HUMAN FACTORS, 1994, 36(4), 700–717

Tactile Stimulation of the Human Head for Information Display

KIRBY GILLILAND[1] *and* ROBERT E. SCHLEGEL, *University of Oklahoma, Norman, Oklahoma*

A series of three studies was conducted to explore the use of tactile stimulation or light tapping of the human head to inform a pilot of possible threats or other situations in the flight environment. Study 1 confirmed that subjects could achieve 100% detection of the tactile stimuli. Localization performance, measured in Study 2, depended on the number of different stimulus sites and ranged from 93% accuracy for 6 sites to 47% accuracy for 12 sites across the parietal meridian of the head. In Study 3 we investigated the effect of performing the localization task simultaneously with a dual memory/tracking task or an air combat simulation task. These studies demonstrated that tactile information display could be an integral contributor to improved situation awareness, but not without cost to other task performance. The results of Study 3 were also examined with reference to popular models of attention and workload.

INTRODUCTION

In recent years there have been numerous technological advances designed to improve pilot performance. Collectively, the resulting electronic systems have saturated the pilot with information and, paradoxically, may now be detrimental to performance. This problem is magnified by increased aircraft speeds, which reduce the time available to process information. Consequently, the pilot's immediate knowledge of events in the cockpit environment is crucial because the difference between success and failure can be measured in seconds.

[1] Requests for reprints should be sent to Kirby Gilliland, Department of Psychology, University of Oklahoma, Norman, OK 73019.

The information saturation of the cockpit environment has led to an increased interest in situation awareness. In aviation, *situation awareness* refers to the pilot's perception and comprehension (i.e., internal model) of elements in the environment at any specific point in time (Endsley, 1988). Heightened interest in measuring and improving situation awareness has prompted researchers to explore new methods of presenting information so that it can be processed rapidly and easily.

This article summarizes a series of three studies that tested the feasibility of presenting spatial information through the use of a helmet-mounted tactile information display. A recently proposed system suggests the integration of onboard radar systems with cutaneous stimulation of the scalp as a means of conveying critical situation information

© 1994, Human Factors and Ergonomics Society. All rights reserved.

Downloaded from hfs.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/24832

through the pilot's helmet (system patented by Meir Morag; U.S. Patent Number 4,713,651, December 15, 1987). Tactile stimulators inside the pilot's helmet are used to signal the spatial locations of targets or threats. The specific point of the scalp that is stimulated represents the pilot-referenced angular direction of the event. Amplitude or frequency modulations can be used to represent parameters of the event such as distance or urgency. High-priority spatial information regarding important targets could thus be transmitted to the pilot without using the overburdened visual and auditory channels.

The helmet-mounted tactile target display meets several design guidelines for improving situation awareness (Endsley, 1988). In particular, Endsley suggests the use of auditory or tactile modes to provide information, especially about critical events. Use of the pilot's head as the display surface provides an egocentric view of the environment that pilot's can rapidly relate to their cognitive maps and their orientations within them. Because the tactile helmet would be used simultaneously with existing visual tactile displays, it emphasizes the urgency of the threat environment.

Although there has been considerable research on the perception of tactile stimuli (Geldard, 1972; Heller and Schiff, 1991; Sherrick and Cholewiak, 1986), there have been few attempts to integrate this knowledge within operational settings (Sherrick, 1975), especially with regard to enhancing communication (Geldard, 1974). Notable exceptions are the development of tactile stimulation technology for the visually impaired (Bliss, Katcher, Rogers, and Shepard, 1970; Craig, 1974), attempts to incorporate tactile feedback in vehicle control (Fenton, 1966; Gilson and Fenton, 1974), and a cutaneous tactile communicator sleeve developed by Northrup Corporation (Zlotnik, 1988). The latter device is worn on the forearm and uses tactile stimulators to present information on such aircraft flight parameters as airspeed, angle of attack, altitude, corner velocity, energy state, and shoot cues.

From a theoretical standpoint, tactile display performance may be assessed within the framework of two attention models that have been applied to workload assessment. Wickens (1980, 1992) introduced a popular multiple-resources model that proposes a structural analysis of information processing based on three factors: stages of processing, cerebral hemispheres, and modalities of processing. This model has been reasonably successful in delineating the patterns of task variables that interfere with one another during concurrent task performance. Based on this model, tactile stimulation should not interfere to a large degree with memory processing or psychomotor performance. Alternatively, the cognitive-energetical stage model (Gopher and Sanders, 1984; Sanders, 1983) stresses the role that energetic concepts play in mediating performance and workload. Examples of energetic concepts are arousal, activation, and effort. This model suggests that the arousing characteristics of tactile stimulation may interfere with other modes of communication or performance.

The objective of the series of studies reported here was to explore the use of tactile stimulation of the scalp as a means of presenting information to the subject. Of particular importance were questions concerning (1) whether tactile stimuli presented to the scalp could be reliably detected given the intense stimulus competition in the cockpit environment, (2) the number of sites on the scalp that could be reliably localized, and (3) the detectability and localizability of such sites under conditions of high workload. From a basic research perspective, the first two studies in this series provide information

Downloaded from hfs.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

regarding the tactile responsivity of the scalp, about which only modest amounts of data exist (Shimizu and Wake, 1982; Weber, 1834/1978; Weinstein, 1968). The third study provides information relevant to theories of attention, especially multiple-resource models (Navon and Gopher, 1979; Norman and Bobrow, 1975).

GENERAL METHODS

Tactile Display Equipment

Guardian tubular solenoids (model TP4x16, intermittent duty, 24 V DC) were used to provide the tactile stimuli. The body of each solenoid was 5.0 cm long and 1.3 cm in diameter. The external shaft of the solenoid was 3.1 cm long with a diameter of 1.3 mm at the tip. For each stimulus trial, one of the solenoids was driven by a B & K Precision Model 3011 square-wave function generator at the rate of 4 Hz for a duration of 1 s. For the specific equipment configuration and voltages used in these studies, the impact force of the solenoid ranged from 110 g (0.05 lb) at 15 V to 150 g (0.07 lb) at 20 V. The tips were coated with silicon rubber to provide a modest amount of padding for impact with the subject's scalp. A secondary circuit initiated subject response timing with a Commodore 64 microcomputer. The computer recorded the subject's location response and the response time in milliseconds.

The tactile stimulation helmet was constructed from a modified Bell automobile racing helmet as illustrated in Figure 1. Large portions of the anterior and posterior sections of the helmet were removed, leaving a thin strip of helmet material running from ear to ear along the vertex of the coronal meridian. This strip supported the helmet while it was in place on the head. An aluminum band (61.5 cm long, 2.5 cm wide, and 0.3 cm thick) served as the anchoring platform for the solenoids. The band was bent to approximate the outer contour of the helmet and was attached with bolts and wing nuts at the ear positions on the helmet. Loosening the wing nuts allowed the band to pivot forward and backward and, thus, provided front-to-back adjustment of the solenoids along the sagittal plane of the skull.

