ELSEVIER

Biological Psychology 42 (1996) 361–377

BIOLOGICAL PSYCHOLOGY

Psychophysiological responses to changes in workload during simulated air traffic control

Jeffrey B. Brookings[a], Glenn F. Wilson[b,*], Carolyne R. Swain[c]

[a]*Wittenberg University, Springfield, OH, USA*
[b]*Armstrong Laboratory/CFHP, 2255 H Street, Wright Patterson Air Force Base, OH 45433-7022, USA*
[c]*Logicon Technical Services, Inc., Dayton, OH, USA*

Abstract

In this investigation, eight Air Force air traffic controllers (ATCs) performed three scenarios on TRACON (Terminal Radar Approach Control), a computer-based air traffic control (ATC) simulation. Two scenarios were used each with three levels of difficulty. One scenario varied traffic volume by manipulating the number of aircraft to be handled and the second scenario varied traffic complexity by manipulating arriving to departing flight ratios, pilot skill and mixture of aircraft types. A third scenario, overload, required subjects to handle a larger number of aircraft in a limited amount of time. The effects of the manipulations on controller workload were assessed using performance, subjective (TLX), and physiological (EEG, eye blink, heart rate, respiration, saccade) measures. Significant main effects of difficulty level were found for TRACON performance, TLX, eye blink, respiration and EEG measures. Only the EEG was associated with main effects for the type of traffic. The results provide support for the differential sensitivity of a variety of workload measures in complex tasks, underscore the importance of traffic complexity in ATC workload, and support the utility of TRACON as a tool for studies of ATC workload.

Keywords: Cognitive workload; Air traffic control; EEG; Heart rate; Eye blinks; Respiration

1. Introduction

It has long been recognized that air traffic control (ATC) is a complex and demanding job (see Noland, 1990, for a discussion of ATC tasks and responsibilities). Over the past decade, amid growing concerns that increased

* Corresponding author.

SSDI 0301-0511(95)05167-Z

traffic in the national airspace system threatens to overwhelm air traffic controllers (ATCs) and compromise the safety of air travel (Danaher, 1985), more attention has been focused on ATC workload. The Federal Aviation Administration (FAA), for example, developed the National Airspace System plan that proposed, among other things, to increase the automation of ATC (Federal Aviation Administration, 1985). It was anticipated that automation would decrease ATC workload, so that projected increases in the demand for air traffic services could be met safely and efficiently.

The notion that decreased workload should result in improved ATC service is intuitively reasonable, but Hopkin (1989) noted that the salutary effects of workload reduction "... seem to be assumed rather than proved" (p. 103). Indeed, studies of workload and ATC operational errors have yielded mixed results. Morrison and Wright (1989) analyzed reports from the Aviation Safety Reporting System of NASA, and found that controller errors (e.g., monitoring failures, improperly executed handoffs, wrong heading or altitude assignments) were associated most frequently with increases in workload factors such as traffic volume and frequency congestion. On the other hand, analysis of FAA data for 1989 performed by Redding (1992) indicated that ATC operational errors occurred more frequently under moderate – rather than high – workload conditions (see also Stager, Hameluck & Jubis, 1989).

Statistical analyses of incident data are useful for characterizing overall ATC system performance and for generating workload evaluation hypotheses. However, the reliance of such analyses on accurate and complete incident reports, coupled with the fact that in retrospective analyses controller workload must be inferred from traffic data, rather than assessed directly, underscores the need for controlled studies of ATC workload.

Much of the research using physiological measures to study ATC has been focused upon issues related to the effects of long-term stress on controller health. While this is quite important, the focus of this paper is on the effects of the momentary workload experienced by controllers as they perform their duties. This approach is more directly related to the concerns stated above about the effects of increased air traffic as well as the effects of automation upon the workload of the individual controller.

Psychophysiological measures have been used in a number of real world environments to monitor operator workload (for a review, see Wilson & Eggemeier, 1991). Heart rate has been reported to vary as a function of the mental load imposed by the operator's task. Wilson (1993) reported that increased heart rate was associated with more difficult aspects of fighter aircraft air-to-ground missions and the heart rate of the pilot increased more than that of the accompanying weapons systems officer except when this officer was in control of the aircraft. Heart rate measures have been used in the process of certification of civilian airline crew (Roscoe, 1987; Speyer, Fort, Fouillot & Blomberg, 1988). The beat-to-beat variability of the heart rhythm has also been used as a measure of mental effort and has been reported to decrease with increases in mental demand in environments ranging from simulated flight

(Itoh, Hayashi, Tsukui & Saito, 1989) to car driving (Egelund, 1982). However, not all reports have found heart rate variability to be sensitive to different levels of mental demand (Wilson, 1992). Respiration measures have also been applied to the work environment and have been reported to change with differing levels of mental effort (see Caldwell et al., 1994; Wientjes, 1992, for reviews). Several investigators have reported decreased rates of eye blinking in situations with high visual demands (Stern & Dunham, 1990). Measures of the electrical activity of the brain are seen by many to be direct ways of determining the cognitive demands placed upon an operator. The brain is responsible for processing information, making decisions and initiating actions on the external environment. Because ATC is a complex task, no doubt involving numerous brain areas, recording from multiple sites over the brain would be appropriate. Because the psychophysiological measures are continuously present and their recording is not intrusive to the operator's job performance it was decided to use them as measures of controller workload as suggested by Kalsbeek (1971).

Many laboratory and simulator studies have used untrained subjects or college students who were trained on part-task ATC simulators. This points out a recurrent problem when investigating controller workload. It is both expensive and time-consuming to select and train subjects in order to establish a test group equivalent to actual controllers. In the present investigation, Air Force controllers were used. They had already been through the required selection and training process. Because these subjects were trained, they already possessed the skills and strategies used by controllers to perform their jobs.

