An Eye-tracking Study of Information Sampling and Decision-making Under Stress: Implications for Alarms in Aviation Emergencies

Aleksandra Stankovic[a], Michael R.F. Aitken[b], Luke Clark[c]

[a] *Department of Psychology, University of Cambridge, Cambridge, UK*
[b] *Department of Psychology, University of British Columbia, Vancouver, B.C.*
[a] *Department of Psychology, King's College London, London, UK*

The objective of this study was to probe the cognitive processing of cockpit warning displays in emergency situations by assessing the effects of acute stress on information sampling and decision-making using eye tracking equipment. A novel image-matching computer task based on the Matching Familiar Figures Task (MFFT) was designed to provide a measure of cognitive impulsivity. The stress induction procedure involved a challenging manual response task coupled with unpredictable and uncontrollable bursts of loud, aversive noise, and a matched neutral control task. Healthy participants (n=40) completed the task under two conditions: neutral and stress. Participants under stress made more image matching errors and visually sampled less in terms of both saccade count and dwell time on the MFFT, and made a greater number of responses without having first sampled all information areas displayed on the screen at least once ('premature closure'). The findings of this study may have useful implications for the design of visual information displays across a variety of industries, particularly aviation.

INTRODUCTION

This research is motivated by recent major aircraft accidents which have highlighted the importance of meaningfully displayed emergency information during periods of increasing criticality. For instance, during the final few minutes of the Air France 447 flight in 2009, over 20 visual and auditory warning indications were triggered. The overwhelming number of textual messages and the intermittent nature of the audible alarms added to the confusion and stress in the cockpit, and the lack of clear information presentation coupled with the ordering of the error messages may have impaired the crew's ability to make sense of the situation. In this particular accident, ice crystal blockage of the pitot tubes resulted in unreliable airspeed indications, initiating the disconnect of the autoflight system; at the same time, a cascade of over a dozen additional error messages was initiated to signal problems with related systems which rely on accurate airspeed input, without ever foregrounding the root cause of these problems (frozen pitot tubes). The crews' inability to correctly assess the situation and take corrective action led to the initiation of a stall due to improper control inputs, resulting in the tragic crash. The objective of this research, therefore, was to examine cognitive subprocesses underlying how humans sample information to reach conclusions, and how stress may modulate these processes, with the goal of informing future design of warning systems.

Stress, Decision-Making & Cognitive Impulsivity

Current research has shown the influence of mood states and emotional experiences, including stress, on behavior and decision making (Bechara & Damasio, 2005; Leith & Baumeister, 1996; Starcke, Wolf, Markowitsch, & Brand, 2008; Preston, Stansfield, Buchanan, & Bechara, 2007). Previous investigations have indicated a stress-induced propensity for impulsive decision-making, as reflected by hasty patterns of selection. For example, the work of Porcelli and Delgado (2009) demonstrated that acute stress can enhance the influences of behavioral biases and lead to a reliance on automatized reactions. Findings by Braver, Paxton, Locke, and Barch (2009) have pointed to vigilance and attention as being important cognitive processes that may impact decision making under different mood states. Furthermore, Keinan (1987) showed that stress-exposed participants spent less time evaluating information than non-stressed controls, and that they demonstrated the phenomenon of premature closure—the tendency to reach a decision before sampling all available information (Janis & Mann, 1977).

Building upon these previous investigations, this study explores the effects of acute stress on decision-making as demonstrated by control over cognitive impulsivity, which is defined broadly as the dimension of executive functioning governing cognitive inhibition of overly hasty decision-making (Donfrancesco, Mugnaini, & Dell'Uomo, 2005).

We propose that both cognitive impulse control and measures of attention and information sampling thoroughness will be significantly reduced following stress induction.

METHODS

Modified Matching Familiar Figures Task

This study employed a modified version of the Matching Familiar Figures Test (MFFT), an established measure of cognitive impulsivity (Kagan, Rosman, Kay, Albert, & Phillips, 1964; Cairns, 1977). In this task, subjects are presented with one template picture and a series of visually-similar options, among which is an identical match; they are instructed to identify the correctly matching figure.