A slot (40.0 cm long and 0.6 cm wide) was milled down the center of the mounting band. To mount each solenoid, an L-shaped bracket extending outward from the helmet was first secured to the band with a bolt and wing nut. Loosening the wing nut allowed the L bracket to be positioned along the channel from the left ear to the right ear in the transverse plane. Attached to the L bracket, perpendicular to the helmet surface, was a short section (3.0 cm) of polyvinyl chloride (PVC) tubing (1.4 cm inner diameter). The PVC tubing served as a mounting sleeve for the solenoid, which was held in place by a set screw tapped through the wall of the tubing. Loosening the hand-tightened set screws allowed rapid

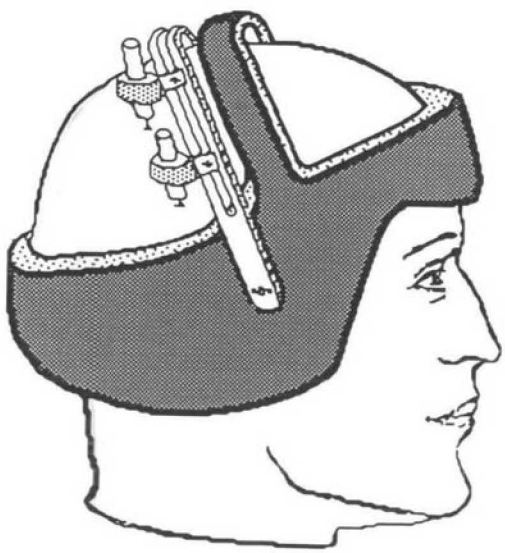


Figure 1. *Vibrotactile helmet showing solenoid mounting.*

Downloaded from hfs.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

inward/outward adjustment of the solenoids for proper alignment with the scalp.

Test Procedure

The subject was instructed to respond to the tactile pulses representing a stimulus as quickly as possible by pressing a labeled key on the computer. The correct key corresponded spatially with the solenoid location that had been stimulated. Both response location and response time were recorded in a data file. Subjects also provided a verbal confidence rating concerning identification of the stimulus location using the following three-point scale: 1—unsure of the location, 2—moderately sure of the location, and 3—very sure of the location.

STUDY 1: STIMULUS DETECTION

Study 1 was designed to ensure that the force intensity of tactile stimulation used in Studies 2 and 3 would be sufficiently above the detection threshold. Although the scalp as a site of tactile stimulation has not been extensively studied, Weber (1834/1978) reported that the forehead was the most sensitive area of the scalp and the crown was the least sensitive area. More recently, Weinstein (1968) and Shimizu and Wake (1982) provided data supporting the sensitivity of the forehead region. However, they did not comprehensively explore other scalp regions. Therefore, Study 1 served as a preliminary investigation of the tactile sensitivity of a limited number of sites on the scalp.

Method

Subjects. The subjects in this experiment were four volunteers (two women and two men). Their average age was 31.7 years (SD = 9.22 years), and all reported no major visual or auditory dysfunction. No subject reported the use of any medication that would alter perceptual ability (i.e., medication incorporating central nervous system [CNS] stimulants or depressants).

Equipment. It was discovered during preliminary testing that the activation of any solenoid caused slight whole-helmet vibrations that were detectable by the subject. In addition, von Békésy (1959) noted that when he used an air puff to help the subject locate a sound source, the subject often could not distinguish between the tactile sensation and the sound. To assess these sources of possible confounding, the helmet was instrumented with four active solenoids and one "dummy" solenoid. The dummy solenoid had the same operating characteristics (e.g., noise spectrum) as the other solenoids. However, the tip was shortened so that it did not make contact with the subject's scalp. By comparing performance on the dummy trials with performance with the active solenoids, a better assessment of detection and localization performance could be made.

The aluminum mounting band was positioned just posterior to the center helmet support area, placing the solenoid array over the parietal meridian of the head (see Figure 1). The dummy solenoid was mounted in the middle of the solenoid array immediately over the midline of the head at approximately the P_z location of the standard 10-20 electroencephalographic (EEG) electrode placement system (Jasper, 1958). Two active solenoids spaced 4.5 cm apart were mounted on each side of the dummy solenoid. Thus two of the active solenoids were located just anterior to locations P_3 and P_4 and the other two were located about midway in the T_3–T_5 and T_4–T_6 regions (of the 10–20 EEG electrode placement system), respectively.

Weber (1834/1978) demonstrated that the two-point limen along the line of points extending from the vertex of the skull toward the sides of the head was larger than that along the line of points equidistant from the vertex. Weber also stated that the crown is

Downloaded from hfs.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

among the least sensitive regions of the scalp. Therefore, a demonstration of accurate detection and localization in the crown area and along a line to the sides of the head could probably be generalized to other scalp locations. This mounting plane was selected for use throughout the current studies because it also allowed cross-study comparisons and provided the broadest amount of scalp surface area (especially needed in Study 2), the greatest ease of solenoid adjustment, and the least amount of obtrusiveness to the subject.

Procedure. Following a briefing on the general nature of the experiment, each subject was instructed to place the helmet on the head in a comfortable position and to secure it firmly with the chin strap. Once the helmet was in place, the solenoids were adjusted to the subject's scalp level to provide equal subjective intensities of stimulation. A small rubber dampener was used to adjust the dummy solenoid so that it provided whole-helmet vibration and acoustic noise approximately equal to that of the other active solenoids without contacting the scalp.

Once the helmet was fully adjusted and calibrated, the subject responded to approximately 40 practice stimuli. The response keyboard was placed immediately in front of the subject. Five number keys were labeled as response buttons corresponding to the five solenoid positions. Response location, response time, and confidence rating were recorded for each stimulus as described in the test procedure section under General Methods. The subject provided a verbal confidence rating following each stimulus trial. Negative values were used to denote trials in which the subject thought a dummy stimulus had been presented. These practice trials served to familiarize the subject with the task and the solenoid locations. Each subject then responded to 240 stimuli at either the low-intensity (15 V) or the high-intensity (20 V) level. Stimuli were administered such that the order of presentation of the solenoid locations was block randomized. Each block included the four possible active solenoid locations plus two dummy trials. After subjects completed the sequence, and following a short rest break while response data were stored, a second series of 240 stimuli was presented at the remaining intensity level. The order of intensity levels was random for each subject. Total test time per subject was approximately 2 h.