The purpose of the current investigation was to assess workload changes associated with variations in the difficulty of TRACON (Terminal Radar Approach Control), a simulated ATC task. Air Force ATCs were recruited to participate in the study, and multiple measures, which included performance, subjective, and physiological metrics, were collected to provide a comprehensive perspective on controller workload.

2. Method

2.1. Subjects

The subjects were 8 Air Force volunteer ATCs (7 male, 1 female) who were paid for their participation. All subjects were right-handed, ranged from 21 to 29 years of age, and reported ATC experience of 2.5–7.5 years.

2.2. Apparatus

The TRACON scenarios were administered on a Unisys 386/25 MHz processor and ViewSonic 7 color VGA monitor (17 in, 640H × 480V resolution graphics). Subjects entered their commands via keyboard and a two-button mouse. Pilots' voices, which were generated using a SoundBlaster interface card, were presented to subjects over a speaker.

Electrophysiological recording of peripheral psychophysiological measures was accomplished using the Psychophysiological Assessment Test System (Wilson & Oliver, 1991). EOG was monitored for vertical blink and horizontal saccade activity via Ag/AgCl electrodes positioned above and below the left eye and at the outer canthus of both eyes. EOG signals were filtered at 0.1-30 Hz and amplified by 5000 using Grass P511 amplifiers. Electrodes for heart activity were positioned on the sternum and fifth intercostal space on the left side of the body. Signal amplification was 2000 and the data were filtered at 10-100 Hz. The ground electrode was on the right side of the ribcage in the fifth intercostal space. Respiration was monitored using a Respitrace system with elastic transducer bands on the chest and abdomen which were summed for output with respiration amplitude calibrated prior to data collection for each subject. Nineteen channels of brain wave data, using the 10-20 electrode system, were also recorded with an ElectroCap and Biologic Brain Atlas III with a linked mastoid reference. The amplifier gain was 30 000 with bandpass filters at 0.1 and 30 Hz. Electrode impedences were 5K 6 or less.

2.3. Task

The simulated ATC task was TRACON for Windows (Version 1.03, Wesson International). The TRACON display consisted of: (a) a color radarscope depicting Los Angeles International and four surrounding airports; (b) flight strips listing 'active' and 'pending' aircraft; (c) a communications box that provided an ongoing visual record of controller commands and pilot responses; and (d) the controller's score for the current scenario. All information displayed on the radarscope (e.g., airways, VOR radio beacons, sector boundaries) represents accurately the information contained in government airspace charts.

The task of the subjects was to handle a series of aircraft requesting ATC services by issuing commands (e.g., turn, descend) that enabled the safe and efficient execution of their respective flight plans. The flights, representing a variety of aircraft types flying under either instrument or visual flight rules, included arrivals, departures and overflights. Subjects received points for successfully handling aircraft, minus any points deducted for operational errors (separation conflicts, hand-off errors, missed approaches).

In the experimental sessions, each subject completed three scenarios. The rationale for the workload manipulations underlying these scenarios was derived from a review of ATC literature and conversations with civilian and Air Force ATCs. One scenario consisted of 36 aircraft presented over a 45 min period and included 15 min each of low, medium and high workload segments differing in traffic volume (6, 12 and 18 aircraft, respectively). All other factors, such as aircraft type and ratio of arrivals to departures and overflights, were held constant across the segments. The second scenario was also 45 min long and consisted of 36 aircraft divided among three workload segments, but the number of aircraft (12) was held constant across segments. Instead, the segments varied in degree of traffic complexity, operationalized as a function of: (a) the ratio of arriving flights to departures and overflights; (b) the probability that pilots either did not hear or failed to execute properly the commands of the controller; and (c) the heterogeneity of aircraft type. In both scenarios, the workload segment transitions were arranged so that subjects had sufficient time at the conclusion of a segment (approximately 1 min) to complete subjective workload ratings before the next segment began. The third scenario, an overload condition, presented controllers with 15 aircraft in 5 min. The purpose of this scenario was to overwhelm the controllers or, in ATC parlance, make them 'lose the picture' with particular attention to physiological activity accompanying this phenomenon.

2.4. Dependent measures

The principal performance measure was the ratio of TRACON points earned to the total points possible for that segment. The subjective workload measures were the NASA TLX scores (Hart & Staveland, 1988), based on ratings collected at the end of each segment. The TLX contains six subscales that are scored from a low of 0 to a high of 100: mental demand, physical demand, temporal demand, performance, effort and frustration. A composite score is also derived based upon combining the scores from all of the subscales. The peripheral physiological measures, calculated from 5 min mid-point workload periods within each segment, included: (a) heart rate and heart rate variability; (b) eye blink rate; (c) saccade rate and amplitude; and (d) respiration rate and amplitude. Since the workload in each condition takes time to build up in order to keep the scenarios realistic, FFTs were performed on the 2 min of EEG data following the 'peak' workload of each segment. The EEG data were corrected for eye artifacts prior to calculating the FFTs using a modified version of the eye movement correction procedure developed by Gratton, Coles & Donchin (1983) and portions containing other artifacts were rejected. The FFT data were then divided into five bands for further analysis. The bands were delta (1.1–3.9 Hz), theta (4.3–7.8 Hz), alpha (8.3–11.9 Hz), beta 1 (12.3–15.8 Hz), and beta 2 (16.2–24.9 Hz). In order to reduce EEG power variations between subjects, all power spectral band values were expressed as the percentage of the total power between 1.1 and 24.9 Hz. For statistical comparisons, SAS was used to compute repeated measures analyses of variance, with the Geiser-Greenhouse correction, to assess main effects. Pairwise contrasts (t-tests) were used to test differences between individual conditions, and those significant at $p < 0.05$ are reported.

