. Note that for the unstable tracking task, lower average absolute tracking error scores correspond to better performance levels.

Speech Intelligibility

The first question the authors asked was whether the chopping circuit settings produced the desired levels of speech intelligibility. Table 1 presents the mean intelligibility levels (calculated as the mean of the first and second MRT) for each intelligibility level condition. Replicating Whitaker, Peters, and Garinther (1989), the observed intelligibility levels were close to the desired intelligibility levels, indicating that the HEL chopping circuit was quite successful at producing the desired speech intelligibility levels.

Table 1

Experiment 1: Mean Percentage Correct Responses During the MRT for the Four Intelligibility Conditions

Unstable tracking difficulty level	Intelligibility condition 80%	60%	40%	20%
Easy	78	70	50	26
Difficult	71	61	46	23

A 2 (unstable tracking difficulty: easy versus difficult) x 4 (intelligibility level: 80%, 60%, 40%, 20%) mixed factor analysis of variance (ANOVA) of these data indicated a significant effect of intelligibility level, $F(3,78) = 407.1$, $MSe = 33.7$ $p < .0001$. (Unless otherwise noted all, effects called significant had $p < .05$.) Newman-Keuls pairwise comparisons indicated that the observed intelligibility level in each condition differed significantly ($p < .01$) from all other conditions. Neither the main effect of unstable tracking difficulty level nor the unstable tracking difficulty x intelligibility level was significant.

Single Task Trials

The next set of issues addressed concerns the effectiveness of the task difficulty manipulations for the unstable tracking task and the auditory Sternberg task. The most straightforward measure of whether the easy and difficult versions of these two tasks actually differed from one another comes from the performance levels observed in the single task trials.

Unstable Tracking

Figure 1 presents the mean error scores for the single task trials in the easy and difficult unstable tracking conditions. Replicating previous studies that have employed the CTS unstable tracking task (e.g., Schlegal & Schingledecker, 1985), there were large differences in the performance levels for these two tracking tasks. Furthermore, there was no effect of the level of speech intelligibility for the block of trials within which this single task trial was run. The lack of an effect of speech intelligibility is, of course, not surprising since subjects were not performing an auditory task during the single task unstable tracking task trial.

These observations were supported by the results of a 2 (tracking task difficulty: easy versus difficult) x 4 (speech intelligibility levels: 80%, 60%, 40%, 20%) mixed factor ANOVA. There was a significant main effect of tracking task difficulty, $F(1,26) = 143.2$. $MSe = 77.3$. Neither the main effect of Speech Intelligibility nor the Tracking Task Difficulty x Speech Intelligibility interaction was significant (F's < 1.0).

Auditory Sternberg

Performance of the single task auditory Sternberg trials was examined by comparing recognition accuracy and response latency (see Table 2). These data were analyzed using separate 2 (tracking task difficulty: easy versus difficult) x 2 (number of Sternberg target items: 2 versus 4) x 4 (intelligibility level: 80%, 60%, 40%, 20%) mixed factorial ANOVAs, one for each performance measure. These two analyses produced completely parallel results. As expected, performance levels were affected by the level of intelligibility and by the number of target items.

Table 2 presents the mean percent correct responses and mean reaction time for the two- and four-target item trials conducted at each of the intelligibility levels. There was a significant main effect of number of target items for both the accuracy, $F(1,26) = 49.7$, $MSe = 57.1$, $p < .001$ and reaction time measures, $F(1,26) = 35.2$, $MSe = .034$, $p < .001$. There was no effect of tracking task difficulty for either the percent correct measure, $F(1,26) = 2.00$, $MSe = 211.1$, or the reaction time measure $F(1,26) = 2.45$, $MSe = .246$. There were no other significant main effects in either analysis, nor were there any significant interactions.