Results and Discussion

In addition to response time and confidence rating, two accuracy measures were used. Absolute accuracy represents the percentage of correct responses to either an active solenoid or to the dummy solenoid. Relative accuracy represents how close the subject came to designating the correct site—that is, a response was scored as correct if it was no more than one site away from the actual location. The importance of this measure increases as the number of sites increases in Studies 2 and 3.

Means and standard deviations for the dependant variables are presented in Table 1. The summary statistics are shown for each site (collapsed across subjects and intensity levels) and for each intensity level and collapsed across all four active solenoid sites and both intensity levels. The means and standard deviations for each active solenoid site are based on 320 stimuli (40 stimuli × 2 intensity levels × 4 subjects), and the overall summary (without the dummy solenoid) is based on a total of 1280 stimuli (320 stimuli × 4 sites).

Two-factor repeated measures analyses of variance testing the main effects and interaction of solenoid site and intensity level revealed no statistically significant differences between the low-intensity (15 V, 110 g) and high-intensity (20 V, 150 g) levels for any of the dependent measures. It is important to

Downloaded from hfs.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

TABLE 1

Means (Standard Deviations) of Response Measures for Detection Study

Measure	Site					Intensity		Four Active[a]
	D	1	2	3	4	Low	High	
Absolute accuracy (%)	100	90	95	96	83	91	91	91
Relative accuracy (%)	100	100	100	99	100	100	100	100
Response time (ms)	899	857	1028	965	886	989	881	935
	(293)	(528)	(1563)	(738)	(479)	(1254)	(423)	(936)
Confidence rating	−3.0	2.9	2.9	2.9	2.9	2.9	2.9	2.9
	(0.0)	(0.4)	(0.4)	(0.3)	(0.5)	(0.4)	(0.4)	(0.4)

Note. D = dummy solenoid.
[a] Summary statistics for Sites 1–4, excluding dummy solenoid.

point out that for each of the 1920 stimuli (1280 active, 640 dummy trials), subjects were able to discriminate between an active solenoid stimulus and the dummy solenoid. Thus the research question of Study 1—whether the intensity levels used in these studies were sufficient for detection of a tactile stimulus—was easily answered. At both the low- and high-intensity settings, all subjects were able to achieve 100% perfect detection, though the specific location was not always identified correctly (absolute accuracy).

In general, subjects were faster when making correct location identifications as compared with overall identifications (922 ms vs. 1061 ms, respectively). They were also more confident in making correct responses (2.9) than in making incorrect responses (2.2). In other words, subjects were aware of their probable errors and incorporated this uncertainty in their ratings. The reasonably high correlation of absolute accuracy with confidence rating (r = 0.50) confirmed this relationship.

One further observation is that it took very little practice for subjects to learn the task. However, it was important for subjects to recalibrate themselves prior to the start of each run in order to reinforce the locations of the various solenoid sites. Subjects tended to become more accurate with practice from Session 1 (88%) to Session 2 (94%). A repeated measures analysis of variance (ANOVA) testing the main effects and interaction of solenoid site and session revealed a small but statistically significant increase in the confidence ratings in Session 2, $F(1,3) = 10.08$, $p = 0.05$. Subjects were better able to detect their errors in Session 2, as reflected by greater differences between their confidence ratings for correct and incorrect responses. In Session 1, the mean rating was 2.9 for correct responses and 2.3 for incorrect responses. However, in Session 2, the mean rating was 3.0 for correct responses and 2.1 for incorrect responses.

The results of Study 1 confirmed that the coronal parietal meridian of the scalp was sensitive to tactile stimuli in the intensity range used in this series of studies. Absolute accuracy in localizing the stimulus was uniformly high, and relative accuracy was essentially perfect. The lack of any significant differences in response time or accuracy among stimulus locations suggests that no differential sensitivity across this scalp region existed at stimulus intensity levels well above the detection threshold and for the interstimulus distances used in this study (however, see Weber, 1834/1978).

STUDY 2: STIMULUS LOCALIZATION

Study 2 was designed to investigate localization ability in determining the number of

Downloaded from hfs.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

sites that could be reliably identified in the absence of any additional task loading. From Study 1, it was obvious that four stimuli spaced at 4.5-cm intervals spanning the coronal plane could be easily detected and localized when there was no additional task loading. By increasing the number of stimulus sites, more information on the tactile localization ability of this region of the scalp could be obtained.

Following Study 1, preliminary testing was conducted on laboratory personnel to determine the most reasonable range and configuration of solenoids for use in Study 2. During this preliminary testing, it was determined that localization was difficult for 10 stimuli along a coronal plane and extremely difficult for 12 positions. It was also determined that incrementing the number of stimuli from an even number to an odd number (in which case the additional site is placed at the centerline of the head) did not provide as significant an increase in localization difficulty as incrementing from an odd number to an even number. The midsagittal site was relatively easy to localize with respect to the other stimulus locations. Thus overall performance was roughly the same as for the previous even-numbered configuration. For this reason, only even-numbered stimulus configurations were tested in this study.

Method

Subjects. Six volunteers served as subjects in this experiment (three men and three women). Their average age was 23.50 years (SD = 1.29 years). All subjects had normal or corrected vision and reported no visual or auditory dysfunctions. No subjects reported the use of any CNS-active medication that might have influenced their performance.

Equipment. The equipment used in Study 2 was the same as that in Study 1 with the exception that no dummy solenoid was used. The placement of the aluminum anchoring band for the solenoids was the same as that in Study 1 (i.e., across the parietal region). Before any testing occurred, specific endpoints were established on the aluminum band. These endpoints were 17.5 cm along the aluminum band on either side of the midline and encompassed a total of approximately 180 deg on the stimulus plane. By fixing the endpoint spacing on the aluminum band, we could fix the sites at the same angular separation for each subject. Arrays of 6, 8, 10, and 12 solenoids were tested using interstimulus distances of approximately 3.9, 2.5, 2.1, and 1.8 cm, respectively.

Procedure. Data were handled in much the same manner as in Study 1. Subjects were seated in the lab, given a brief introduction to the nature of the investigation, and instructed regarding the helmet adjustment and response procedures. The helmet was then secured to the subject, adjusted, and calibrated. The different solenoid arrays were presented to each subject in random order. Subjects responded in essentially the same manner as in Study 1, with the only major difference being the use of a larger or smaller number of response keys depending on the number of solenoids in the array. Confidence ratings were again obtained.

For each solenoid array, 20 stimuli were presented at each solenoid location (i.e., for the 6-solenoid condition, 120 total stimuli were presented per subject). Following each condition, solenoids were added or removed from the helmet, readjusted, and calibrated for the next test series. All testing was conducted at an intermediate intensity level of 17.5 V (approximately 130 g force), which was based on the results of Study 1. Total test session time per subject was approximately 3 h.