2.5. Practice

In an initial briefing, controllers were introduced to the TRACON simulation and were given instructions on how to issue commands using the keyboard and mouse. Then, during off-duty hours, the controllers completed a series of eight practice simulations arranged in order of increasing difficulty. Subjects were required to complete each simulation with no crashes, separation conflicts, or hand-off errors before proceeding to the next one. The controllers kept performance records for each practice trial, and the mean time required to complete the eight simulations was approximately 6 h. When all practice simulations were completed successfully, controllers were scheduled for their experimental session.

2.6. Procedure

The experimental session for each controller took approximately 4 h. This included: (a) the time required to instrument the subjects (i.e., attach electrodes, affix the respiration bands and EEG cap); (b) a 3 min resting period to collect baseline data for the physiological measures; (c) a 5 min warm-up simulation (four aircraft) to familiarize controllers with the laboratory computer; (d) the two long scenarios; (e) the overload scenario; and (f) 15 min breaks between each of the three principal scenarios.

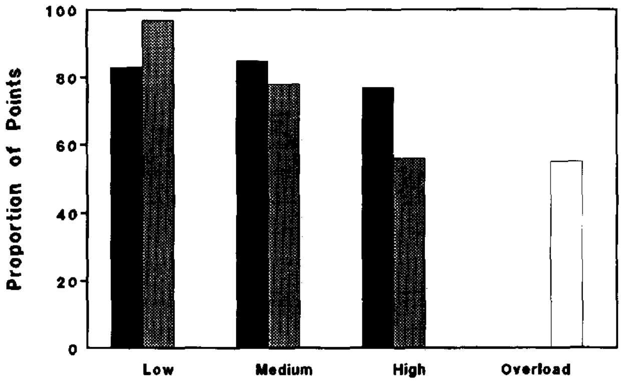
The order in which controllers completed the volume and complexity scenarios was counterbalanced, as was the order of the three workload segments within the two scenarios. The overload scenario was always completed last. Experimental sessions were conducted in a shielded room. The experimenter sat behind and to the right of the subject, and administered the TLX after each workload segment of the long scenarios and at the conclusion of the overload scenario.

3. Results

For the performance and subjective data, repeated measures analyses of variance (ANOVAs) indicated: (a) significant main effects of task difficulty (low, medium, high) on TRACON performance ($F = 4.64$, $p < 0.04$) , all six of the TLX subscales, and the TLX composite score ($F = 25.65$, $p < 0.0001$); (b) no main effects for traffic (volume versus complexity); and (c) no significant difficulty × traffic interactions. TRACON and TLX composite score data are presented in Figs. 1A and 1B, respectively. Performance scores in the high volume and high complexity conditions were significantly lower than in the low complexity condition. Further, the overload condition produced lower performance scores than the low and medium workload conditions of either volume or complexity scenarios. The TLX composite score results mirrored the performance data with the addition of high complexity condition being rated as more difficult than the medium complexity condition and a significant difference between low and medium volume conditions.

Blink rate decreased significantly as the TRACON task became more difficult ($F = 9.37$, $p < 0.01$). However, there was no significant difference between the two traffic manipulations of complexity and volume ($F = 2.75$, $p > 0.05$). As can be seen in Fig. 2A, the blink rates declined with increasing

A.



B.

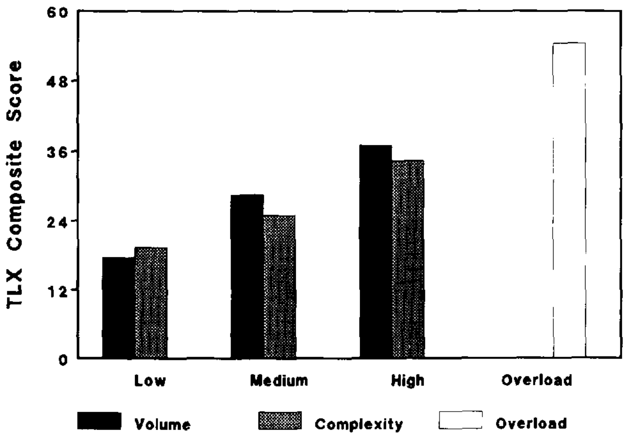
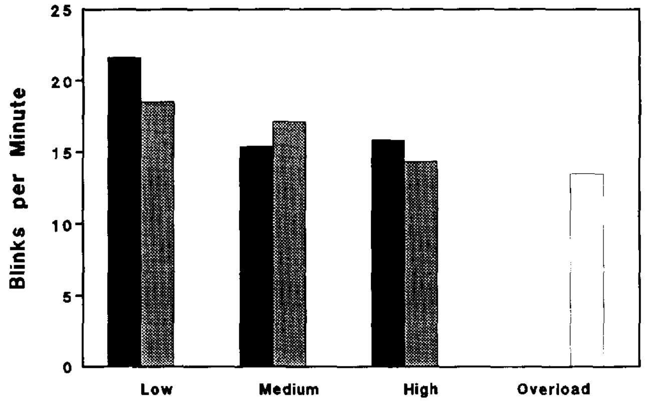


Fig. 1. (A) Performance data expressed as a proportion of total possible TRACON points for each level of task difficulty and traffic type. (B) Composite subjective scores for each level of task difficulty and both traffic types including overload.

TRACON difficulty. Post hoc t-tests ($p < 0.05$) demonstrated that the blink rates in the low level conditions were significantly higher than those observed during the high workload levels or during the overload condition. The medium volume condition blink rates were significantly lower than the low workload level. The medium complexity condition blink rates were significantly lower than the low volume condition and higher than the high complexity condition. The blink rates recorded from the overload condition were significantly lower than both low workload conditions and the medium complexity condition.

A.



B.