Response times on this task are inversely related to choice accuracy, providing a measure of cognitive impulsivity in the form of the speed-accuracy trade-off (Kagan 1966). Several studies, however, have criticized currently-used versions of the MFFT (Banilivy & Gilliland, 1980; Block, Block, & Harrington, 1974; Kojima, 1976), as the task employs recognizable figures such as domestic animals or household objects, which may carry emotional or memory associations. Furthermore, non-match options vary in unsystematic ways, resulting in a lack of standardization across trials in degree of difference between options and templates.

For these reasons, a modified version was developed for this investigation which retained the typical task instructions of the standard MFFT but incorporated abstract figures to offer non-emotional stimuli with visual complexity that could be varied in a systematic and controlled way. These stimuli comprised four sub-elements (e.g. symbols such as “%” and “&”) presented in four different colors. The template image was presented in the center of the screen (surrounded by an orange box), and the four options were simultaneously presented at each corner of the template (so that each option was equidistant from the template, for eye movement purposes). Only one option was an exact visual match to the target (see Figure 1 for a sample trial illustration).

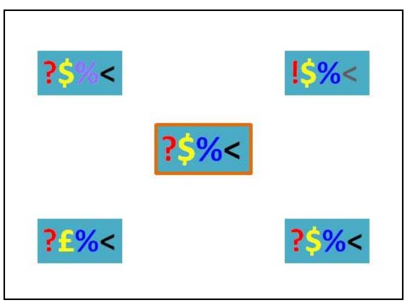

Figure 1. Illustration of a sample trial from the modified MFFT design.

The three non-match options were constructed as follows: one distractor option was composed of three symbols identical to those found in the template image and one symbol that was visually similar to, but distinct from, the fourth figure in the template sequence (for instance, the symbol “<” was substituted by “>”); one had the identical sequence of the template but differed by the color of a single symbol (pink being replaced by red, for example); and the last option presented a sequence in which one figure differed in color and one differed in symbol from the template. This pattern of systematic variation was created in order to discourage search strategies based on encoding single elements of the template, and to ensure uniformity from trial to trial. Across the trials, the position of the match relative to the template was equally distributed to each corner and randomly assigned. After a practice trial, 40 test trials were presented in a randomly generated order sequence. The task duration was approximately 5 min.

Eye-Tracking Equipment

While behavioral and performance measures offer great insight into the cognitive effects of stress, the use of head-mounted eye-tracking equipment in this study offers the additional benefit of allowing the monitoring of saccade (eye movement) counts, fixation areas, and dwell times (duration of gaze fixation on a given area) to give an impression of how stress may affect attention and strategies of information sampling leading up to decision-making. Pupillometry data collected by the eye-tracking equipment may also be used to assess magnitude of stress response, as pupil dilation has been shown to reflect arousal levels (Bradley, Miccoli, Escrig, & Lang, 2008; Beatty & Lucero-Wagoner, 2000).

In this study, participants wore a head-mounted eye-tracker (EyeLink I) for the duration of the session to record eye movements and pupil dilation. The eye-tracker sampled at a rate of 250 Hz and 0.2° of resolution. Subjects were seated approximately 1 ft away from a color monitor computer screen (21 inches, resolution set to 1024:768 pixels). Before beginning the eye movement tasks, a calibration and validation procedure was completed for the EyeLink I. Data was recorded from both eyes when available, or, in cases when stable binocular recording was not sustainable, from whichever eye showed the strongest recording capability at calibration. Analysis of eye movement data was performed on recordings from the right eye by default.

During the MFFT, a drift correction was performed immediately before each trial to ensure accurate equipment calibration. Drift correction was not performed during the stressor task and neutral response task (see below) as this would impose a delay on each trial, and the rapid pace of the task was considered a critical feature for the stress effect. Pupillometry was collected during the stressor task and neutral response task, in order to validate the stress manipulation, but eye movements were not obtained. Pupil size data was sampled at 250 Hz and was collected from the onset of the trial until the feedback sound was given. Values were measured in terms of total pupil area using Eyelink II software-determined arbitrary integer units, with typical pupil area being between 800 and 2000 units (SR Research Ltd, 2002). The period for the stress induction response task was maximally 1450 ms, and for the control version 2000 ms at most (exact trial timings depended upon the speed of the participant’s response).