Results and Discussion

Large portions of the data from two subjects were lost because of computer hardware

Downloaded from hfs.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

malfunctions; thus the summaries and analyses were based on the four remaining subjects. Means and standard deviations for the dependent variables based on 80 stimuli per solenoid site are summarized in Table 2. Figures 2 and 3 present the accuracy and response time means, respectively, for Studies 1 and 2. In the table and figures, response times and confidence ratings for correct trials are for only those stimuli for which the subject identified the exact site of the stimulus. Overall summary statistics include all responses, whether accurate or inaccurate.

The data were analyzed using repeated measures ANOVAs with number of sites (6, 8, 10, 12) as the primary within-subjects factor. Performance differences among the various site configurations were highly significant for all dependent measures. With respect to absolute accuracy, $F(3,9) = 18.93$, $p = 0.0003$, localizing 6 sites resulted in significantly higher absolute accuracy (93%) than that obtained for 10 sites (60%) or 12 sites (47%). The accuracy for 8 sites (76%) was also significantly higher than that for 12 sites. As expected, the relative accuracy followed a similar trend (from 100% for 6 sites to 87% for 12 sites) but with lower statistical significance, $F(3,9) = 3.59$, $p = 0.0595$.

Response times differed significantly, $F(3,9) = 10.71$, $p = 0.0025$, with the time for 6 sites (1028 ms) being substantially shorter than that for 8 sites (1520 ms), 10 sites (1564 ms), or 12 sites (1838 ms). Although there was little difference in response time between 8 sites and 10 sites, there was a substantial decline in absolute accuracy of 16%. The changes in accuracy and response time are most likely attributable to the combined impact of (1) the reduced interstimulus distance accompanying an increase in the number of sites (i.e., a decrease in signal discriminability) and (2) the increase in transmitted information requirements associated with a greater number of stimulus-response alternatives (i.e., Hick's law). The relative contribution of these two factors could not be determined with the current experimental design.

Confidence ratings decreased from 2.9 for 6 sites to 1.6 for 12 sites, $F(3,9) = 12.09$, $p = 0.0017$. The largest and only statistically significant difference was between 10 sites (2.4) and 12 sites (1.6). We concluded, based on subject performance and subject comments, that localizing 12 sites along a single plane is a difficult task. This is true even when one looks at the relative accuracy score, though at 12 sites the relative accuracy was still a respectable 87%. However, subjects were able to detect that a stimulus occurred in all cases and were able to reliably indicate at least the approximate location of the stimulus.

TABLE 2

Means (Standard Deviations) of Response Measures for Localization Study

	All Trials				*Correct Trials*	
Number of Sites	*Absolute Accuracy*	*Relative Accuracy*	*Response Time (ms)*	*Confidence Rating*	*Response Time (ms)*	*Confidence Rating*
6	93%	100%	1028 (457)	2.9 (0.2)	1007 (448)	3.0 (0.2)
8	76%	98%	1520 (847)	2.7 (0.5)	1471 (809)	2.8 (0.5)
10	60%	92%	1564 (805)	2.4 (0.6)	1538 (676)	2.5 (0.6)
12	47%	87%	1838 (651)	1.6 (0.6)	1840 (630)	1.9 (0.6)

Downloaded from hfs.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

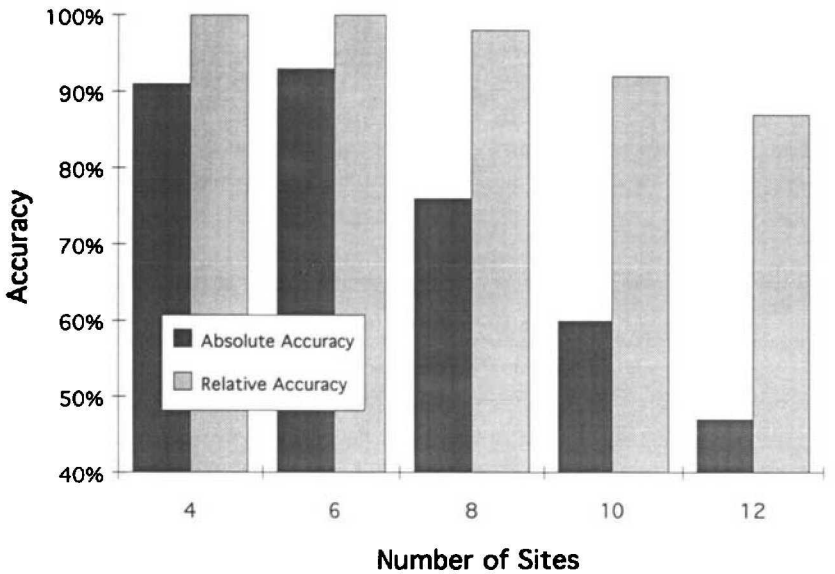


Figure 2. *Absolute accuracy and relative accuracy vs. number of sites. Four-site data are from Study 1; other data from Study 2.*

As in Study 1, subjects were slightly faster and more confident when making correct identifications. Once again, this was also evident from the correlations of dependent measures. Based on the raw data, the correlation between absolute accuracy and confidence rating was $r = 0.35$ across all site conditions. Likewise, the correlation for response

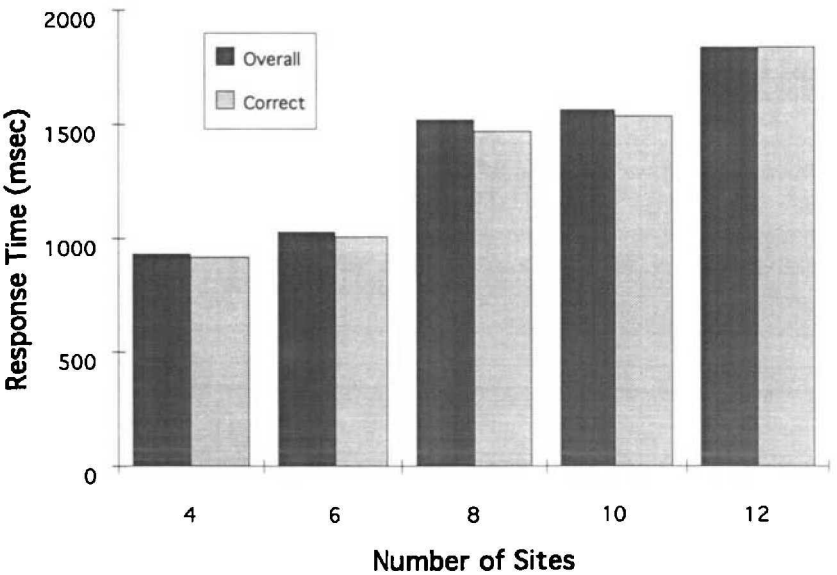


Figure 3. *Response time for all responses and correct responses vs. number of sites. Four-site data are from Study 1; other data from Study 2.*

Downloaded from hfs.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

time and confidence rating was $r = -0.37$. There appeared to be little relationship between accuracy and response time except in the 6-site condition, for which higher accuracy was accompanied by shorter response times ($r = -0.18$).