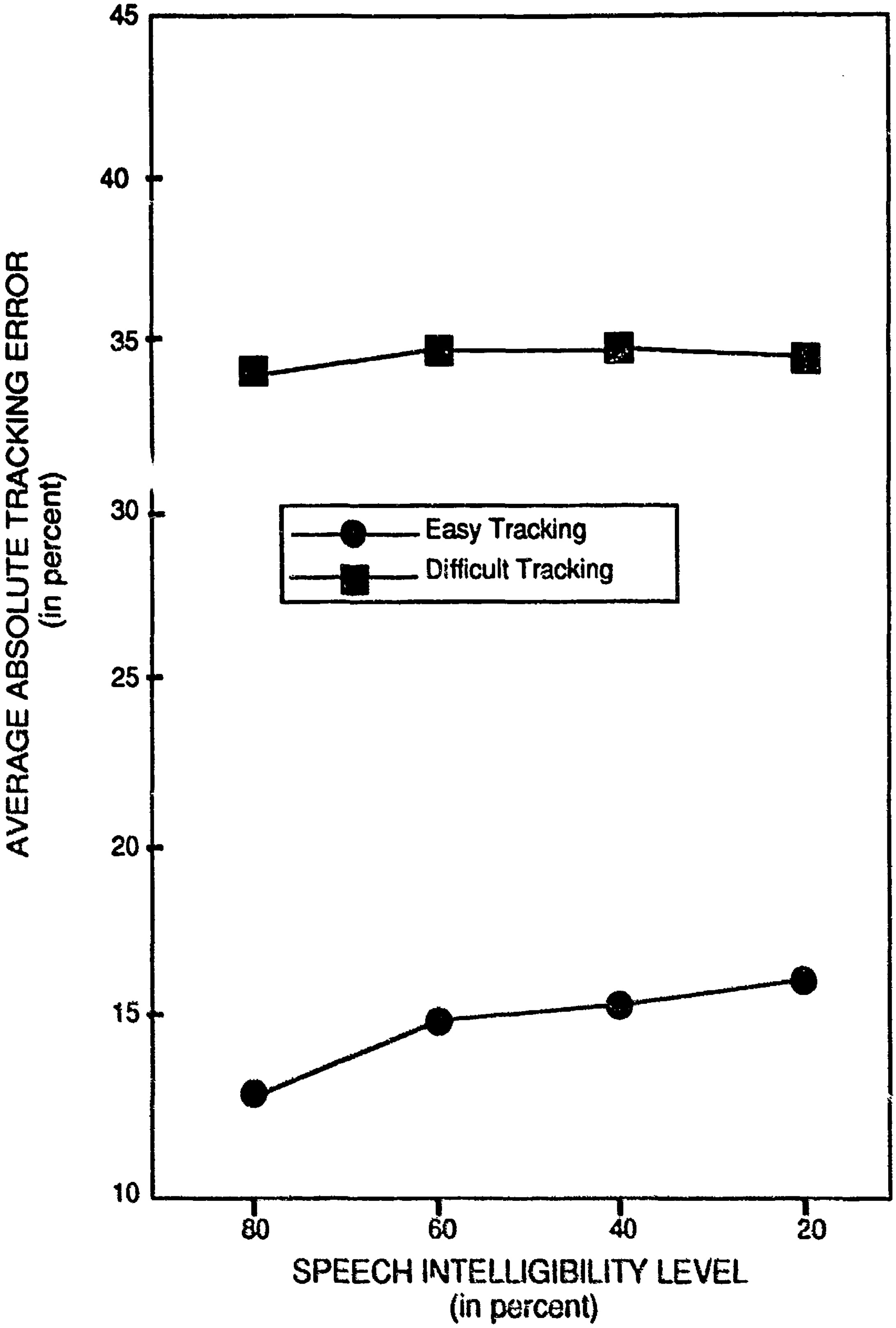

Figure 1. Experiment 1: Performance levels in the unstable tracking single task trials for the easy and difficult unstable tracking conditions.

Table 2

Experiment 1: Mean Percent Correct Responses (upper panel) and Mean Reaction Time (in seconds, lower panel) for the Single Task Auditory Sternberg Trials

Number of target items	Intelligibility condition 80%	60%	40%	20%	Mean
	Mean Percent Correct Response				
2	99	98	91	60	85
4	95	98	91	60	83
	Mean Reaction Time				
2	1.148	1.165	1.327	1.574	1.303
4	1.276	1.370	1.493	1.658	1.449

The data from the single task trials thus indicate that the manipulations of speech intelligibility and tracking difficulty had the desired effect of producing large differences in the observed performance levels. As such, these two tasks allow us to examine whether changes in speech intelligibility level will impact performance in the tracking task when the auditory and visual tasks are performed concurrently. Before the effect of speech intelligibility on the tracking task is considered, however, performance levels in the auditory Sternberg task must be examined to confirm that intelligibility manipulation affected performance in this task, and that varying the number of target items impacted task difficulty.

Dual Task Trials

Auditory Sternberg

The accuracy and response latency data from the auditory Sternberg task in the dual task trials were analyzed using separate 2 (number of target items: 2 versus 4) x 4 (intelligibility: 80%, 60%, 40%, 20%) x 2 (tracking task difficulty: easy versus difficult) mixed factorial ANOVAs, one for each dependent variable. Table 3 presents the mean percent correct responses and mean reaction times for the two and four alternate auditory Sternberg trials conducted at the four intelligibility levels. As with the single task trials, speech intelligibility level exerted a significant effect on both response accuracy, $F(3,78) = 158.2$, $MSe = 98.4$, and response latency, $F(3,78) = 14.9$, $MSe = .067$. The number of target items also affected both accuracy $F(1,26) = 5.3$, $MSe = 85.8$, and reaction time, $F(1,26) = 7.6$, $MSe = .018$, indicating that subjects were faster and more accurate in the two-target condition than the four-target condition.

Table 3

Experiment 1: Mean Percent Correct Responses (upper panel) and Mean Reaction Time (in seconds lower panel) for the Dual Task Auditory Sternberg Trials

Number of target items	Intelligibility condition 80%	60%	40%	20%
	Mean Percent Correct Responses			
2	95	94	89	60
4	93	94	82	58
	Mean Reaction Time			
2	1.169	1.244	1.238	1.535
4	1.245	1.287	1.362	1.492

There were two additional significant effects in these analyses. First, for the response accuracy measure, there was a significant main effect of visual task difficulty. When performing the auditory Sternberg task in conjunction with the easy unstable tracking task, subjects made 85.9% correct responses in contrast to 79.9% correct with the difficult version of the tracking task, $F(1,26) = 12.8$, $MSe = 157.6$. However, the difficulty of the unstable tracking task did not interact with speech intelligibility level, $F < 1.0$.

The second significant effect was that for the reaction time measure there was a significant Number of Target Items x Intelligibility interaction, $F(3,78) = 3.4$, $MSe = .02$. Two points must be noted regarding this interaction. First, the proportion of correct and incorrect responses varied across the levels of speech intelligibility and the number of target items, and it has been shown in studies using visual Sternberg tasks that the response latency for negative items is typically longer than that obtained with positive items (e.g., Sternberg, 1967). In the present experiment, it was not possible to provide an accurate measure of response latency separately for correct and incorrect responses since when accuracy was either very high or very low, there are too few data points for, respectively, the incorrect responses and the correct responses. Second, the form of the interaction is unusual in that the difference in response latency between the difficult and easy unstable tracking conditions was .076, .043, .124, and -.043, for the 80%, 60%, 40%, and 20% speech intelligibility conditions, respectively. Given these two facts, it is unclear how to best interpret this interaction.

Overall, the results from the auditory Sternberg dual task condition indicated that performance was impacted by the number of target items and by the level of speech intelligibility. The main effect of unstable tracking task difficulty on the reaction time measure suggests that the tracking task difficulty did vary with the overall joint task difficulty. Importantly, in the auditory Sternberg task, the tracking task difficulty did not interact with speech intelligibility level, suggesting that either (a) the tracking task and the auditory Sternberg task tapped into different resources, or (b) the two tasks tapped into the same resources, but the resulting performance decrement was reflected only in the unstable tracking task. Was there any effect of speech intelligibility on tracking task performance?