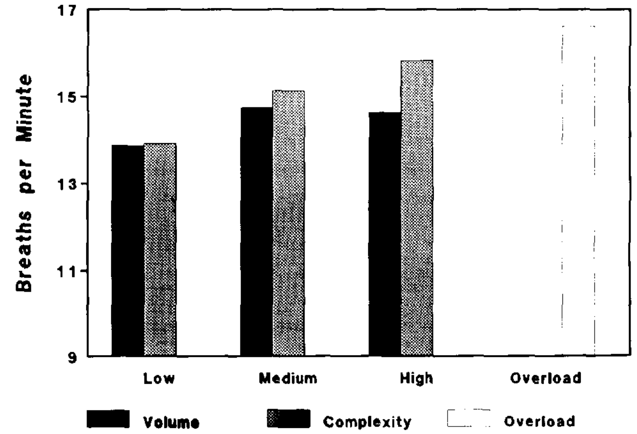


Fig. 2. (A) Mean blinks per min for each TRACON segment. (B) Mean respiration rate for each simulation condition.

Although respiration amplitude was unaffected, respiration rate was higher as the TRACON difficulty increased ($F = 17.88$, $p < 0.0015$) but were not affected by the traffic manipulation ($F = 1.84$, $p < 0.22$). The mean respiration rates are depicted in Fig. 2B. The high complexity condition was associated with significantly higher respiration rates than both low conditions and the medium volume condition. Both medium conditions and the overload condition showed significantly higher rates compared to the low complexity condition while the overload respiration rates were also significantly higher than those of the low volume condition.

Heart rate and saccade measures did not demonstrate significant differences to either difficulty or traffic manipulations. All TRACON conditions were associated with heart rates higher than the resting baseline by 1.77–3.73 beats per min. Heart rate variability in the 0.15 to 0.4 Hz band did approach significance as a result of the difficulty manipulation ($F = 3.54$, $P < 0.066$) but there were no task related differences in the 0.06–0.14 Hz band.

The FFTs for the two 1 min segments following the 'peak' of each condition did not differ statistically and were pooled for the analyses. Fig. 3 summarizes the results of the ANOVAs performed on the EEG data from the nineteen channels for the five bands. The EEG activity in the theta band was sensitive to task difficulty manipulations of the TRACON simulation. The effects of difficulty were evident at central, parietal, one frontal and one temporal site. Percent theta power at all of these sites showed significant increases as task difficulty increased as can be seen in Fig. 4. The frontal and temporal sites were limited to the right hemisphere while the central and parietal were found more on the midline. Pairwise contrasts indicated that percent theta power in the high difficulty conditions was significantly larger than that obtained during low difficulty conditions at F8, C3, Cz, T4, P3, Pz and P4. Medium difficulty theta responses were significantly larger than the low difficulty responses at C3, P3 and Pz. The overload condition was associated with significantly increased theta activity compared to the low volume task condition at electrode sites F3, C3 and Cz. Theta in the overload condition was also significantly larger than in the volume medium condition at T6 and O2. Sites F7, F3, Fz and C3 evidenced significantly greater theta during the overload condition compared to the high volume condition.

Alpha band power changes were driven primarily by the interactions between the difficulty and traffic manipulations at FP1, F7, Fz, F4 and Cz sites with F3 and P4 showing marginally significant interactions ($p < 0.055$ and

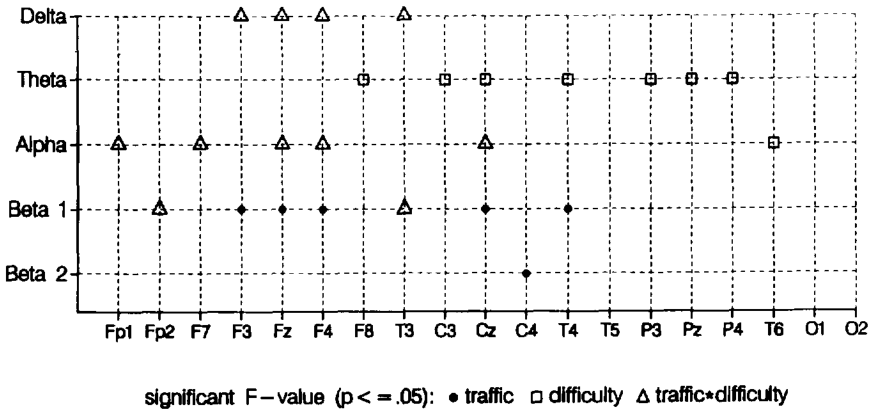


Fig. 3. Significant F values, $p < 0.05$, are indicated by the dot (traffic manipulation), box (difficulty manipulation), or triangle (interaction) for each combination of electrode and EEG band. Intersections without a symbol did not reach significance.