Analyses of individual site differences for 6, 8, 10, and 12 sites were conducted. For 6 sites, there were significant differences in the response times among sites, $F(5,15) = 4.81, p = 0.008$. Responses for the extreme sites (1 and 6) were faster. This pattern was also evident for 8 sites, though it was not statistically significant. The pattern did not occur for 10 and 12 sites. With the exception of occasional differences in ratings for specific sites, no other significant site differences were observed for the other dependent measures.

To examine performance changes with training during the course of a session, means were computed for the first half and the second half of the trials for each site configuration. For all site configurations, there was a consistent 5% to 8% improvement in absolute accuracy and a 50- to 225-ms improvement in response time.

The data from Study 2 are consistent with those of Weber (1834/1978), who reported that the minimum perceptual tactile interstimulus distance for the top of the head was 3.4 cm. Absolute accuracy for the 6-site configuration involving a 3.9-cm interstimulus distance was quite good (93%). As the interstimulus distance decreased with increasing set size, absolute accuracy declined in near-linear fashion.

In an operational environment, the relative accuracy measure used in this series of studies becomes more useful as a measure of general functional effectiveness when examining an increasing number of stimulus sites. Operational environments would require a fairly high density of tactile stimulators with small interstimulus distances to provide adequate coverage across the scalp surface. However, because they would not require precise localization to provide a functional tactile information system, the relative accuracy measure portrays a reasonably accurate view of the functional performance of the proposed tactile system. In this regard, relative accuracy was reasonably good even at the densest stimulus level (87%), for which absolute accuracy was only 47%. Response time did not appear to degrade in linear fashion, but it clearly was optimal for the 6-site configuration, for which performance was comparable to that for the 4-site configuration from Study 1, in which interstimulus distance was maximized.

STUDY 3: DEMANDING WORKLOAD

Maintaining accurate situation awareness is presumably more difficult in environments that require the management of increasing amounts of complex information. Although it seemed clear from Studies 1 and 2 that subjects could detect and localize tactile stimuli under nondistracting circumstances, it was unclear whether a demanding workload would compromise this ability. Also, if there was a compromising influence, it was unclear whether the type of workload involvement might mediate this effect. Therefore Study 3 was an investigation of the ability to detect and localize tactile stimuli during two types of concurrent task loading.

One task loading condition was designed to explore the influence of standard human performance loading tasks used in typical secondary-task techniques. However, the testing paradigm was not strictly a secondary-task application. Two features were unique. First, in an attempt to increase the ecological validity of the loading tasks, the two selected tasks (memory search and unstable tracking) demanded resources commonly required in the cockpit environment. Thus the subjects performed a dual secondary task. The second variation was in the testing approach. Unlike

Downloaded from hfs.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

standard secondary-task approaches, in the present study the subjects were presented with a balanced secondary-task approach (Gopher and Donchin, 1986) in which they had to balance performance on the primary task (tactile response) with performance on the dual secondary task. The second task loading condition provided even greater ecological validity by presenting a low-fidelity combat flight simulation. This task increased not only cognitive workload but also the general motivational and emotional engagement of the subjects. Differences in the nature of the selected loading tasks provided the opportunity to assess the performance results within the framework of the two attention models discussed in the introduction.

Method

Subjects. Eight subjects (four men and four women) participated in this study. All had normal or corrected vision, and no subject reported taking any medication that would affect visual or auditory acuity. The average age of the subjects was 27.21 years (SD = 7.36 years). Four of the subjects had participated in Study 2.

Equipment and tasks. The results of Study 2 demonstrated that the 8-site configuration provided a reasonable trade-off among speed, accuracy, and confidence ratings. This array represented a challenging task but one in which accuracy was still acceptable (76% absolute, 98% relative). Arrays larger than 8 solenoids evidenced declines in both absolute accuracy and relative accuracy. In addition, the 8-site array largely avoided site-specific response time and accuracy differences. Therefore, the helmet display was instrumented for Study 3 in the same manner as for Study 2 with 8 solenoids mounted on the aluminum band.

In addition to the tactile stimulation task, two task loading conditions were designed incorporating (1) the standard memory search/unstable tracking dual task from the U.S. Air Force Criterion Task Set (CTS; Shingledecker, 1984) and (2) a low-fidelity simulation of a flight mission using the gunnery practice scenario from the commercial World War II air combat flight simulator game Battlehawks 1942 (copyright Lucasfilm Ltd., 1988).

The CTS is a battery of nine basic human performance tasks designed to place selective demands on the elementary mental resources and information-processing functions of the human operator. The CTS model is based on Wickens' (1981) multiple-resources theory and Sternberg's (1969) processing stage theory of human performance. The dual-task version of memory search/unstable tracking was selected to provide additional task loading that relates to the cognitive processing and tracking tasks encountered in flying. The memory search task places demands on a subject's short-term memory retrieval function. Unstable tracking is primarily a psychomotor-output task that requires the subject to keep a vacillating screen cursor between right and left limit lines. It is similar to the critically unstable tracking task developed by Jex, McDonnel, and Phatak (1966). In Study 3, the CTS dual-task version paired the most difficult level of the memory search task (memory set size = 6) with the most difficult level of unstable tracking (λ = 3). The CTS tasks were presented on a Commodore 64 microcomputer and required a custom response pad and rotary controller. The memory search stimuli were presented directly above the unstable tracking display on the video screen, requiring very little gaze aversion to perform both tasks simultaneously.

Battlehawks 1942 was presented on a Zenith 248 microcomputer with an accompanying joystick controller (Flight Stick, CH Products). In Battlehawks 1942, the subject's task was to locate, pursue, and destroy eight enemy aircraft. The subject monitored standard

Downloaded from hfs.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

aircraft instruments to avoid stalls, maintain altitude, and so on, while attending visually to the enemy airspace. The subject sought out the enemy aircraft and attempted to shoot down each one during the test session.

Procedure. Subjects performed at least five standard 3-min CTS trials on each of the memory search and unstable tracking tasks, and then they performed at least five dual-task practice trials. Previous work by Schlegel and Gilliland (1990) suggested that five trials are sufficient to produce reasonably stable and reliable performance on these tasks. The subject responded to this dual-task version by controlling the rotary knob with the preferred hand for the unstable tracking task and pressing the yes/no memory search buttons with the nonpreferred hand. These practice trials were performed at least 24 h prior to the test trials.

