2005

Operators' Time Perception Under Stress

R. V. Nayeem

T. Oran-Gilad

P. A. Hancock

OPERATORS' TIME PERCEPTION UNDER STRESS

R.V. Nayeem[1], **T. Oron-Gilad**[1&2], & **P. A. Hancock**[1&2]
Department of Psychology[1] and [2]Institute for Simulation and Training
University of Central Florida

Time perception is extremely important to the understanding, design and use of complex aviation systems. This experiment focused on differences in time estimation, flight performance, and monitoring tasks. In a between-subjects experiment, participants navigated through a flight scenario while monitoring a switch and listening to white noise at either 55dBA or 85dBA. Flight performance data and monitoring data were collected throughout the task. Participants also completed the NASA-TLX and the DSSQ-S. Statistical analyses showed that the noise condition did not significantly affect workload, monitoring abilities, task completion and subjective stress questionnaires for the dual task. However, the 85dBA condition significantly affected prospective time estimation. These results suggest that the dual task was not demanding enough, and the stress was not adequate to push participants out of the comfort range and experience a performance decrement.

Introduction

The temporal domain is stress-sensitive in a similar manner to the spatial domain and comparable narrowing occurs, resulting in distortions of perceived time (Hancock & Weaver, 2005). Understanding time perception and time distortion under stress is of high importance for operating complex systems. Time perception affects how operators react to visual, auditory and tactile alarms. If time estimation is inaccurate due to stress, the alarm may go unattended for a critical length of time which exceeds the time available for solution. Time estimation is affected by the stress and mental effort that operators experience. If stress conditions are sufficiently high, such conditions will induce time distortion (Block, Zakay, & Hancock, 1999). This has been seen in Eastern Airlines Flight 401 in the Everglades where the crew got fixated on a landing light, and in the infamous John Denver fatal aircraft accident.

One possible solution to this problem is adaptive automation. Adaptive function allocation or dynamic automation is an approach in which control of tasks dynamically shifts between humans and machines, and is an alternative to traditional static allocation in which task control is assigned during system design and remains unchanged during operations. With adaptive automation, if operators are under too much stress to respond optimally to the aircraft, tasks will automatically get shifted to the automation rather than the operator. If automation states are switched too quickly, operators can become confused and decision making can be degraded, and allocation induced oscillations car occur (Hancock & Scallen, 1996). Also, when a trigger is operator-driven, the added workload and decision making can hinder performance (Hildebrandt & Harrison, 2002). Using dynamic function scheduling, an operator can prioritize re-allocation by assigning tasks both a temporal value, when the task should be re-allocated along the system timeline, and a qualitative value, what contribution the task makes to the system goal and the quality of the eventual solution. The Hancock and Warm (1989) model shows different zones in which a person can be when stressed.

The purpose of this study was to investigate differences on time estimation due to differences in the audible level of white noise, while performing a flight navigation task and a dual monitoring task. Stress was manipulated by task complexity, time constraints, and the addition of one of two levels of white noise: standard high frequency white noise at either 55dBA or 85dBA. This study attempted to investigate time distortion associated with stress and monitoring tasks in a basic flight navigation task. It was anticipated that participants would be pushed out of the comfort zone (Hancock & Warm, 1989) and would experience performance decrements to workload, monitoring abilities, subjective stress questionnaires, and prospective time estimation. However, given the relatively simple nature of the task and the qualifications for participation, it was expected that the noise would not affect task completion.

Method

Participants. 32 pilots, 28 male and 4 female students enrolled in psychology and human factors courses at Embry-Riddle Aeronautical University. The mean age was 21.19 (SD=3.61). Participants were required to have at least a private pilot's license. Total hours as PIC

[1] Corresponding Address: Razia V. Nayeem, Department of Psychology, PO Box 161390, University of Central Florida, Orlando, FL 32816-1390 E-mail: rnayeem@ist.ucf.edu

ranged from 40 to 700, with a mean of 133.1 (SD=134.1).Students who chose to participate in this research effort were compensated via extra credit. The participants all filled out personality questions as well. On a rating scale of 1-5, they uniformly rated themselves as able to pace themselves to get things done (M=4.0, SD=.6), able to handle stress (M=2.3, SD=.9), as moderately methodical people (M=2.4, SD=.8), and as having fast-paced lives (M=3.3, SD=1). About half of the participants responded that they waste a lot of time before settling down to work, and about half responded that they do not (M=3.9, SD=.7).