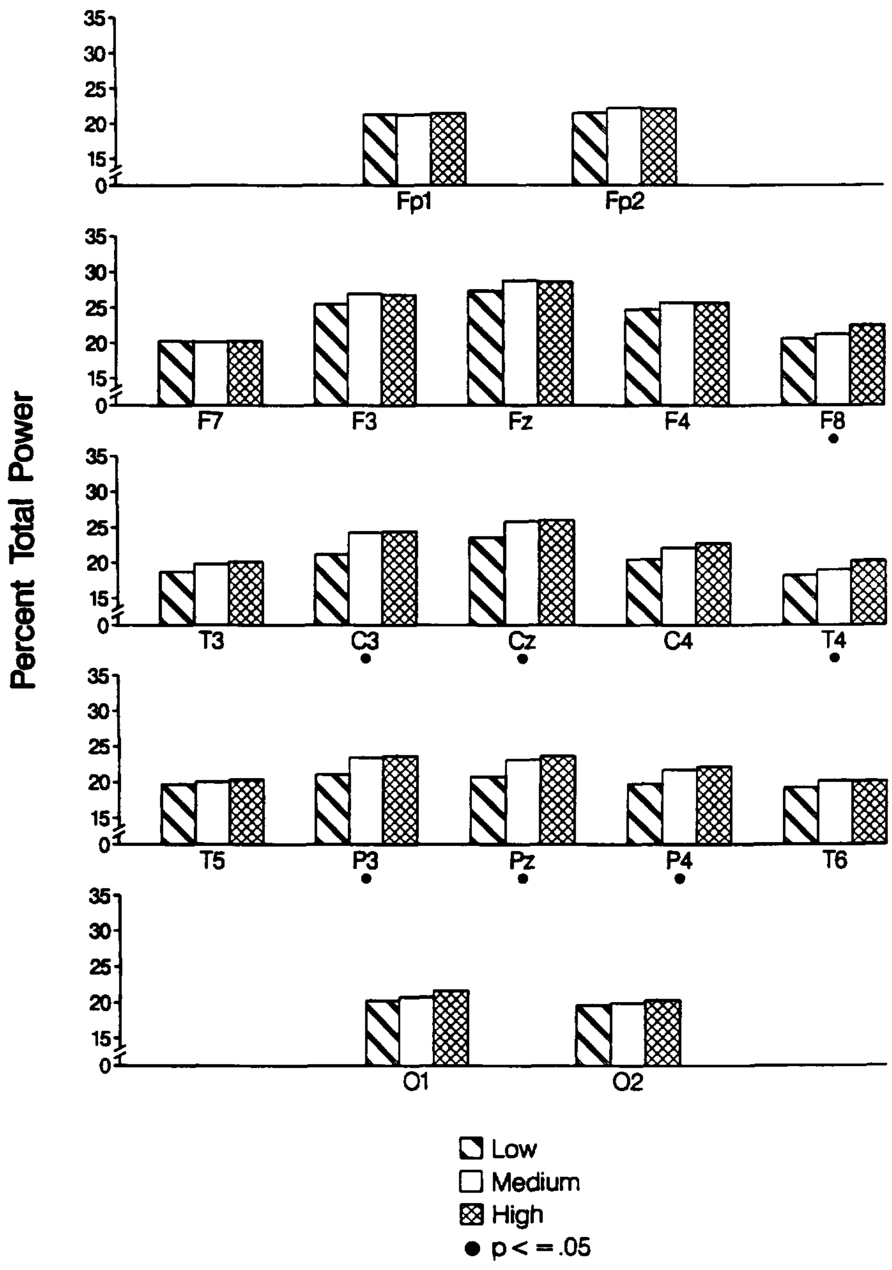


Fig. 4. Percent total power data from the theta band for each electrode at each level of volume and complexity. The filled circles below an electrode label indicate statistical significance at $p < 0.05$ or less.

$p < 0.06$). Percent power was highest for the low complexity conditions, lower for the high complexity conditions and lowest for the medium complexity conditions. Pairwise contrasts revealed significantly higher percent alpha power in the low compared to both the high and medium conditions at sites Fz, F4 and Cz. For the volume manipulations the three workload levels were statistically equivalent. Percent alpha power at T6 also yielded a significant main effect due to task difficulty with greater alpha percent power in the low difficulty conditions compared to both the medium and high difficulty conditions which were not significantly different (see Fig. 3).

The beta 1 band revealed significant differences between the traffic conditions at F3, Fz, F4, Cz, T4 with C3, C4 and P4 marginally significant ($p < 0.055$ to $p < 0.066$). The complexity manipulation was associated with a higher percent of beta 1 power than the volume manipulation. Significant beta 1 interactions were found at sites FP2 and T3. Pairwise contrasts indicated that the overload condition was associated with significantly larger percent beta 1 than with the low volume condition at F7 and T4 or the high volume condition at T6. However, low complexity and medium complexity conditions were associated with greater beta 1 relative power than overload at sites Fz, F3 and Pz, and at sites F7 and T4, respectively. For the beta 2 band there was significantly greater percent power in the complexity traffic than in the volume traffic conditions at site C4 but this band was unaffected by manipulations of workload.

The delta band activity also demonstrated significant differences due to the interaction between traffic and difficulty manipulations at electrode sites F3, Fz, F4 and T3. For the complexity conditions the highest percent delta power was found during the medium difficulty conditions with lowered values associated with the high conditions. The low difficulty conditions evidenced the least delta power with significant differences between medium and low at Fz and F4. For volume, the low difficulty segment was associated with higher power than medium and high difficulty segments which were approximately the same magnitude. Pairwise contrasts indicated low difficulty responses to be significantly enhanced compared to high difficulty responses only at electrode site T3.

4. Discussion

The results of this investigation provide support for the sensitivity of a variety of workload indexes to manipulations in the difficulty of a simulated ATC task using Air Force controllers. Specifically, changes in difficulty produced differences in TRACON performance, TLX ratings, eye blink rate, respiration rate, and EEG frequencies. Changes in the traffic manipulation, complexity or volume, produced changes in the EEG only. These differences were consistent with expectations derived from the workload and ATC literature. While the overload condition did represent the highest workload level based on subjective ratings and performance only one subject reported to have 'lost the picture' as anticipated.

The decrease in eye blink rate with higher workload indicates that part of the increased demands placed upon the controllers involves visual attention. This interpretation is supported by other investigations reporting decreased blink rates when operators were under conditions of increased visual load (Stern & Dunham, 1990; Wilson, 1993; Wilson, Purvis, Skelly, Fullenkamp & Davis, 1987). As the visual demands increased, the controllers blinked less often so as not to miss important events on the radar scope.