On the day of testing, each subject performed two 3-min baseline trials on the memory search and unstable tracking tasks separately, followed by two baseline trials on the dual task. After performing the CTS baseline trials and receiving an orientation to the experimental procedures, the subject secured the helmet, which was then adjusted and calibrated.

Subjects performed three 3-min trials (i.e., total of 9 min) of the CTS dual-task while simultaneously responding to the tactile helmet stimuli in the same manner as in Studies 1 and 2. This multitude of tasks presented the subjects with a scheduling dilemma when the presentation of task stimuli overlapped and required two and sometimes three simultaneous responses. Subjects were forced to schedule responses, typically retaining continuous control of the unstable tracking task and trading off responses on the memory search keypad and the tactile task response keys.

For the Battlehawks 1942 combat mission scenario, subjects were instructed to use the joystick controller with their preferred hands while responding to the tactile stimuli with their nonpreferred hands. In addition to the color video display of the combat simulation and the accompanying auditory stimuli generated by the game (e.g., engine noise), subjects also listened to sound tracks of air-to-air combat cockpit activity (overdubbed twice to reduce clarity) presented at 80 dB (A) SPL. This created an environment that demanded intense concentration and psychomotor activity compounded by intrusive, emotion-laden background noise.

Results and Discussion: Helmet-Based Performance

Means and standard deviations for the dependent variables based on 40 stimuli per subject for each workload condition are presented in Table 3. Figures 4 and 5 present the accuracy and response time means, respectively, for the three task loading conditions. To provide a valid comparison, data from only the four subjects common to both Studies 2 and 3 are presented in the figures. As seen in Table 3, data from the other four subjects were similar, though these subjects were not as proficient given that they had less helmet experience.

The effect of the three loading conditions (helmet only, CTS dual task, Battlehawks 1942) on the tactile performance measures was evaluated using a separate repeated measures ANOVA for each dependent variable. Performance differences among the three task loading conditions were significant for all dependent measures. With respect to absolute accuracy, $F(2,10) = 10.75$, $p = 0.0032$, there was a significant decline in localization accuracy from 76% for the helmet task alone to approximately 60% when combined with the Battlehawks 1942 loading task. The accuracy was roughly equivalent to the 10-site condition with no additional task loading. Helmet accuracy with the CTS loading task was

Downloaded from hfs.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

TABLE 3

Means (Standard Deviations) of Helmet Response Measures for Multiple Task Study

	All Trials				Correct Trials	
	Absolute Accuracy	Relative Accuracy	Response Time (ms)	Confidence Rating	Response Time (ms)	Confidence Rating
Helmet only[a]	76%	98%	1520 (847)	2.7 (0.5)	1471 (809)	2.8 (0.5)
CTS (n = 4)	77%	99%	2452 (1031)	2.4 (0.6)	2389 (1101)	2.5 (0.6)
BH (n = 4)	61%	92%	2587 (1581)	1.9 (0.6)	2467 (1529)	2.1 (0.7)
CTS (n = 8)	76%	99%	2506 (1085)	2.3 (0.7)	2462 (1138)	2.4 (0.7)
BH (n = 8)	58%	93%	2944 (1895)	2.1 (0.7)	2872 (1969)	2.2 (0.7)

Note. CTS = Criterion Task Set; BH = Battlehawks 1942.
[a] Data from Study 2.

almost the same as that with the helmet task alone. A similar result existed for the relative accuracy score, $F(2,10) = 6.14$, $p = 0.0182$. Response time means for both task loading conditions were approximately one full second longer than for the helmet-only condition, $F(2,10) = 5.04$, $p = 0.0307$. There was no difference in response time between the two task loading conditions. Differences in confidence ratings were significant, $F(2,10) = 7.30$, $p = 0.0111$, with an overall lower confidence for the task loading conditions. This difference was particularly evident for the Battlehawks 1942 task. For all three

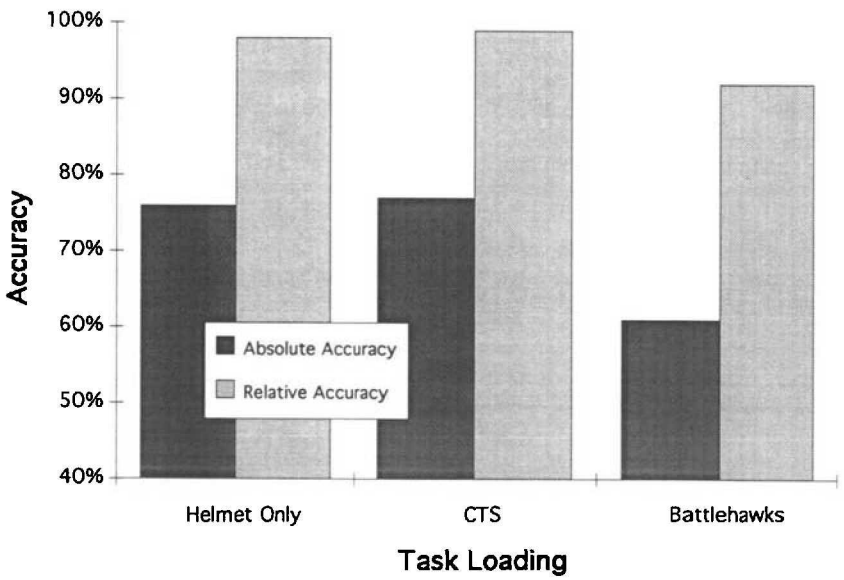


Figure 4. *Absolute accuracy and relative accuracy vs. task loading. Helmet condition represents eight-site data from Study 2.*