Apparatus and Tasks. The experimental setup consisted of two independent tasks generated in the AirBook® by Simigon F-16 flight simulator: a) the monitoring task and b) navigation task. The flight simulator was installed on a Dell laptop model # PP08L, with a 17" screen. Static white noise was emitted through headphones plugged into the computer speakers.

White Noise. There were two levels of white noise – 55dB and 85dB. The sound was delivered via Maxell headphones, model # HP/NC-II, and was measured before each participant via a Sper Scientific sound meter, model # 840029.

Flight Controls. Participants manipulated the flight environment using a GF yoke, model # G60503A.

Navigation Task. The navigation task objective was to take off from an airport, navigate to three waypoints, and land at the initial airport, as shown in Figure 1. A skyscraper marked each waypoint. After the participant crossed over the waypoint at 5000 feet MSL, the avionics of the aircraft changed to guide the participant to the next waypoint. When a participant crossed all three waypoints, the task was considered successfully complete. The waypoints were strategically placed throughout Arizona terrain. Two different navigational missions were used for the practice trial and the dual task trial so that participants would not become too familiar with any one mission. The navigation task difficulty did not vary from mission to mission. Additionally, at the end of the second mission, participants were asked to estimate how long they spent completing the dual task.

Monitoring Task. The monitoring task consisted of periodically checking the master switch in the aircraft console. The switch was set to turn to the "off" position 30 seconds after takeoff, and 30 seconds after the participant crossed the second waypoint. Switch states are shown in Figure 2. Reaction time and accuracy were recorded for the monitoring task. If the participant did not turn set the switch to the "off" position after the first change, the switch remained in this position for the entire flight. Similarly, if a participant did not set the switch to the "off" position after the second change, the switch remained in this position for the remainder of the flight. If participants did not cross the second waypoint at all, or higher than 5000 feet, the switch remained in the position it was prior to these events.

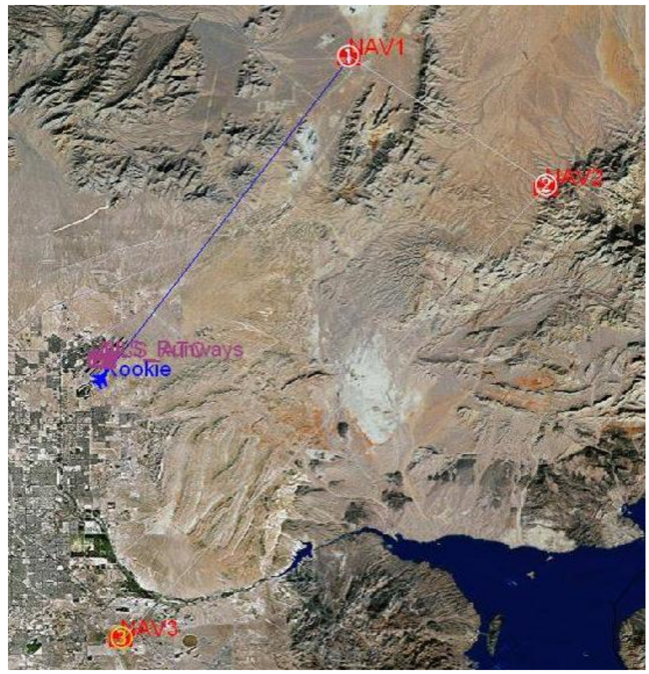

Figure 1. *The navigation route in the AirBook® environment. The objective of the navigation task was to take off from an airport, navigate to three waypoints, and land at the initial airport.*

(a) (b)

Figure 2. *The Monitoring task required pilots to respond to the status of the master switch. In (a) the master switch is armed in (b) it is off.*

Time Estimation. During the flight participants were required to estimate the duration of the task they performed (Zakay & Block, 2004). This prospective duration estimation can serve as a secondary task for workload measurement (Block, Zakay, & Hancock, 1999).