Unstable tracking

Figure 2 presents the mean absolute tracking error scores for the easy and difficult unstable tracking tasks for each of the four intelligibility measures. Replicating the single-task data, there was a significant effect of tracking task difficulty, $F(1,26) = 156.8$, $MSe = 187.7$. More importantly, however, there was no evidence of a main effect of speech intelligibility, nor did intelligibility interact with any other variable (all F's < 1.0).

Analyses of these data revealed that there was a significant main effect of the number of target items in the auditory Sternberg task, $F(1,26) = 8.5$, $MSe = 40.3$, as well as a significant Tracking Task Difficulty x Number of Target Items interaction, $F(1,26) = 4.3$, $MSe = 40.3$. The source of these effects lies primarily in the difficult tracking task condition; in the easy tracking task condition, there was a nonsignificant difference between the two- and four-target item conditions (16.4 versus 15.7), $F < 1.0$, while there was a significant difference in the difficult tracking task (41.1 versus 36.8), $F(1,26) = 12.5$, $MSe = 40.3$. Thus, subjects performed more poorly during the tracking task when there were two target items than during the Sternberg task than when there were four items. This difference may reflect a tradeoff in task emphasis in that as the Sternberg task increased in difficulty, subjects elected to focus more on the tracking task. Note, however, that there is no evidence of such a change in task emphasis as a function of speech intelligibility level. This is consistent with the hypothesis that speech intelligibility impacts resources not shared by the unstable tracking task.

EXPERIMENT 2

The results of Experiment 1 indicate that varying the level of speech intelligibility had no effect on the performance levels in a concurrent visual tracking task. This result can be interpreted within the multiple resource framework as indicating that the auditory and speech tasks tapped into different resources, with the tracking task loading on visual encoding and manual responding and the auditory Sternberg task loading on auditory encoding and central processes. Experiment 2 was conducted to determine if a visual task that places demands on central processing (e.g., working memory, decision making) resources, which are presumably involved in performing the auditory Sternberg task, would be affected by changes in speech intelligibility. Toward this goal, the authors conducted a replication and extension of Experiment 1, this time using the spatial processing task from the CTS. In the spatial processing task, subjects were presented with pairs of histograms, one presented in a vertical orientation and the second presented at a 0°, 90°, 180°, or 270° orientation. The subjects' task was to decide if the two histograms were identical in overall shape, regardless of the orientation of the second, comparison stimulus.

This spatial processing task has several important features. First, task difficulty can be manipulated by varying the number of bars comprising the histogram. Second, the spatial or imaginal processing required to perform the task is similar to many tasks performed by mobile vehicle crews (e.g., following directions, locating targets in the field of view). Finally, this task requires subjects to encode or store stimuli and to compare a subsequent stimulus to the initial one. This is similar to tasks performed in operational settings (e.g., visually identifying friend versus foe), and it also has many of the cognitive processes involved in performing the auditory Sternberg task.

If the multiple resource theory provides a viable framework for examining the impact of speech intelligibility on visual task performance, changes in speech intelligibility would be expected to negatively affect performance levels in the spatial processing task. Alternatively, it is possible that changes in speech intelligibility will not interfere with the performance of visual tasks; if this is the case, the results of Experiment 2 should parallel those from Experiment 1.

Subjects, Design, Apparatus

Twenty-eight subjects were selected from the same source as in Experiment 1. The experimental design and apparatus were the same as in Experiment 1 with one exception. The visual task employed was the CTS spatial processing task in which subjects viewed computer-generated pairs of histograms presented on the Commodore screen. Each histogram bar could assume any of six arbitrary heights; the first histogram in each pair appeared in the vertical orientation and was labeled with a 1. The second histogram was presented at either the same vertical orientation, rotated 90° left or right, or rotated 180° and was labeled with a 2. Approximately 50% of the comparison histograms were the same shape as the original histogram, and the remaining histogram's shape was altered slightly. Each "unit" of histogram bar height was approximately 0.85 cm. Bars were 0.5 cm wide and separated by 0.4-cm spaces. Subjects indicated their responses by pressing one of two response keys (for same or different) on a four-key response box. The left key was used to respond "same" and the right key to respond "different."

Difficulty level in the spatial processing task was manipulated by varying the number of bars comprising the histograms. For subjects in the easy version of the spatial processing task, each histogram contained four bars, while for the difficult condition, there were six bars. The first target stimulus was presented for 3.0 seconds followed by a short pause. The comparison stimuli were presented for a maximum of 2.5 seconds in the easy condition and 3.5 seconds in the difficult condition. The comparison stimulus was removed as soon as the subject made a response, and the screen remained blank until the next target stimulus was presented. The spatial processing task was performed for 3 minutes. In all regards other than those associated with the visual task, Experiment 2 was an exact replication of Experiment 1.

Results

Speech Intelligibility

Table 4 presents the mean intelligibility levels for the mean of the first and second MRT for each intelligibility level. Replicating Experiment 1 there were large differences among the observed intelligibility levels for the four conditions, and these intelligibility levels were close to the target intelligibility levels for each condition.