Downloaded from hfs.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

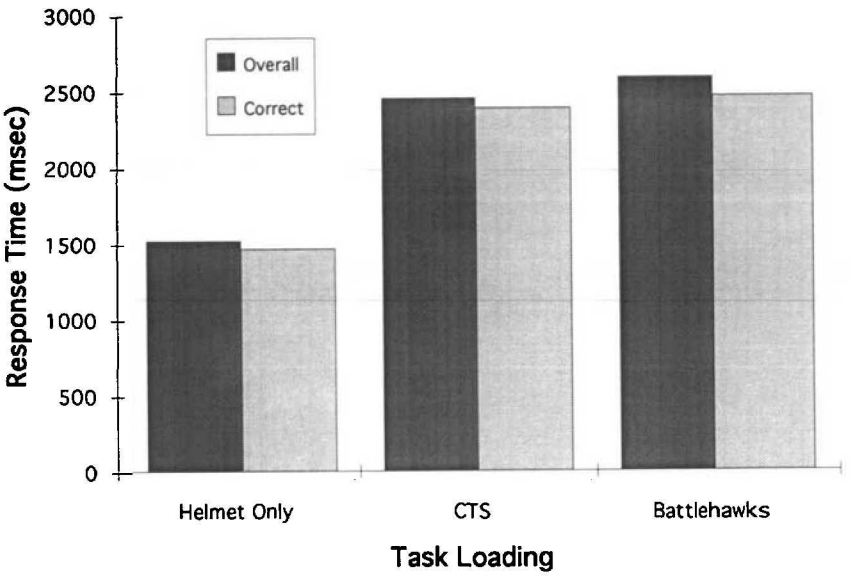


Figure 5. *Response time for all responses and correct responses vs. task loading. Helmet condition represents eight-site data from Study 2.*

conditions, subjects were faster and more confident when making correct localizations. Analyses of individual site differences for the various task loading conditions revealed no clear pattern of performance. Although performance for the extreme sites (1 and 8) tended to be better than for other sites, the differences were not statistically significant.

These results reflect a fundamental difference in the two tasks used for additional loading. The CTS dual task did not appear to interfere with the detection and localization ability of the subject. However, the need to divert visual attention and provide a manual response (competing with the manual responses for the CTS) caused difficulty in scheduling the response activity and required additional hand movement, which was included in the response time measure. This resulted in significantly longer response times but no reduction in accuracy, a remarkable result considering the difficulty of time-sharing all three tasks (memory search, unstable tracking, and helmet monitoring). Viewed within the context of the multiple resource model of attention (Wickens, 1980), these results appear to support the notion of relative independence of the central processing resources allocated to these various tasks. The major source of interference appeared to be in scheduling and performing the motor output responses, an activity that added to the overall multitask demand but did not appear to influence accuracy. A verbal (or other nonmanual) response to the tactile stimulus would more than likely reduce this response time difference.

In contrast to the CTS condition, the Battlehawks 1942 task represented a highly engaging activity in which subjects tended to detect that a tactile stimulus had occurred but apparently assigned a lower priority to its occurrence compared with the battle scenario. The subjects appeared to be less attentive to the stimulus location and delayed their responses to the tactile stimuli because they were motivationally drawn to the demands of the battle scenario. This resulted in

Downloaded from hfs.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

lower accuracy and longer response times. These results were surprising in that the Battlehawks 1942 condition, like the CTS condition, introduced combined memory and tracking demands. Unlike the CTS condition, it did not require the sharing of manual response resources between the two tasks (i.e., separate hands were used for each task), though the need to shift visual attention to the tactile response keys remained. This suggests that the critical factor that may have been introduced by the Battlehawks 1942 condition was the energetic concepts associated with the heightened battle scenario motivational state. For this reason, the cognitive-energetical stage model (Gopher and Sanders, 1984; Sanders, 1983) may play an increasingly important role in understanding complex multitask performance.

Results and Discussion: CTS Performance

To assess the possible intrusion of the tactile helmet task on operational performance, baseline single-task and dual-task CTS performance was compared with CTS performance with the tactile helmet stimuli. Means and standard deviations for the CTS performance task measures across all eight subjects are presented in Table 4. The single-task and dual-task summaries include two 3-min baseline trials per subject. The helmet condition summaries are based on three 3-min CTS trials for each subject. All four dependent measures demonstrated a similar pattern. CTS performance in conjunction with the tactile stimulation task was significantly worse than baseline single-task or dual-task performance. In addition, there were no statistically significant differences between single-task and dual-task performance, though dual-task performance was slightly poorer for all measures.

Compared with baseline performance without helmet stimuli, memory search responses were approximately 50% slower under the helmet condition, $F(2,14) = 51.02$, $p = 0.0001$, and the percentage of correct responses dropped drastically, from 93% in the baseline condition to 78% in the helmet condition, $F(2,14) = 16.23$, $p = 0.0002$. Thus subjects took longer to respond because of the manual response scheduling problem described previously, and their responses were less accurate. Part of the accuracy decline was attributable to the subjects' failing to respond within the response time limit.

TABLE 4

Means (Standard Deviations) of CTS Response Measures for Multiple Task Study

Response Measure	Baseline		With Helmet
	Single Task	Dual Task	
Memory search response time	714 (69)	843 (219)	1211 (229)
Memory search percentage correct	94 (7)	93 (6)	78 (14)
Unstable tracking RMS error	31 (9)	32 (9)	42 (3)
Unstable tracking edge violations	27 (30)	34 (37)	101 (29)

Again, compared with baseline conditions, unstable tracking performance while responding to helmet stimuli was decidedly worse in terms of the root mean square error score, $F(2,14) = 15.58$, $p = 0.0003$, and number of edge violations or control losses, $F(2,14) = 61.94$, $p = 0.0001$. The performance difference was particularly evident for the control loss score, which went from 34 for the dual-task baseline condition to 101 for the helmet condition. The most likely cause for this performance decrement relates to the visual demands of responding to the tactile task stimuli—that is, the subject needed to visually direct the nonpreferred hand from the CTS memory search response buttons to the tactile task response keyboard, make the response, and return the hand to the memory search keypad. Although much of this control

Downloaded from hfs.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

could be done with peripheral vision, this additional task detracted from the subject's visual attention to the unstable tracking cursor. During the tactile response, control losses would typically accumulate until control of the cursor was regained.

The CTS performance differences must also be considered in light of the additional information presented to the subject by the helmet. An alternative display system (visual or auditory) to the tactile system used in this study for presenting the same information would probably affect CTS performance to a greater extent and might be less effective in terms of providing situation awareness information (alerting and spatial information) to the subject.

CONCLUSIONS

The findings from these studies answer many basic questions regarding the use of tactile stimuli to provide situational information in operational environments. One important finding was that detection of a tactile stimulus applied to the scalp was not a difficult task for the intensity of stimuli used in this study. Throughout approximately 5440 stimuli presented during the studies, no subject ever failed to respond to a stimulus. Subjects were typically faster and more confident in their responses when they responded correctly. Incorrect responses were usually slower and less confident, probably reflecting dissonance, indecision, or other processes (e.g., distraction) that may have influenced accuracy.

The number of sites in the solenoid array determined complex patterns of variation in the subject's response characteristics. As expected, smaller arrays (4 and 6 sites) provided the fastest response times, the highest degree of accuracy, and the highest confidence ratings. However, the 8-site array appeared to optimize the trade-off between the higher precision afforded by more sites and overall response effectiveness.