Eye movement data was processed online during the decision-making task using SR Research DataViewer software. Eye movements were recorded from trial onset until the decision response. Movements were assessed within and between pre-designated fixed areas of interest where the target and four options appeared on the screen.

Stressor

Due to the nature of the eye-tracking equipment involved, a computer-based tone-avoidance stressor was chosen for this study, with a simple no-tone neutral response task as the control condition. Short-term aversive noise has been validated as an effective acute stressor: for example, Hébert et al. (2009) found that exposure to a 20 min cycle of noise at amplified low frequency resulted in a sustained increase in salivary cortisol and self-reported stress in healthy participants.

The stress induction procedure used in this study was modeled on a tone avoidance reaction-time task used by Snieder et al. (1997). The task consisted of 60 trials (plus 5 practice trials). On each trial, a fixation cross (500 ms) was followed by an "X" flashing in one of the four corners of the monitor. Participants were instructed to hit a button on a keypad corresponding to the diagonally *opposite* corner from where the "X" appeared (see Figure 2 for an illustration of the task). In the stress condition, incorrect responses received a loud (90 db) 500 ms burst of white-noise, presented through headphones to ensure standardized volume. In the neutral procedure, incorrect responses received a short, quiet buzz. Correct responses in both conditions received a soft chime. In the neutral procedure, the response window for the diagonal key press was 1000 ms, making the task fairly easy. In the stress procedure, this response window was only 450 ms, thus increasing the speed and difficulty of the task, and increasing errors. This added a dimension of frustration and experienced failure to the stress induction procedure.

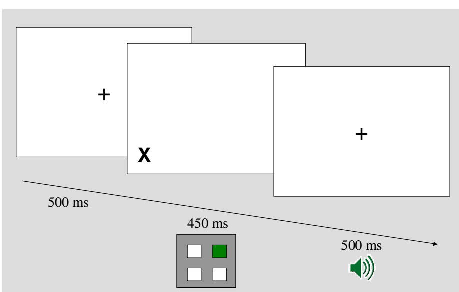

Figure 2. Illustration of the tone avoidance stress induction procedure.

Subjects and Experimental Design

Forty healthy participants (12 male, 28 female) aged 19 to 55, completed a one-hour long study conducted in two blocks. Participants were free of current or prior history of psychiatric illness, and were recruited from a university sample to take part in a study involving stress and decision-making. The study was approved by the local Human Research Ethics Committee, and all participants gave written informed consent prior to engaging in the project.

Participants were randomly assigned to one of two conditions (20 per group). Those in condition 1 were first fitted with eye-tracking equipment and completed the "neutral" response procedure followed by the modified figure-matching task. Then, following a rest, these participants completed the tone avoidance stress procedure, followed by a secondary behavioral task (which is outside the scope of the current report and therefore not included in this presentation) under stress. Subjects in condition 2 were assigned to the reverse pairing, completing the neutral response procedure followed by the secondary task; then, after a rest period, they completed the stress procedure followed by the matching task. This design enabled between-subjects comparisons to be made for the effect of stress on each of the two tasks, while avoiding the possibility of practice effects by only having each participant perform each task once. All participants were thoroughly debriefed after completing the study.

RESULTS

Psychophysiological Measures

Pupillometry measures (pupil size area) were analyzed during the neutral and stress-induction tasks and during the MFFT. Average pupillometry readings across all trials were higher during the stress-induction task ($M = 1031.94$; $SD = 354.50$) compared to the neutral task ($M = 924.00$; $SD = 393.95$), $t(39) = -2.47$, $p = 0.02$. These pupillometry responses were short-lived, as the pupillometry data across all trials of the MFFT did not differ significantly between the neutral condition ($M = 1010.60$; $SD = 509.23$) and the stress condition ($M = 1056.65$; $SD = 555.61$), $t(35) = -0.26$, $p = .79$

	Neutral, Pre-task Mean (SD)	Neutral, Post-task Mean (SD)	Stress, Pre-task Mean (SD)	Stress, Post-task Mean (SD)
Heart Rate	72.28 (11.94)	73.48 (12.17)	72.37 (11.59)	72.07 (10.68)
Systolic BP	116.48 (8.41)	118.03 (10.78)	117.93 (10.03)	119.07 (10.10)
Diastolic BP	76.38 (8.38)	77.24 (9.29)	76.87 (9.37)	76.87 (10.10)

Table 1. Mean and standard deviation values for HR and blood-pressure before and after the neutral and stress-induction versions of the tone avoidance task.