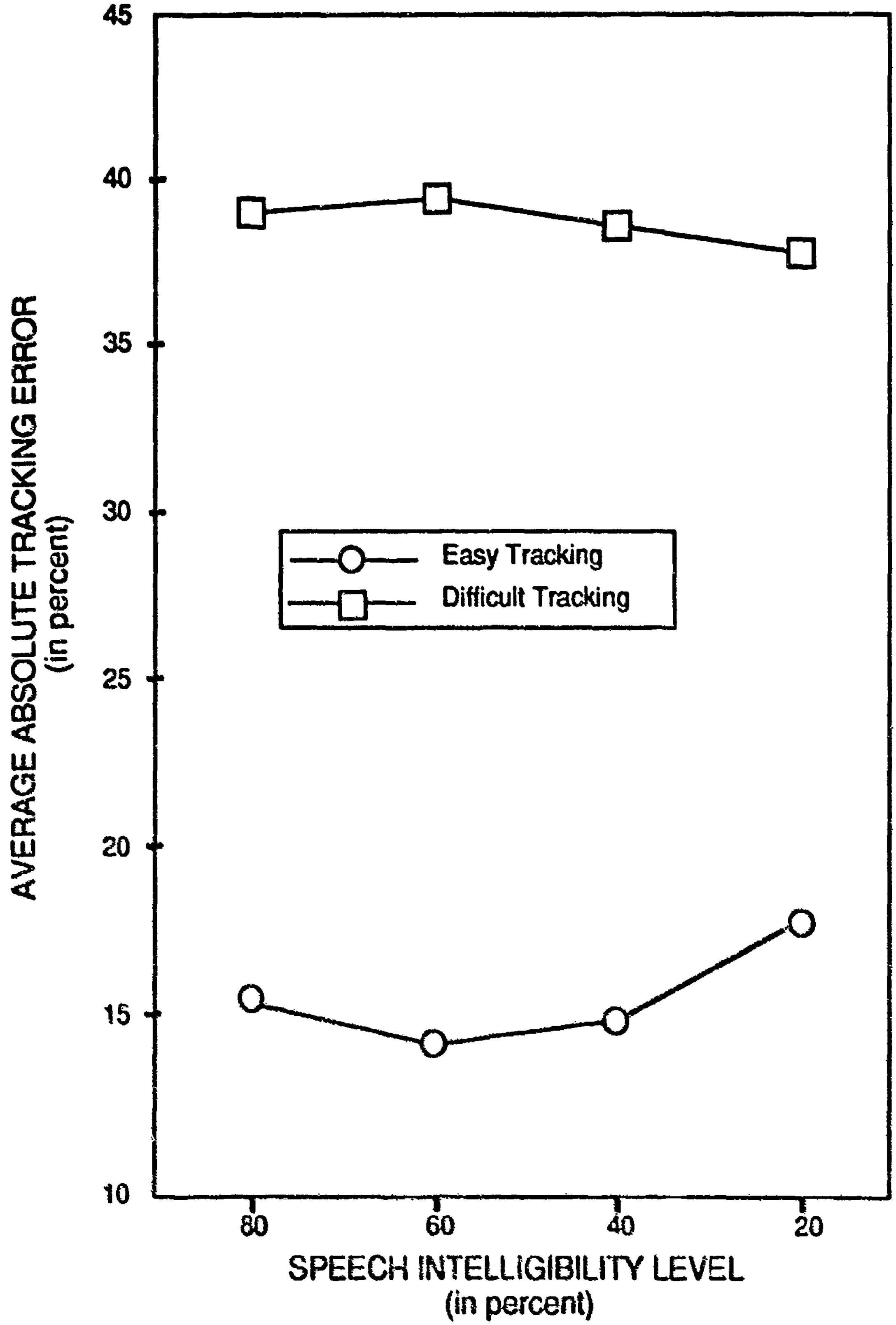

Figure 2. Experiment 1: Performance levels in the unstable tracking dual task trials for the easy and difficult unstable tracking conditions.

Table 4

Experiment 2: Mean Percentage Correct Responses on the MRT for the Four Intelligibility Conditions

	Intelligibility Condition			
Spatial processing difficulty level	80%	60%	40%	20%
Easy	75	63	47	25
Difficult	73	66	43	22

Furthermore, the intelligibility levels were comparable for the easy and difficult spatial processing conditions.

These conclusions were supported by the results of a 4 (intelligibility level) x 2 (spatial processing task: easy versus difficult) mixed factor ANOVA. There was a main effect of speech intelligibility level, $F(3,78) = 360.8$, $MSe = 38.9$, $p < .0001$. Neither the main effect of spatial processing difficulty or the Spatial Processing Difficulty x Intelligibility interaction were significant, F's < 1.65. Post hoc Newman-Keuls tests indicated that each Intelligibility Level was significantly different from all other levels, p's $< .01$.

Single Task Trials

Spatial Processing

Figure 3 presents the mean percent correct responses in the single task spatial processing trials. As indicated in Figure 3, subjects were more accurate in the easy spatial processing condition than in the difficult condition, $F(1,26) = 10.97$, $p < .01$. In addition, speech intelligibility also affected performance levels, $F(3,78) = 2.7$, $p = .049$. Although the Spatial Processing Difficulty Level x Speech Intelligibility interaction was not significant, $F < 1.0$, visual inspection of the data suggests that the difficult spatial processing condition was more affected by changes in speech intelligibility than was the easy spatial processing condition. Simple effects tests revealed that there was a significant effect of speech intelligibility for the difficult spatial processing task ($p = .03$), but not for the easy spatial processing task.

Response latency in the spatial processing task was slightly but not significantly, $p = .26$, longer in the difficult spatial processing condition than in the easy spatial processing condition (1099 versus 1004 msec, respectively). There was no evidence of speech intelligibility affecting response times in the spatial processing tasks, with mean reaction times of 1063, 1071, 1040, and 1034 msec for the 20%, 40%, 60%, and 80% speech intelligibility conditions, respectively.