Additional task loading of any type that resulted in competition for manual response resources generally increased response latency to the tactile stimuli. Accuracy was not significantly affected by simultaneous performance of an additional, fairly complex dual task. However, a task similar in its absorbing quality to that of an actual flight mission scenario did significantly affect accuracy. The effect of the tactile task on concurrent human performance tasks was significant: ongoing task performance was degraded in comparison with that under baseline conditions.

These studies have also served to highlight and generate a number of important future research questions. The basic issues of response mode and instruction set should be addressed. For example, can some of the response time decrements observed with additional task loading be reduced by using oral responses to the tactile stimuli? This would reduce some of the manual response competition and the concurrent visual demands. Can varying the instructions given to the subject mediate some of the response competition differences? For example, helmet performance might improve significantly if the subject is instructed to assign greater importance to the tactile stimuli. Performance on ancillary tasks may also improve.

Although accuracy is an important factor, the current studies suggest that absolute accuracy becomes less important than speed. Assuming a reasonable level of relative accuracy, speed becomes the predominant factor if the tactile helmet is to be used effectively in an operational environment (i.e., maximizing the ability to rapidly impart information through the disruptions of ongoing cockpit functions).

It is also important to point out that the helmet's disruptive effects must be compared with those of alternative methods of providing

Downloaded from hfs.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

the same information. A visual or an auditory display performing the same alerting functions might intrude on the concurrent task to the same, or an even greater, degree. The ultimate question is which display mode results in the largest information transfer and the smallest performance decrement.

The meaning of the tactile stimuli to the pilot is another crucial question. The lower accuracy observed in the combat simulation (Battlehawks 1942) condition compared with that in the CTS dual-task condition strongly suggests the need for more research on the effect of the absorbing nature of the concurrent task. In this regard, the basic research mentioned previously could be creatively designed to determine whether response mode or instruction set could effectively minimize adverse effects. Regardless, the tactile helmet display appears to have substantial potential for providing pilots with important information about their environment. However, considerable research is needed to further explore the effectiveness of this new device.

ACKNOWLEDGMENTS

This research was funded by the Boeing Military Airplane Company, Wichita, Kansas, under contract AB0211. The authors thank Mary Hornsby and Chris Goetsch of BMAC for providing their support for the effort. Appreciation is extended to research assistants Randa Shehab, Staci Allen, and Scott Mills for their contributions in data collection, reduction, and analysis and to Marie Gomes for her work in preparing a preliminary review of the literature on tactile and vibrotactile stimulation.

REFERENCES

Bliss, J. C., Katcher, M. H., Rogers, C. H., and Shepard, R. P. (1970). Optical-to-tactile image conversion for the blind. *IEEE Transactions on Man-Machine Systems, MMS-11,* 58–64.

Craig, J. C. (1974). Pictorial and abstract cutaneous displays. In F. A. Geldard (Ed.), *Cutaneous communication systems and devices.* Austin, TX: Psychonomic Society.

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 97–101). Santa Monica, CA: Human Factors and Ergonomics Society.

Fenton, R. E. (1966). An improved man-machine interface for the driver-vehicle system. *IEEE Transactions on Human Factors Engineering, HFE-7,* 150–157.

Geldard, F. A. (1972). *The human senses* (2nd ed.). New York: Wiley.

Geldard, F. A. (1974). *Cutaneous communication systems and devices.* Austin, TX: Psychonomic Society.

Gilson, R. D., and Fenton, R. E. (1974). Kinesthetic-tactual information presentations—Inflight studies. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-4,* 531–535.

Gopher, D., and Donchin, E. (1986). Workload—An examination of the concept. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 2, pp. 41-1–41-49). New York: Wiley.

Gopher, D., and Sanders, A. F. (1984). S-OH-R: OH stages! OH resources! In W. Pritz and A. F. Sanders (Eds.), *Cognition and motor processes* (pp. 231–253). Berlin: Springer-Verlag.

Heller, M. A., and Schiff, W. (1991). *The psychology of touch.* Hillsdale, NJ: Erlbaum.

Jasper, H. H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology, 10,* 371–375.

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

Navon, D., and Gopher, D. (1979). On the economy of the human processing system. *Psychological Review, 86,* 214–255.

Norman, D. A., and Bobrow, D. J. (1975). On data limited and resource limited processes. *Cognitive Psychology, 7,* 44–64.

Sanders, A. F. (1983). Toward a model of stress and human performance. *Acta Psychologica, 53,* 61–97.

Schlegel, R. E., and Gilliland, K. (1990). *Evaluation of the Criterion Task Set* (Tech. Report AAMRL-TR-90-007). Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory.

Sherrick, C. E. (1975). The art of tactile communication. *American Psychologist, 30,* 353–360.

Sherrick, C. E., and Cholewiak, R. W. (1986). Cutaneous sensitivity. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1, pp. 12-1–12-58). New York: Wiley.

Shimizu, Y., and Wake, T. (1982). Tactile sensitivity to two types of stimulation: Continuous and discrete shifting of points stimulus. *Perceptual and Motor Skills, 54,* 1111–1118.

Shingledecker, C. A. (1984). *A task battery for applied human performance assessment research* (Tech. Report AFAMRL-TR-84-071). Wright-Patterson Air Force Base, OH: Air Force Aerospace Medical Research Laboratory.

Sternberg, S. (1969). Scanning mental processes revealed by reaction time experiments. *American Scientist, 57,* 421–457.

von Békésy, G. (1959). Similarities between hearing and skin sensations. *Psychological Review, 66,* 1–22.

Weber, E. H. (1978). *The sense of touch* (H. E. Ross and D. J. Murray, Trans.). London: Academic. (Original work published 1834)

Weinstein, S. (1968). Intensive and extensive aspects of tactile sensitivity as a function of body part, sex, and laterality. In D. R. Kenshalo (Ed.), *The skin senses* (pp. 195–222). Springfield, IL: Charles C. Thomas.

Wickens, C. D. (1980). The structure of attentional

Downloaded from hfs.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

resources. In R. Nickerson and R. Pew (Eds.), *Attention and performance VIII* (pp. 239–257). Hillsdale, NJ: Erlbaum.

Wickens, C. D. (1981). *Processing resources in attention, dual task performance, and workload assessment* (Tech. Report EPL-81-3). Urbana-Champaign, IL: University of Illinois, Engineering Psychology Research Laboratory.

Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd ed.). New York: HarperCollins.

Zlotnik, M. A. (1988). Applying electro-tactile display technology to fighter aircraft—Flying with feeling again. In *Proceedings of the IEEE 1988 National Aerospace and Electronics Conference NAECON 1988* (pp. 191–197). New York: IEEE Aerospace and Electronics Systems Society.

Date received: March 16, 1993
Date accepted: March 18, 1994

Downloaded from hfs.sagepub.com at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016