Subjective Measures. Questionnaires included a demographic questionnaire, the DSSQ-S (Matthews, et al., 1999), and a computer version of the NASA-TLX (Hart & Staveland, 1988).

Procedure. After a brief description of the experiment by the experimenter, participants signed an informed consent form and completed a demographic questionnaire. Once informed consent was received, participants were familiarized with the AirBook® navigation task. They were given written and verbal instructions on how to maneuver through the AirBook® environment using the keyboard and the yoke provided for flight inputs. Then, participants had a five-minute practice. After the practice session, participants completed an initial NASA-TLX to be used as a baseline, and the DSSQ-S pre. Participants were then given the option to take a five-minute break. After this, participants were given brief instructions, and the dual task began. Participants navigated through a predefined route while performing the monitoring task for ten minutes. Performance on the monitoring task was recorded. In addition, during the dual task, participants received white noise through headphones. Half of the participants received noise at 55 dB, and the other half received noise at 85 dB. After completion of the dual task, participants estimated how long they were doing the dual task, completed a second NASA-TLX and the DSSQ-S post and were completely debriefed, having all their questions answered.

Results

All analyses were conducted using SPSS 11.5 for Windows. All alpha levels were set to .05.

Flight Performance

Navigation Task. Twenty-three of the 32 participants (72%) successfully completed the navigation task. Seven of those who did not complete the task had lower total flight hours than the mean. The other two had 140 and 700 flight hours, respectively. A one-way ANOVA on task completion with noise as a between subject variable was not significant, $F(1,30) = 0.146$, $p = 0.705$.

Monitoring Task. The mean reaction time for alarm 1 was 100.28 seconds (SD=186.36), and the mean reaction time for alarm 2 was 240.96 (SD=93.07). Twenty-eight of the 32 participants responded to alarm 1, and 23 of the 32 participants responded to alarm 2. All participants that failed to respond to alarm 1 had less flight hours than the mean except one participant, who had 700 hours. All participants that failed to response to alarm 2 had less flight hours than the mean except two participants, who had 140 hours and 700 hours.

A paired-samples T-test was conducted on the reaction times for alarm 1 and alarm 2 revealed that there was a significant difference in response time, $t(22) = -11.6$, $p < .0005$, Cohen's $d = -3.3$, effect size $r = -.86$. This difference is probably due to the fact that as the flight progressed participants paid less attention to the monitoring task. Two one-way ANOVAs on response to the first and the second alarm with noise as a between subject variable were not significant, $F(1,30) = 1.111$, $p = 0.3$, $F(1,30) = 0.146$, $p = 0.705$, respectively. The mean time to turn off alarm 1 in the 55dBA condition was 71.69 seconds (SD=138.9), and 128.88 seconds (SD=225.3) in the 85dBA condition. The mean time to turn off alarm 2 in the 55dBA condition was 208.75 (SD=94.5) seconds, and 276.09 (SD=81.4) seconds in the 85dBA condition. Two one-way ANOVAs on response time to alarm 1 and alarm 2 with noise as the between-subject variable were also not significant, $F(1,30) = 0.747$, $p = 0.394$, $F(1,21) = 3.322$, $p = 0.083$, Cohen's d = , effect size r = , respectively.

Stress and Workload.

NASA-TLX. A repeated-measures ANOVA before the task (M=52.40, SD=15.68) and after the task (M=53.25, SD=20.58) with noise as a between subjects variable revealed that the overall pre- and post- TLX measures were not affected by noise, $F(1,30) = 0.113$, $p = .739$.

A one-way ANOVA of the TLX scores after completion of the task indicated that there were no significant difference in NASA-TLX scores between the 55dBA condition and the 85dBA condition, $F(1,30) = 0.908$, $p = .348$. Additionally, the six sub-scales were analyzed for significance and effect size using one-way ANOVAs. The statistical results can be seen in Table 1.