Auditory Sternberg

The mean percent correct responses and reaction times for the two and four alternate single task Sternberg trials are presented in Table 5. Replicating Experiment 1, both accuracy and reaction time were affected by the level of speech intelligibility, $F(3,78) = 103.8$, $MSe = 157.2$, and $F(3,78) =$

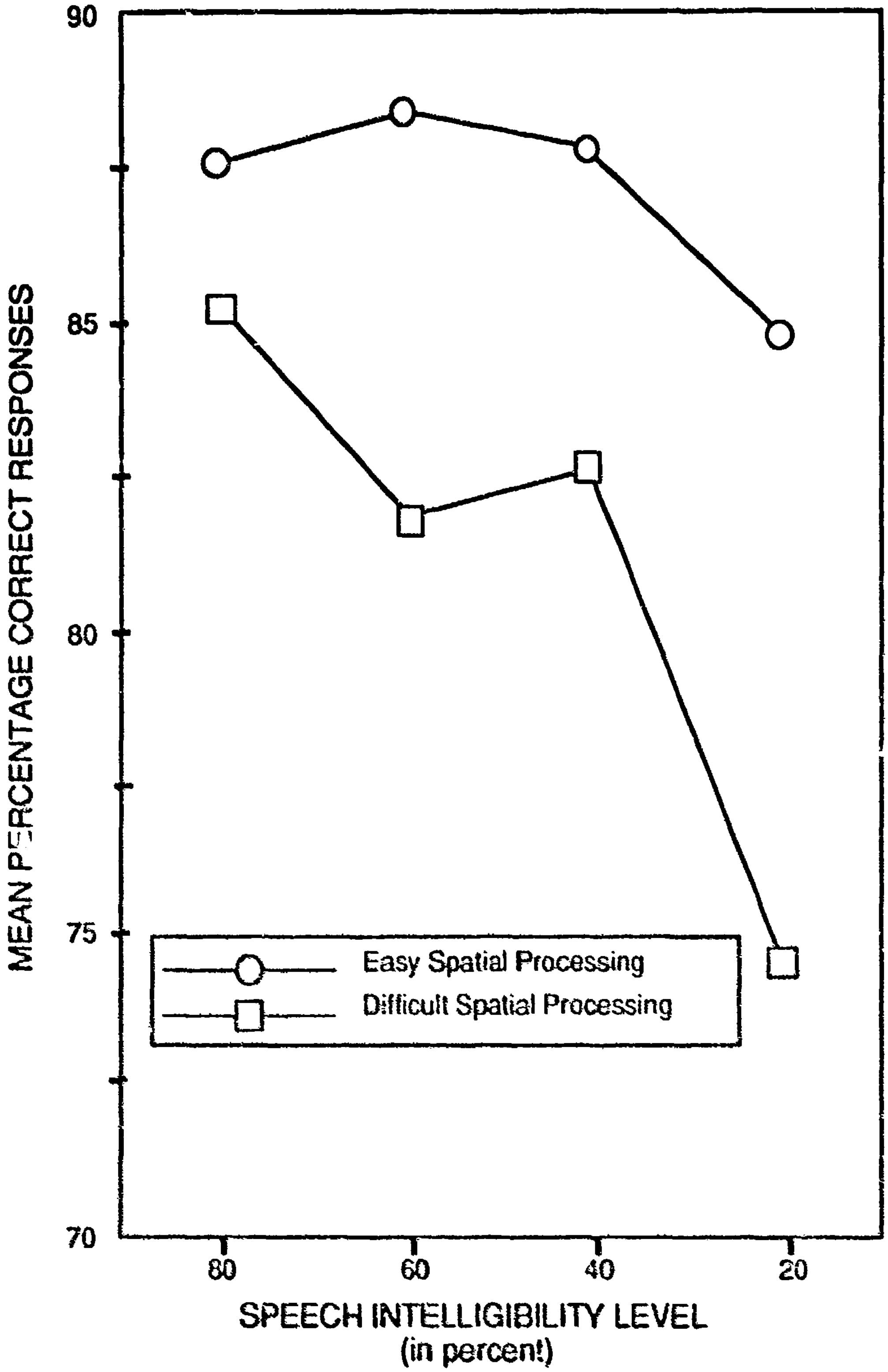

Figure 3. Experiment 2: Mean percent correct responses in the spatial processing single task trials for the easy and difficult spatial processing conditions.

Table 5

Experiment 2: Mean Percent Correct Responses (upper panel) and Mean Reaction Time (in seconds lower panel) for the Single Task Auditory Sternberg Trials

	Intelligibility Condition			
Number of target items	80%	60%	40%	20%
	Mean Percent Correct Responses			
2	95	96	82	60
4	91	90	78	55
	Mean Reaction Time			
2	1.128	1.219	1.527	1.650
4	1.263	1.352	1.430	1.590

13.4, MSe = .152, respectively. The number of target items affected subjects' accuracy, with 83.4% correct responses in the two alternate conditions and 78.6% correct responses in the four alternate conditions, $F(1,26) = 8.2$, MSe = 151.0. There were no other significant effects for either the percent correct or the reaction time measures.

Dual Task Trials

Auditory Sternberg

Table 6 presents the mean percent correct responses and mean reaction times for the two and four alternate auditory Sternberg trials. These data were analyzed with separate 2 (number of target items: 2 versus 4) x 4 (intelligibility: 80%, 60%, 40%, 20%) x 2 (visual task difficulty: easy versus difficult) mixed factorial ANOVAs, one for each dependent variable. Speech intelligibility level affected both response accuracy, $F(3,78) = 172.6$, MSe = 91.4, and response latency, $F(3,78) = 3.19$, MSe = .126. There was also a significant main effect of the number of target items in the accuracy data, with 82.4% correct responses in the two alternate condition versus 76.8% current responses in the four alternate condition, $F(1,26) = 21.6$, MSe = 82.4, $p < .001$. Finally, the Number of Target Items x Speech Intelligibility interaction did not achieve statistical significance, $F(3,78) = 2.52$, MSe = 61.31, $p = .06$. Note, however, that the Number of Target Items x Speech Intelligibility Level was significant in Experiment 1. Simple effects tests revealed that in Experiment 2, this interaction is attributable to the fact that there was no significant difference between the two and four alternate conditions at 60% intelligibility ($p > .20$) while there was either a significant or marginally significant (all p's $< .09$) difference at the other three intelligibility levels.