Modified Matching Familiar Figures Task

Behavioral Results. Two participants with an excessive number of incorrect responses on the modified MFFT (>16 incorrect, equaling more than 1/3 responses wrong) were excluded from analysis for having either not understood the directions or not followed the task instructions, and behavioral responses were unavailable for a further subject due to a recording malfunction. Eye-tracking data were unavailable for one subject.

Participants committed significantly more errors on the matching task in the stress condition ($M = 5.05$; $SD = 3.61$) compared to the neutral condition ($M = 2.89$; $SD = 2.45$), $t(35) = -2.12$, $p = .04$ (Figure 3). There was no significant difference between latency to response times (in seconds) in the neutral condition ($M = 5.51$; $SD = 1.31$) compared to the stress condition ($M=5.23$; $SD =1.97$), $t(35) = 0.51$, $p =.62$.

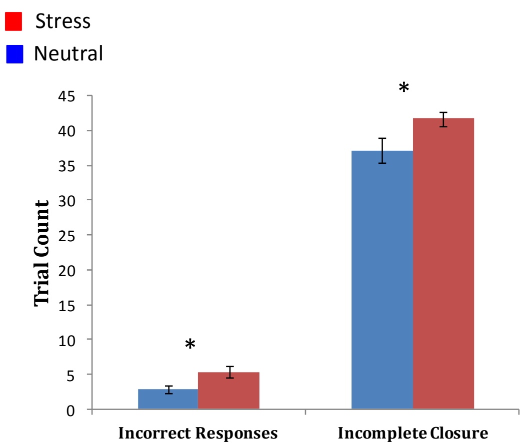

Figure 3. More errors were committed on the modified Matching Familiar Figures Task under stress. Under stress, match decisions were reached more often before all four option images had been sampled at least once ("incomplete closure," Keinan, 1987).

Eye Tracking Effects. Average saccade measures (defined as a movement in focus of the eye from one physical point on the screen to another) were examined separately for the template and 4 options and were compared between the neutral and stress conditions using independent samples t-tests. In the stress condition, participants sampled from the four match options significantly fewer times under stress (fixations per option: M = 1.20; *SD* = 0.33) than in the neutral condition (M = 1.59; *SD* = 0.63), t(35) = 2.36, *p* = .02. They also looked fewer times at the template under stress (M = 5.70, *SD* = 1.13) compared to the neutral condition (M = 6.84; *SD* = 1.86), t(35) = 2.25, *p* = .03 (Figure 4).

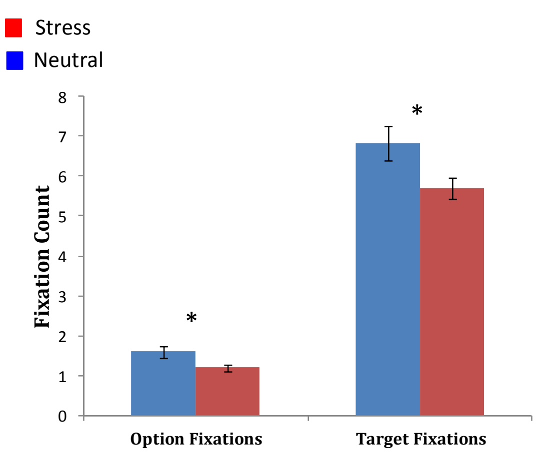

Figure 4. Under stress, fewer fixations were made on each of the four option images and on the template.