NASA-TLX after completion of task	55dBA	85dBA	F	df	p	Cohen’s d	Effect Size r
Global TLX	M=55.7 SD=21.5	M=50.8 SD=20.0	0.908	(30)	.348	.24	.12
Mental Demand	M=66.6 SD=24.4	M=58.4 SD=24.3	1.256	(1,30)	.271	.33	.16
Physical Demand	M=30.0 SD=18.3	M=35.5 SD=24.3	1.533	(1,30)	.225	-.25	-.12
Temporal Demand	M=56.9 SD=21.1	M=41.6 SD=25.5	0.058	(1,30)	.811	.65	.31
Performance	M=45.0 SD=33.2	M=36.3 SD=18.5	2.837	(1,30)	.103	.31	.15
Effort	M=70.3 SD=22.7	M=60.3 SD=30.0	2.034	(1,30)	.164	.38	.18
Frustration	M=35.6 SD=26.6	M=39.4 SD=30.9	0.244	(1,30)	.625	-.13	-.06

Table 1. *Analysis of the differences in TLX scores after completion of the dual task by noise condition .*

Time Estimation. The actual duration of the task was 10 minutes. The estimated times had a mean of 7.7 minutes (SD=3.0). Zakay (1998) suggested that underestimation is more likely to occur when time estimation is not the primary task at hand. As more attention is being focused on temporal information processing, more time signals are processed and the judgment is more likely to be accurate. Allocating fewer attentional resources to temporal information processing causes fewer time signals to accumulate, resulting in a decrease in the estimated duration (Zakay & Block, 2004).

Time estimates were translated into the duration judgment ratio (DJR; Block, Zakay, & Hancock, 1999). Correlations were calculated for the DJR (M=77.3, SD=29.7) and the NASA-TLX score for participants in the 55dBA condition after the dual task, and the DJR and the NASA-TLX score for participants in the 85dBA condition to determine whether or not time estimates were correlated with noise condition. The correlation between the DRJ and the 55dBA condition was not significant, $r = .164$, $p = .543$. However, the correlation between the DJR and the 85dBA condition was significant, $r = -.57$, $p<.05$. This suggests that the 85dBA noise condition affected the DJR.

DSSQ-S. The DSSQ-S is scored into engagement, pre (M=21.3, SD=3.7) and post (M=22.7, SD=4.0), distress, pre (M=8.4, SD=5.8) and post (M=7.0, SD=4.4), and worry, pre (M=6.6, SD=4.6) and post (M=5.8, SD=4.3). Paired-samples t-tests were performed for engagement pre-post, distress pre-post, and worry pre-post, $t(31) = -1.97$, $p = .058$, Cohen’s $d = -.33$, effect size $r = -.17$, $t(31) = 1.48$, $p = .15$, Cohen’s $d = .27$, effect size $r = .13$, and $t(31) = 1.45$, $p = .157$, Cohen’s $d = .18$, effect size $r = .09$, respectively. The difference in the pre and post engagement scores was not statistically significant, though it was very close. However, the effect size was not large enough to warrant more examination.

Repeated-Measures. ANOVAs were performed for each sub-scale with noise as the between subjects variable, $F(1,30) = 3.79$, $p = .061$ Cohen’s $d = .37$, effect size $r = .18$, $F(1,30) = 2.15$, $p = .153$, Cohen’s $d = -.23$, effect size $r = -.11$, and $F(1,30) = 2.104$, $p = .157$, Cohen’s $d = .06$, effect size $r = .03$, respectively. Thus, noise was not a moderating condition for the DSSQ-S sub-scales.

Correlations were performed for the DSSQ-S measures before and after the task with noise condition and pre- and post- TLX global and performance scores using the Spearman ranking coefficient. For the correlations computed, the only ones that were statistically significant at the .05 level was the correlation between the post-task NASA-TLX global scores and the DSSQ-S post-task distress measure, $\rho = .384$, $p = .03$, pre-task distress and pre-task mental demand, $\rho = .439$, $p = .012$, pre-task distress and pre-task frustration, $\rho = .434$, $p = .013$, and post-task distress and post-task effort, $\rho = .377$, $p = .033$. These correlations indicate that there is a strong relationship between overall workload, post-task effort and pre-task performance with distress after the task. Additionally, these correlations indicate that there is a strong relationship between pre-task mental demand and pre-task frustration with pre-task distress.