Table 6

Experiment 2: Mean Percent Correct Responses (upper panel) and Mean Reaction Time (in seconds lower panel) for the Dual Task Auditory Sternberg Trials

Number of target items	Intelligibility condition 80%	60%	40%	20%
	Mean Percent Correct Responses			
2	93	95	83	58
4	91	86	76	54
	Mean Reaction Time			
2	1.628	1.613	1.686	1.828
4	1.637	1.651	1.729	1.793

Spatial Processing

Figures 4 and 5 present the mean percent correct responses and the mean reaction times for the easy and difficult spatial processing conditions across four levels of speech intelligibility. For the percent correct measure, there was a significant difference between the easy and difficult conditions, 81.7% versus 75.6%, $F(1,26) = 10.6$, $MSe = 194.7$. More importantly, there was a significant main effect of speech intelligibility, $F(3,78) = 2.97$, $MSe = 126.8$. Averaging performance across the easy and difficult spatial processing task, the mean percent correct responses for the 80%, 60%, 40%, and 20% speech intelligibility levels were 79%, 80%, 79%, 75%, respectively. Thus, performance levels were consistent across the 80%, 60%, and 40% intelligibility levels, with accuracy decreasing only in the 20% speech intelligibility level condition. There were no further significant main effects or interactions.

For the reaction time measures, the only noticeable finding was that subjects in the difficult task were slightly but only marginally, ($F(1,26) = 2.65$, $MSe = .408$, $p = .11$), slower than in the easy spatial processing task. There was no indication that speech intelligibility level affected reaction time ($F < 1.0$). No other effects approached significance.

GENERAL DISCUSSION

There were several important empirical findings in the present research. From a methodological perspective, the data from the MRT and auditory Sternberg task indicate that the chopping circuit used in this research was quite effective in producing the desired target level of speech intelligibility. These results, along with the findings of Whitaker et al. (1990) and Peters and Garinther (1990), indicate that the chopping circuit provides a valuable research tool.

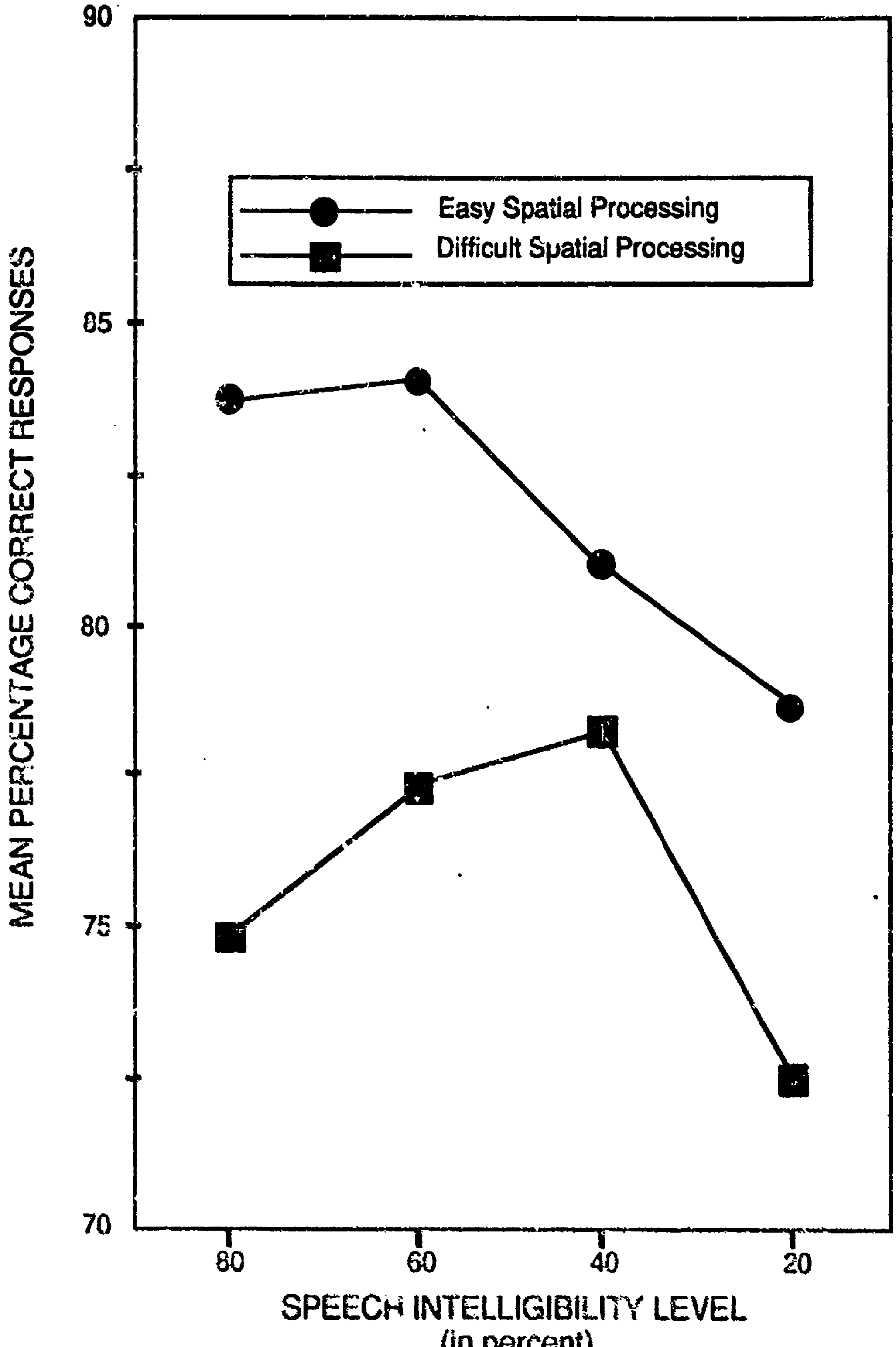

Figure 4. Experiment 2: Mean percent correct responses in the spatial processing dual task trials for the easy and difficult spatial processing conditions.