It is possible that the increased respiration rate associated with increased task demand reflected the increased metabolic demands resulting from the greater mental activity required for task performance. Carroll, Turner and colleagues (Carroll, Turner & Hellawell, 1986; Carroll, Turner & Rogers, 1987; Turner & Carroll, 1985) have reported increases in respiration rate when subjects performed more difficult laboratory tasks. However, they reported correlated changes in heart rate and respiration rate which were not found in the present study. Backs, Ryan & Wilson (1994), on the other hand, reported that respiration was more sensitive to changes in tracking task difficulty than was heart rate. Wientjes (1993) also found respiration rate, among other respiration variables, to be more sensitive to changes in the difficulty of a memory loading task than was heart rate. Wientjes (1992) provides an alternative hypothesis that states that in tasks such as the present one changes in respiration rate are due to psychological demands and not metabolic demands. The current data would support this interpretation since heart rate did not change during the TRACON manipulations. In more applied environments, respiration rate was reported to be more sensitive to simulated flight workload than heart rate in a study by Opmeer and Krol (1973); and Harding (1987) further reported increased respiration rates during the more demanding portions of aircraft flights.

The lack of significant changes in heart rate and heart rate variability suggests that these measures were not sensitive to the differences in the levels of mental demand for this task. Costa (1993) has also reported that heart rate was not reliably related to ATC workload as measured by the number of aircraft handled during three shifts with varying levels of workload. It appears that in the present study one level of heart rate activity was sufficient to supply the metabolic needs of the brain at all levels of task demand.

Because blinks and respiration can be recorded continuously and are non-intrusive, they could be used as on-line monitors of cognitive demand during normal working conditions without interfering with the primary task of the controller. The absence of main effects for traffic indicates that the volume and complexity manipulations had equivalent effects on these dependent measures. Of course, it is not possible to completely separate traffic volume and traffic complexity. In fact, there remains considerable disagreement about the number and relative importance of the elements that comprise airspace complexity (Mogford, Murphy & Guttman, 1994).

The topographically recorded relative EEG power was sensitive to both difficulty and traffic manipulations of the TRACON simulation as well as to the interaction between these manipulations. The EEG bands were differentially sensitive to the TRACON difficulty manipulations. The theta band showed significant increases in relative power with increased task difficulty while increased relative beta activity was associated with the complexity manipulations when compared to the volume manipulations. Delta and alpha band activity was differentially affected by the complexity and volume manipulations.

The enhanced theta band activity was seen primarily at central and parietal sites with similar increases apparent at right frontal and right temporal recording sites. Increased mental workload in both of the traffic manipulations was associated with increased theta relative power. The high difficulty situations were always associated with significantly greater relative power than the low difficulty conditions, and in three of the seven comparisons the medium difficulty conditions were also enhanced relative to the low difficulty conditions. The relationship of theta activity to human mental state is not clear. Some authors report increased theta activity with increased cognitive activity while others report increased theta when cognitive demands decrease in vigilance situations. Lang and colleagues (Lang, Lang, Diekmann & Kornhuber, 1987; Lang, Lang, Kornhuber, Diekmann & Kornhuber, 1988) have reported increased theta over frontal sites during learning tasks compared to a baseline condition. Mecklinger and colleagues (Mecklinger, Kramer & Strayer, 1992; Pennekamp, Bosel, Mecklinger & Orr, 1994) have also reported increased theta in tasks requiring increased attention and memory load, while Gevins et al. (1979) and Ishihara and Yoshi (1972) reported increased theta in several tasks which required memory, geometric design evaluation, reading, and spatial processes. Gundel and Wilson (1992) and Pigeau, Hoffmann, Purcell and Moffitt (1987) found increased theta activity in response to increased task difficulty using several tasks. On the other hand, increased theta has also been reported in studies in which subjects become less vigilant due to the repetitive and long-term nature of their task. For example, Belyavin and Wright (1987) reported increases in theta activity with increased time on task. While seemingly contradictory, these results may actually indicate that increased theta activity reflects at least two different mechanisms, one which is related to certain classes of cognitive activity, for example, learning or complex task performance, and another mechanism which is determined by the arousal state of the subject such that decreased arousal and increased fatigue produce increases in theta activity. The topographic distribution of the these changes may be useful for determining the underlying cortical areas that are engaged in the task specific increased activity. If this premise is correct then not only the changes in theta but the location of change could be predictive of the nature of the cognitive demands experienced in a given situation.

The increased beta activity during the complexity manipulations was seen primarily over frontal and central sites. This is indicative of increased processing requirements in the complexity condition when several variables had to be dealt with during the simulation. In the volume conditions the primary task was to keep track of differing numbers of aircraft while the complexity condition added different types of aircraft and a wider range of ability levels of the pilots which may have caused the subjects to utilize different cognitive strategies. Since beta seems to be more sensitive to traffic manipulations this could be used to differentiate the type of processing rather than the level of difficulty of the task. There is relatively little in the literature concerning changes to beta due to manipulations of task difficulty. These data suggest that beta levels may be associated with the nature of the cognitive activity during complex task performance.

The changes in the alpha band were primarily represented by decreased relative power between the low and medium complexity conditions at sites F7, Fz and F4, and C4 while the cognitive demands of varying levels of traffic volume did not produce changes in alpha. There also was a significant workload related reduction in alpha from low to medium to high difficulty at the T6 electrode site. Numerous reports have shown decreased alpha with increased task difficulty. Earle and Pikus (1982) reported decreased alpha with increased difficulty of mental arithmetic tasks. Using several laboratory cognitive tasks, Gundel and Wilson (1992) and Pigeau, Hoffmann, Purcell and Moffitt (1987) found alpha to decrease with increased task difficulty. It is interesting to note that with the TRACON task which involved complex cognitive activity the significant decreases in alpha were between the low and medium difficulty levels and not between low and high.