Participants were also found to differ in time spent looking at each fixation point (i.e. dwell time) in the stress and neutral conditions: in the stress condition, they spent less time looking at the four options (dwell time per fixation: M = 229.41; *SD* = 63.00), than in the neutral state (M = 315.29; *SD* = 124.82), t(35) = 2.62, *p* = .01); they also spent less time looking at the template (M = 1095.94, *SD* = 227.56) than did participants in the neutral condition (M = 1381.20; *SD* = 424.42), t(35) = 2.53, *p* = .02 (Figure 5).

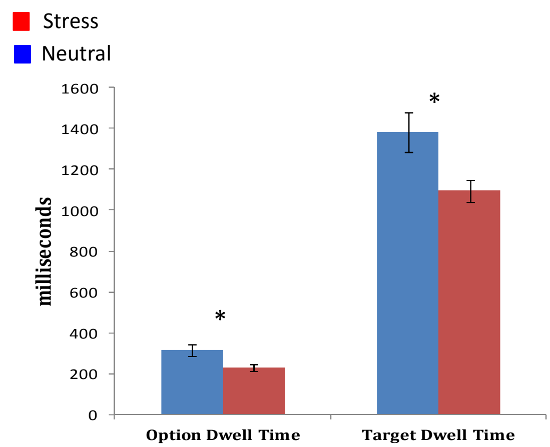

Figure 5. Under stress, less time was spent fixating on each of the four option images and on the template.

Premature closure was examined by extracting the number of trials where the participant responded to a match *before* having directed at least one eye movement to all of the available options. Using an independent samples t-test, participants under stress were found to make significantly more of these sorts of responses (M = 41.72; *SD* = 4.61) than their neutral-task counterparts (M = 37.16; *SD* = 7.76), t(35) = -2.16, *p* = .04. The variable here included both correct and incorrect responses (Figure 3).

To evaluate whether premature closure underlies the increase in errors under stress, this analysis was repeated looking only at trials in which participants chose incorrectly (i.e. premature closure errors as a proportion of total errors). As 4 participants committed no errors, this analysis was run in a restricted sample of those who made at least 1 incorrect response. A significantly higher proportion, t(30) = -2.75, *p* = .01, of incorrect responses were made without sampling all available choices under stress (M = 0.95; *SD* = 0.12) than under the neutral condition (M = 0.71; *SD* = 0.35).

DISCUSSION

The aim of this study was to examine the effects of acute stress on sub-processes of decision-making related to impulsivity. Participants completed a figure matching task (modified MFFT) under either a neutral condition or following an aversive tone stressor. Pupillometry measures confirmed higher arousal levels during the stress induction task as compared to the matched control, supporting the efficacy of this stress manipulation.

It was hypothesized that stress would lead to greater cognitive impulsivity, in terms of both performance on the MFFT and strategies of information sampling as observed using eye-tracking methodologies. As predicted, stress reduced accuracy of performance on the Matching Familiar Figures Task: under stress, participants made more incorrect responses, as compared to in the neutral condition. Furthermore, participants under stress were more likely than their neutral-state counterparts to arrive at a decision before fully sampling all available points of information at least once, thereby demonstrating a tendency towards premature closure under stress. These findings replicated the observations of Keinan (1987), showing that stress leads to more impulsive decision-making. The novel eye-tracking feature of this study allowed for further characterization of these behavioral effects by revealing that participants under stress also spent less time assessing the information, and that they made fewer number of discrete sampling events of the target points.

The findings of this study suggest that acute stress may promote an increased reliance on more cognitively impulsive decision-making strategies. These findings support the conclusions that acute stress leads to impaired information sampling, therefore adversely affecting accuracy and decision-making. The importance of understanding such phenomenon is clear when considering cases such as the Air France 447 accident, in which 24 separate engine-indicating and crew-alerting system (EICAS) messages annunciated problems, all of which originated from the same underlying cause, pitot icing. In this case, over-warning may have contributed to information saturation of the pilots during the accident, resulting in an inability to adequately evaluate all of the information presented in order to correctly identify the root cause of the problem amidst the barrage of alarms. In such an instance, it is possible that some sort of message prioritization function capable of identifying the common origin of cascading failures could have given the crew a better chance of recognizing and responding appropriately to the situation. With this in mind, it is evident that an enhanced understanding of human decision-making, impulsivity, information sampling, and decision-making under stress could potentially have important implications in the implementation and improvement of a variety of visual displays, information consoles, and warning systems in the cockpit, as well as for the design of similar systems across a variety of high-risk, high-stress operational environments.