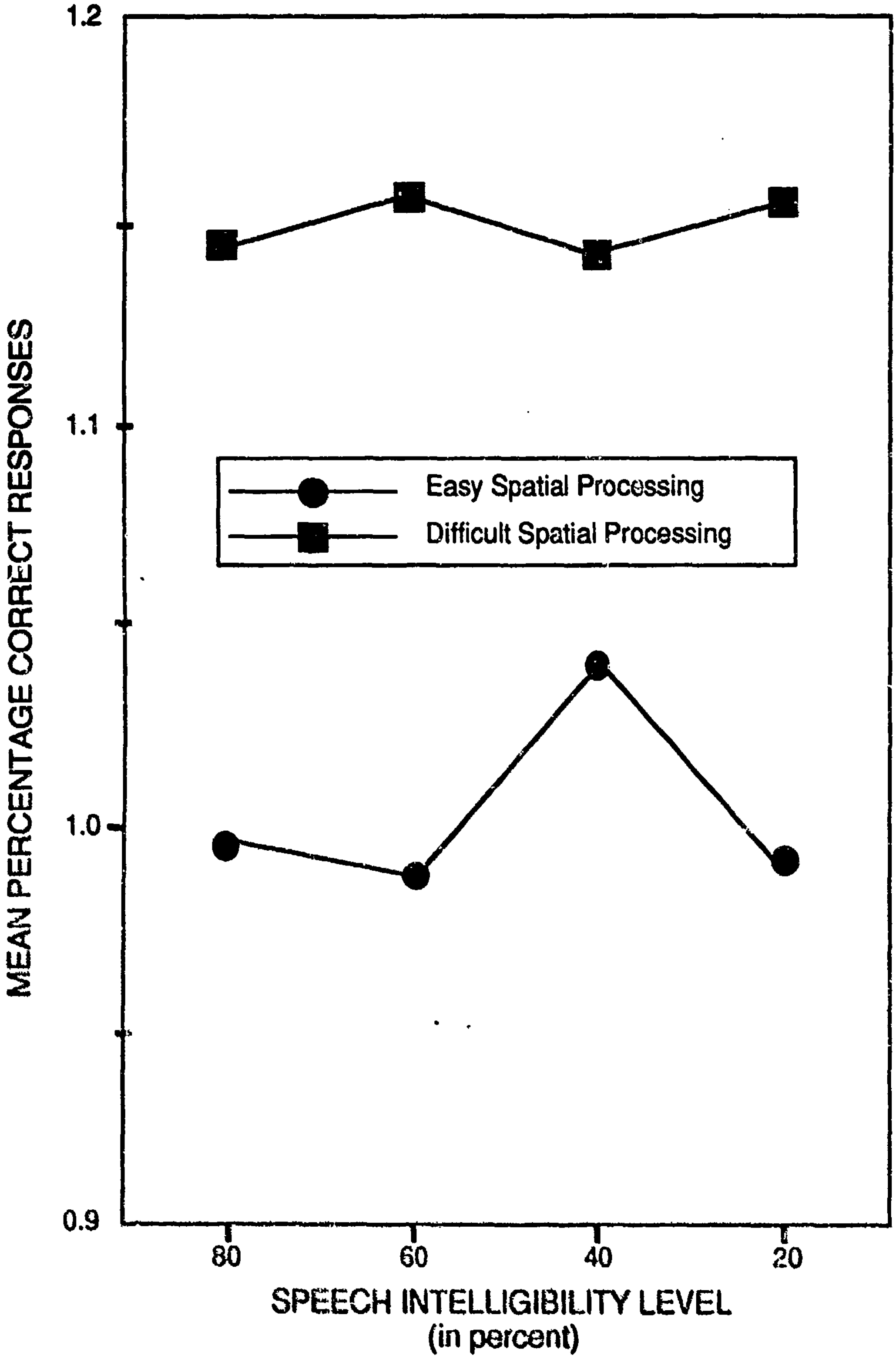

Figure 5. Experiment 2: Mean reaction times in the spatial processing dual task trials for the easy and difficult spatial processing conditions.

A second major finding was that changes in speech intelligibility impacted performance levels in a visual task that required the use of short term memory and decision making (Experiment 2) but did not affect performance in a visual tracking task that presumably loads perceptual motor resources (Experiment 2). These data indicate that (a) the dual task methodology represents a useful research tool for examining the perceptual-cognitive-motor processes impacted by degraded speech, and (b) Wickens' multiple resource model is a viable explanatory framework for describing the performance changes precipitated or caused by degrading speech communication levels. Based on the present results, one would expect to see degraded speech resulting in changes in performance levels in tasks requiring central processing but not in perceptual motor tasks. The findings reported here also suggest that the multiple resource theory interpretation of the selective performance deficits noted by Peters and Garinther (1990) seems reasonable.

Several other points are noted regarding the present research. First, the data suggest that the impact of degraded speech intelligibility is greatest at intelligibility levels of less than 40%. This indicates that further research could profit by concentrating more closely on performance levels obtained with intelligibility levels of less than 50%. Above 50% intelligibility, there is relatively little impact on task performance.

Second, these results indicate clearly that selective patterns of cross-modal interference (i.e., auditory-visual) are produced by varying intelligibility levels. As such, it is important that system designers consider speech intelligibility criteria levels as not simply affecting performance of communications tasks but rather as affecting overall operator performance.