The significant changes in delta activity were found at frontal and one left temporal electrode site (F3, Fz, F4, T3). This suggests that these results could be due to residual eye blink activity not removed by the eye blink correction procedure. However, this does not seem to be the case since the pattern of changes does not mirror the pattern of eye blink activity as one would expect. The patterns due to difficulty were different for the volume and complexity manipulations and were for the most part opposite of that found for the alpha band at the frontal sites. These data indicate that the relative delta power may be dependent on the specific requirements of the task. As was the case with the beta band, there is a paucity of discussion of this band in the literature and the results of the current experiment suggest that further examination of the delta band in complex cognitive situations is warranted.

The ability of a number of psychophysiological measures to distinguish between ATC task difficulty and traffic manipulations suggests that these measures can be used on-line to assess controller workload during normal working conditions. The physiological measures contribute information beyond that provided by performance and subjective measures alone and they have the advantage of being continuous and non-intrusive to the primary responsibilities of the controllers. In general, these data suggest that EEG measures may be more sensitive to task variables than either performance or subjective measures or the other physiological measures used in this study. If true, then the more sensitive EEG measures could be used to evaluate the relative contributions of workload variables not detected by other indexes. The effects of new procedures and equipment could be assessed by the use of physiological measures in conjunction with performance and subjective measures. Questions concerning the workload of different shifts, different ATC locations, and other related variables could be assessed with the same procedures. The feasibility of utilizing psychophysiological data to correctly classify operator workload in the laboratory (Wilson & Fisher, 1995) and in actual flight (Wilson & Fisher, 1991) has been demonstrated. Psychophysiological systems that can provide on-line reduction of multiple physiological variables and also provide estimates of operator workload are currently being developed (Wilson, 1994).

Even though the TRACON task faithfully mirrors several aspects of 'real-world' ATC, it does not incorporate all elements of ATC complexity. Despite these limitations, the results of this study support the utility of TRACON as a tool for conducting controlled studies of ATC workload. As Ackerman (1992) observed, the task helps 'bridge the gap' between the simple information processing tasks used in most laboratory studies and the complex tasks performed by ATCs. In addition, the use of active air traffic controllers as subjects enhances the external validity of these findings. Ackerman (1992) and Ackerman and Kanfer (1993) studied TRACON performance among college students and ATC trainees, respectively, but the current investigation is the first to use TRACON for workload assessment among active ATCs.

Acknowledgements

Dr. Brookings was serving on a National Research Council Research Associateship at Armstrong Laboratory when this research was conducted. The authors wish to acknowledge the assistance of Barbara Palmer and George Reis during data collection.

References

Ackerman, P.L. (1992). Predicting individual differences in complex skill acquisition: dynamics of ability determinants. *Journal of Applied Psychology, 77*, 598–614.

Ackerman, P.L., & Kanfer, R. (1993). Integrating laboratory and field study for improving selection: development of a battery for predicting air traffic control success. *Journal of Applied Psychology, 78*, 413–432.

Backs, R.W., Ryan, A.M., & Wilson, G.F. (1994). Psychophysiological measures of workload during continuous manual performance. *Human Factors, 36*, 514–531.

Belyavin, A., & Wright, N.A. (1987). Changes in electrical activity of the brain with vigilance. *Electroencephalography and Clinical Neurophysiology, 66*, 137–144.

Caldwell, J.A., Wilson, G.F., Centiguc, M., Gaillard, A.W.K., Gundel, A., Lagarde, D., Makeig, S., Myhre, G., & Wright, N.A. (1994). *Psychophysiological assessment methods*, AGARD Advisory Report 324. Paris: AGARD.

Carroll, D., Turner, J.R., & Hellawell, J.C. (1986). Heart rate and oxygen consumption during active psychological challenge: the effects of level of difficulty. *Psychophysiology, 23*, 174–181.