REFERENCES

Banilivy, M., & Gilliland, K. (1980). An alternate form of the Matching Familiar Figures Test. *Journal of Personality Assessment*, 44(4).

Beatty, J., and Lucero-Wagoner, B. (2000) The papillary system. In: Cacioppo, J.T., Tassinary, L.G., and Berntson, G.G. (Eds.), *Handbook of psychophysiology* (2nd ed.), pp. 142-162. New York: Cambridge University Press.

Bechara, A., and Damasio, A.R. (2005). The somatic marker hypothesis: a neural theory of economic decision. *Games and Economic Behavior*, 52(2), 336-372.

Block, J., Block, J.H., & Harrington, D.M. (1974) Some misgivings about the Matching Familiar Figures Test as a measure of reflection impulsivity. *Developmental Psychology,* 10, 611-632.

Bradley, M., Miccoli, L., Escrig, M., and Lang, P. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, 45(4), 602-607.

Braver, T.S., Paxton, J.L, Locke, H.S., and Barch, D.M. (2009). Flexible neural mechanisms of cognitive control within human prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 106(18), 7351-7356.

Cairns, E. (1977). The reliability of the Matching Familiar Figures Test. *British Journal of Educational Psychology*, 47(2), 197-198.

Clark, L., Li, R., Wright, C. M., Rome, F., Fairchild, G., Dunn, B. D., & Aitken, M. F. (2012). Risk-avoidant decision making increased by threat of electric shock. *Psychophysiology*, *49*(10), 1436-1443.

Donfrancesco, R., Mugnaini, D., and Dell'Uomo, A. (2005). Cognitive impulsivity in specific learning disabilities. European Child & Adolescent Psychiatry, 14(5), 270-275.

Hébert, S., and Lupien, S.J. (2009). Salivary cortisol levels, subjective stress, and tinnitus intensity in tinnitus sufferers during noise exposure in the laboratory. International Journal of Hygiene and Environmental Health, 212(1), 37-44.

Janis, I.L., & Mann, L. (1977). Emergency decision making: a theoretical analysis of responses to disaster warnings. *Journal of Human Stress*, 3(2), 35-45.

Kagan, J., Rosman, B., Kay, D., Albert, J., and Phillips, W. (1964). Information processing in the child: significance of analytic and reflective attitudes. *Psychological Monographs: General & Applied*, 78 (1, Whole No. 578).

Keinan, G. (1987). Decision making under stress: scanning of alternatives under controllable and uncontrollable threats. *Journal of Personality and Social Psychology*, 52(3), 639-644.

Kojima, H. (1976). Some psychometric problems of the Matching Familiar Figures Test. *Perceptual and Motor Skills*, 43(3, Pt 1), 731-742.

Leith, K.P., and Baumeister, R.F. (1996). Why do bad moods increase self-defeating behavior? Emotion risk taking, and self-regulation. *Journal of Personality and Social Psychology*, 71(6), 1250-12167.

Porcelli, A.J. and Delgado, M.R. (2009). Acute stress modulates risk taking in financial decision making. *Psychological Science*, 20(3), 278-283.

Preston, S.D., Stansfield, R.B., Buchanan, T.W., and Bechara, A. (2007). Effects of anticipatory stress on decision making in a gambling task. *Behavioral Neuroscience*, 121(2), 257-263.

Snieder, H., Boomsma, D. I., Van Doornen, L. P., & De Geus, E. C. (1997). Heritability of respiratory sinus arrhythmia: Dependency on task and respiration rate. *Psychophysiology*, *34*(3), 317-328.

Starcke, K., Wolf, O.T., Markowitsch, H.J., and Brand, M. (2008). Anticipatory stress influences decision making under explicit risk conditions. *Behavioral Neuroscience*, 122(6), 1352-1360.