Discussion

The majority of the pilots (71.9%) were able to complete the navigation task without any difficulties. However, since the workload estimates indicated that participants were not experiencing a heavy workload, one would think that close to 100% of the participants would have finished the task. Of the participants who could not complete the task, 4 were in the 55dBA condition and 5 were in the 85dBA condition. This suggests that task completion was not strongly affected by noise condition. Thus, task completion must be mediated by other moderators. When looking at the DSSQ-S engagement scores after the task, the mean was 22.6 compared to a maximum possible score of 28. Thus, the average score was only approximately 80% of the possible score. This suggests that participants were simply not engaged enough in the task to complete the task. To remedy this in the future, the navigation task will be comprised of more tasks that are shorter and a more demanding navigation scenario.

Performance on the monitoring task varied among participants, but this variation did not seem to correlate with the noise condition. The data suggests that some people simply ignored the alarms, possibly because there were no consequences, and the changes of the switch were very subtle. To better this task in the future, participants will be given a certain amount of time to comply with the alarm, and after that time expires, a visual alert will be shown to direct attention to the switch.

Interestingly, the mean time to turn off Alarm 2 in both conditions was more than twice the time it took to turn off Alarm 1. This suggests that participants were not monitoring the switch as closely throughout the flight as they were in the initial phases of flight. This is consistent with other monitoring task experiments, in which participants experience a performance decrement after a certain period of time.

The actual results were somewhat different from the expected results, in that workload, monitoring abilities, and subjective stress questionnaires were not affected by noise. However, as expected, task completion and time estimates were affected by noise. In the Hancock and Warm (1989) model, this means the stress was not sufficient to push participants out of the comfort zone, which is reflected in the lack of performance decrement in the flight parameters. Future work will address the issues that were problematic in this study.

Acknowledments

This work was supported by the Department of Defense Multidisciplinary University Research Initiative (MURI) program administered by the Army Research Office under Grant DAAD19-01-1-0621, Dr. Elmar Schmeisser, Technical Monitor, P.A. Hancock, Principal Investigator. The views expressed in this work are those of the authors and do not necessarily reflect official Army or Department of Defense policy.

References

Block, R.A., Zakay, D. & Hancock, P.A. (1999). Developmental changes in human duration judgments: a meta-analytic review. Developmental Review, 19(1), 183-211.

Chignell, M.H., & Hancock, P.A. (1987). A comparison of mental workload and available capacity in complex person-machine systems: A fuzzy-set theoretic approach. In: W. Karwowski and A. Mital (Eds.). Applications of Fuzzy-Set Theory in Human Factors. (pp. 271-288), Amsterdam: Elsevier Publishers.

Gluckman, J. P., Morrison, J. G., & Deaton, J. E. (1991). Complex task performance as a basis for developing cognitive engineering guidelines in adaptive automation. Proceedings of the Human Factors Society 35th Annual Meeting, Santa Monica, CA: Human Factors Society

Hancock, P.A. & Warm, J.S. (1989). A dynamic model of stress and sustained attention. Human Factors, 31, 519-537.

Hancock, P. A., & Weaver, J. L. (2005, in press). Temporal distortions under extreme stress. Theoretical Issues in Ergonomic Science.

Hildebrandt, M. & Harrison, M. (2002). The temporal dimension of dynamic function allocation. Paper presented at the Eleventh European Conference on Cognitive Ergonomics.

Parasuraman, R. & Mouloua, M. (1996). (Ed.). Automation and human performance. Lawrence Erlbaum: Mahwah, N.J.

Scallen, S.F. & Hancock, P.A. (2001). Implementing adaptive function allocation. The International Journal of Aviation Psychology, 11(2), 197-221.