Third, the results obtained in these two experiments suggest that it may be possible to develop a battery of tasks that will allow researchers to identify individuals who are likely to be severely affected by lower intelligibility levels. As Hunt, Pellegrino, and Yee (1989) have documented, large individual differences exist in peoples' abilities to selectively attend to different tasks. Furthermore, previous research has shown that performance of laboratory tasks can be effectively used to predict performance levels in real world tasks such as flying aircraft (e.g., Gopher & Kahneman, 1971) or driving automobiles (e.g., Avolio, Kroeck, & Panek, 1985; Kahneman, Ben-Ishai, & Lotan, 1973). Although there have been some efforts to design a test battery to assess individual differences in information processing (e.g., Avolio, Alexander, Barrett, & Sterns, 1981), most of this research has focused on the use of visual information processing tasks. The present results indicate that the scope of these efforts needs to be expanded to include both auditory and visual tasks. It would be of considerable use if an auditory test battery could be developed that would identify individuals whose performance is likely to suffer when intelligibility levels fall.

Finally, the results of the present experiments illustrate that the conceptual approach taken in this research has considerable potential applicability. It is important, however, that additional investigations be conducted to determine the generalizability of these results. Future research could vary, for example, the types of visual and/or auditory tasks employed in a dual task procedure. It would also be of great importance to examine performance levels in more complex task environments in which operators are required to perform several tasks at the same time. Such a research approach would help to demonstrate the applicability of the present results to more realistic analog tasks.

REFERENCES

Allport, D. A., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single channel hypothesis. Quarterly Journal of Experimental Psychology, 24, 225-235.

Avolio, B. J., Alexander, R. A., Barrett, G. V., & Sterns, H. L. (1981). Designing a measure of selective attention to assess individual differences in information processing. Applied Psychological Measurement, 5, 29-42.

Avolio, B. J., Kroeck, K. G., & Panek, P. E. (1985). Individual differences in information-processing ability as a predictor of motor vehicle accidents. Human Factors, 27, 577-587.

Broadbent, D. (1958). Perception and communication. Oxford: Permagon.

Brooks, L. R. (1968.). Spatial and verbal components of the act of recall. Canadian Journal of Psychology, 22, 349-368.

Carswell, C. M., & Wickens, C. D. (1985). Lateral task segregation and the task-hemispheric integrity effect. Human Factors, 27, 695-700.

Damos, D. (1985). The relation between the Type A behavior pattern, pacing, and subjective workload under single and dual task conditions. Human Factors, 27, 675-680.

Department of Defense (1989). Human engineering design criteria for military systems, equipment and facilities (MIL-STD-1472D). Washington, DC: Author.

Gopher, D., & Kahneman, D. (1971). Individual differences in attention and the prediction of flight criteria. Perceptual and Motor Skills, 33, 1335-1342.

Hanson, R. H., Payne, D. G., Shively, R. J., & Kantowitz, B. H. (1981). Process control simulation research in monitoring analog and digital displays. Proceedings of the Human Factors Society 25th Annual Meeting, 154-158.

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

Hunt, E., Pellegrino, J. W., & Yee, P. L. (1989). Individual differences in attention. In G. H. Bower (Ed.), The psychology of learning and motivation. New York: Academic Press.

Jex, H. R., McDonnell, J. D., & Phatak, A. V. (1966). Critical tracking task for manual control research, IEEE Transactions on Human Factors Engineering, HFE-7, 138-145.

Kahneman, D. (1973). Attention and effort. Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., Ben-Ishai, R., & Lotan, M. (1973). Relation of a test of attention to road accidents. Journal of Applied Psychology, 58, 113-115.

Kerr, B. (1973). Processing demands during mental operations. Memory and Cognition, 1, 401-412.

Navon, D., & Gopher, D. (1979). On the economy of the human information processing system. Psychological Review, 86, 214-255.

Norman, D., & Bobrow, D. (1975). On data-limited and resource limited processes. Cognitive Psychology, 7, 44-64.

Ogden, G. D., Levine, J. M., & Eisner, E. J. (1979). Measurement of workload by secondary tasks. Human Factors, 21, 529-548.

Peters, L., & Garinther, L. J. (1990). The effects of speech intelligibility on crew performance in an M1A1 tank simulator (Technical Memorandum 11-90). Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratory.

Proctor, R. W., & Proctor, J. D. (1979). Secondary task modality, expectancy, and the measurement of attentional capacity. Journal of Experimental Psychology, Human Perception and Performance, 5, 610-624.

Roediger, H. L., Knight, J. L., & Kantowitz, B. H. (1977). Inferring decay in short-term memory: The issue of capacity. Memory and Cognition, 5, 167-176.

Schlegel, R. E., & Shingledecker, C. A. (1985). Training characteristics of the criterion task set workload assessment battery. Proceedings of the Human Factors Society 29th Annual Meeting, 770-773.

Shaffer, L. H. (1975). Multiple attention in continuous verbal tasks. In P.M.A. Rabbit and S. Dornic (Eds.). Attention and performance V. New York: Academic Press.

Shingledecker, C. A., Crabtree, M. S., & Acton, W. H. (1982). Standardized tests for the evaluation and classification of workload metrics. Proceedings of the Human Factors Society 26th Annual Meeting, 648-651.

Sternberg, S. (1969). The discovery of processing stages: Extension of Donder's method. In W.G. Koster (Ed.), Attention and performance II, Amsterdam: North Holland.

Tyler, S., Hertel, P., McCallum, M., & Ellis, H. (1979). Cognitive effort and memory. Journal of Experimental Psychology: Human Learning and Memory, 5, 607-617.

Whitaker, L. A., Peters, L. J., & Garinther, G. (1989). Tank crew performance: Effects of speech intelligibility on target acquisition and subjective workload assessment. Proceedings of the Human Factors Society 33rd Annual Meeting, 614-617.

Whitaker, L. A., Peters, L. J., & Garinther, G. (1990). Effect of communication degradation on military crew task performance (Discussion Paper 90-49). Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratory.

Wickens, C. D. (1980). The structure of attentional resources. In P. S. Nickerson (Ed.), Attention and performance VIII, Hillsdale, NJ: Erlbaum.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman and R. Davies (Eds.), Varieties of attention. New York: Academic Press.