Carroll, D., Turner, J.R., & Rogers, S. (1987). Heart rate and oxygen consumption during mental arithmetic, a video game, and graded static exercise. *Psychophysiology*, *24*, 112–118.
Costa, G. (1993). Evaluation of workload in air traffic controllers. *Ergonomics*, *36*, 1111–1120.
Danaher, J.W. (1985). Human error in ATC system operations. *Human Factors*, *22*, 535–545.
Earle, J.B.B., & Pikus, A. (1982). The effect of sex and task difficulty on EEG alpha activity in association with arithmetic. *Biological Psychology*, *15*, 1–14.
Egelund, N. (1982). Spectral analysis of heart rate variability as an indicator of driver fatigue. *Ergonomics*, *25*, 663–672.
Federal Aviation Administration (1985). *National air-space system plan: facilities, equipment and associated development*. Washington, DC: Federal Aviation Administration.
Gevins, A.S., Zeitlin, G.M., Doyle, J.C., Yingling, C.D., Schaffer, R.E., Callaway, E., & Yeager, C.L. (1979). Electroencephalogram correlates of higher cortical functions. *Science*, *203*, 665–668.
Gratton, G., Coles, M.G.H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*, 468–484.
Gundel, A., & Wilson, G.F. (1992). Topographical changes in the ongoing EEG related to the difficulty of mental tasks. *Brain Topography*, *5*, 17–25.
Hart, S.G., & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In P.A. Hancock, & N. Meshkati (Eds.). *Human mental workload* (pp. 139–183). Amsterdam: North-Holland.
Harding, R.M. (1987). Human respiratory responses during high performance flight. In AGARDograph. AGARD-AG-312. Neuilly Sur Seine, France: NATO.
Hopkin, V.D. (1989). Implications of automation on air traffic control. In R.S. Jensen (Ed.). *Aviation psychology* (pp. 96–108). Brookfield, VT: Gower.
Ishihara, I., & Yoshi, N. (1972). Multivariate analytic study of EEG and mental activity in juvenile delinquents. *Electroencephalography and Clinical Neurophysiology*, *33*, 71–80.
Itoh, Y., Hayashi, Y., Tskui, I., & Saito, S. (1989). Heart rate variability and subjective mental workload in flight task validity of mental workload measurement using HRV method. In M.J. Smith, & G. Salvendy (Eds.). *Work with computers: organizational, management, stress and health aspects* (pp. 209–216). Amsterdam: Elsevier.
Kalsbeek, J.W.H., (1971). Standards of acceptable load in ATC tasks. *Ergonomics*, *14*, 641–650.
Lang, M., Lang, W., Diekmann, V., & Kornhuber, H.H. (1987). The frontal theta rhythm indicating motor and cognitive learning. Current trends in event-related potential research. In R. Johnson, J.W. Rohrbaugh, & R. Parasuraman (Ed.). *Electroencephalography and clinical neurophysiology*, (suppl. 40) 322–327.
Lang, W., Lang, M., Kornhuber, A., Diekamnn, V., & Kornhuber, H.H. (1988). Event-related EEG-spectra in concept formation task. *Human Neurobiology*, *6*, 295–301.
Mecklinger, A., Kramer, A.F., & Strayer, D.L. (1992). Event related potentials and EEG components in a semantic memory search task. *Psychophysiology*, *29*, 104–119.
Mogford, R.H., Murphy, E.D., & Guttman, J.A. (1994). Using knowledge exploration tools to study airspace complexity in air traffic control. *The International Journal of Aviation Psychology*, *4*, 29–45.
Morrison, R., & Wright, R.H. (1989). ATC control and communications problems: an overview of recent ASRS data. *Proceedings of the Fifth International Symposium on Aviation Psychology*, 902–907.
Nolan, M.S. (1990). *Fundamentals of air traffic control*. Belmont, CA: Wadsworth.
Opmees, C.H.J.M. & Krol, J.P. (1973). Towards an objective assessment of cockpit workload: I. Physiological variables during different flight phases. *Aerospace Medicine*, *44*, 527–542.
Pennekamp, P., Bosel, R., Mecklinger, A., & Orr, H. (1994). Differences in EEG-theta for responded and omitted targets in a sustained attention task. *Journal of Psychophysiology*, *8*, 131–141.
Pigeau, R.A., Hoffmann, R., Purcell, S., & Moffitt, A. (1987). The effect of endogenous alpha on hemispheric asymmetries and the relationship of frontal theta to sustained attention. In *Electric and magnetic activity of the central nervous system: research and clinical applications in aerospace medicine*, AGARD-CP-432, 20/1–20/16. Neuilly sur Sein, France: NATO.

Redding, R.E. (1992). Analysis of operational errors and workload in air traffic control. *Proceedings of the 36th Annual Meeting of the Human Factors Society*, 1321–1325.

Roscoe, A.H. (1987). In-flight assessment of workload using pilot ratings and heart rate. In A.H. Roscoe (Ed.). *The practical assessment of pilot workload*, AGARDograph No. 282 (pp. 1–14). Neuilly sur Seine, France: AGARD.

Speyer, J.J., Fort, A., Fouillot, J., & Blomberg, R.D. (1988). Dynamic methods for assessing workload for minimum crew certification. In A.H. Roscoe, & H.C. Muir (Eds.). *Workload in transport operations* Report number 1B 316-88-06 (pp. 196–220). Cologne: DFVLR.

Stager, P., Hameluck, D., & Jubis, R. (1989). Underlying factors in air traffic control incidents. *Proceedings of the 33rd Annual Meeting of the Human Factors Society*, pp. 43–46.

Stern, J.A., & Dunham, D.N. (1990). The ocular system. In J.T. Cacioppo, & L.G. Tassinary (Eds.). *Principles of psychophysiology: physical, social and inferential elements* (pp. 513–553). Cambridge: Cambridge University Press.

Turner, J.R., & Carroll, D. (1985). Heart rate and oxygen consumption during mental arithmetic, a video game, and graded exercise: further evidence of metabolically-exaggerated cardiac adjustments? *Psychophysiology, 22*, 261–267.

Wientjes, C.J.E. (1992). Respiration in psychophysiology: methods and applications. *Biolgical Psychology, 34*, 179–204.

Wientjes, C.J.E. (1993). *Respiration and stress.* Doctoral dissertation. Tilburg: University of Tilburg.

Wilson, G.F. (1992). Applied use of cardiac and respiration measures: practical considerations and precautions. *Biological Psychology, 34*, 163–178.

Wilson, G.F. (1993). Air-to-ground training missions: a psychophysiological workload analysis. *Ergonomics, 36*, 1071–1087.

Wilson, G.F. (1994). Workload assessment monitor (WAM). *Proceedings of the Human Factors and Ergonomics Society*, 944.

Wilson, G.F., & Eggemeier, F.T. (1991). Psychophysiological assessment of workload in multi-task environments. In D.L. Damos (Ed.). *Multiple-task performance* (pp. 229–260). London: Taylor and Francis.

Wilson, G.F., & Fisher, F. (1991). The use of cardiac and eye blink measures to determine flight segment in F4 crews. *Aviation, Space, and Environmental Medicine, 62*, 959–962.

Wilson, G.F., & Fisher, F. (1995). Cognitive task classification based upon topographic EEG data. *Biological Psychology, 40*, 239–250.

Wilson, G.F., Purvis, B., Skelly, J., Fullenkamp, P., & Davis, I. (1987). Physiological data used to measure pilot workload in actual flight and simulator conditions. *Proceedings of the Human Factors Society* (pp. 779–783). Santa Monica, CA: Human Factors Society.

Wilson, G.F., & Oliver, C.G. (1991). PATS: Psychophysiological Assessment Test System. In E. Farmer (Ed.). *Stress and error in aviation: proceedings of the XVIII WEAAP Conference* (vol. 2, pp. 15–25). Worcester: Billing & Son Ltd.