How Stress Affects Functional Near-Infrared Spectroscopy (fNIRS) Measurements of Mental Workload
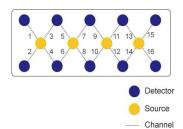

Figure 1: Sensor layout for the Biopac fNIRS used.

Norah H. Alsuraykh
School of Computer Science
University of Nottingham
Nottingham, UK
psxna10@nottingham.ac.uk

Horia A. Maior
School of Computer Science
University of Nottingham
Nottingham, UK
horia.maior@nottingham.ac.uk

Max L. Wilson
School of Computer Science
University of Nottingham
Nottingham, UK
max.wilson@nottingham.ac.uk

Paul Tennent
School of Computer Science
University of Nottingham
Nottingham, UK
paul.tennent@nottingham.ac.uk

Sarah Sharples
Human Factors Research Group
Faculty of Engineering
University of Nottingham
Nottingham, UK
sarah.sharples@nottingham.ac.uk

Abstract

Recent work has demonstrated that functional Near-Infrared Spectroscopy has the potential to measure changes in Mental Workload with increasing ecological validity. It is not clear, however, whether these measurements are affected by anxiety and stress of the workload, where our informal observations see some participants enjoying the workload and succeeding in tasks, while others worry and struggle with the tasks. This research evaluated the effects of stress on fNIRS measurements and performance, using the Montreal Imaging Stress Task to manipulate the experience of stress. While our results largely support this hypothesis, our conclusions were undermined by data from the Rest condition, which indicated that Mental Workload and Stress were often higher than during tasks. We hypothesize that participants were experiencing anxiety in anticipation of subsequent stress tasks. We discuss this hypothesis and present a revised study designed to better control for this result.

Author Keywords
fNIRS; BCI; Stress; Mental Workload; SSSQ; MIST; Distress; Engagement; Worry; Anxiety

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous;

The Montreal Imaging Stress Task: MIST is used to induce a stress response. It has been proficiently applied in various experiments that demonstrate its effectiveness. MIST contains different levels of mental arithmetic challenges that participants must complete. Initially these are completed without any stressors in place. The Stress condition has the same levels of arithmetic difficulty, but places participants under time pressure alongside negative social-comparison evaluations. MIST runs on an automated schedule, putting participants in rest, workload, and stress conditions at given times.

Introduction

The evaluation of Mental Workload (MWL) for measuring human performance has been a significant aim for Human Computer Interaction (HCI) community for decades. MWL has been defined as *"the perceived relationship between the amount of mental processing capability or resources and the amount required by the task"* [3]. MWL, therefore, is experienced in relation to the demand of the task, and the capacity that they have to achieve it, which may be mitigated by other tasks, experience, or physiological states like alertness. MWL has traditionally been measured with subjective scales [3, 18] or indirectly with other physiological changes [6]. In comparison to other brain sensing techniques, recent work in HCI has shown that MWL can be directly and reliably evaluated in conditions of higher ecological validity using Functional Near Infrared Spectroscopy (fNIRS), which measures blood oxygenation changes in the brain [10, 16].

In our prior work, we have anecdotally observed the difference between participants who enjoy a task, and those that are worried or stressed about achieving them [11], where stressed participants tended to perform worse. It has been suggested that Selye coined the term 'stress' in 1936 [15]. There are, however, different forms of stress. Much research focuses on stress as a long-term chronic disorder [5]. This work, however, is focused on more immediate experiences of stress in relation to a task. This kind of stress is typically captured in the conceptual models of how MWL relates to task demand, and is also loosely captured into often-used subjective measures of Mental Workload, where NASA TLX [3] includes stress in the description of the Frustration subscale. Numerous quantitative and qualitative approaches, however, have

been developed to measure Stress, as a concept, more directly, such as the Short Stress State Questionnaire (SSSQ) [4]. We do not know, however, whether stress is measured by fNIRS, nor whether experienced stress confounds our fNIRS measurements of MWL.

This work aims to examine how stress affects fNIRS measurements of Mental Workload in the pre-frontal cortex. If stress does affect fNIRS measurements of MWL, then a) it may have confounded prior studies, b) stress must be controlled in future work, and c) future work could potentially measure both using fNIRS. Below we present the results of our first study. Against our expectations, however, MWL and Stress measurements were often highest during the Rest condition prior to the main study conditions. We concluded that this was because participants had anxiety about the stress-based tasks they were expecting to experience next. We conclude by discussing a revised study protocol for ongoing work, that better controls for these unexpected findings.

Experiment Design

The major aim of our initial study was to evaluate the impact of stress on fNIRS measurements of MWL. To do this, we manipulated Stress using the Montreal Imaging Stress Task [1], to create three conditions: Rest, Control (MWL only), and Experimental (MWL+Stress). While we expected to see objective and subjective differences in MWL and Stress, our primary hypothesis was: There will be a difference in fNIRS data between MWL and MWL+Stress conditions. If rejected, then fNIRS can be confidently used without controlling for experience of stress. If accepted, however, then Stress must be controlled or measured when using fNIRS to measure mental workload.
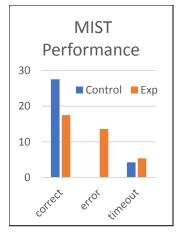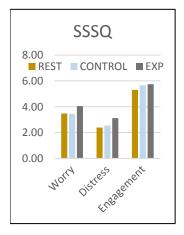
Figure 2: MIST Performance

Figure 3: SSSQ Score

Participants and Protocol

A total of 20 participants (11 females and 9 males) ranging in age from 20 to 34, with a mean age of 26, were recruited. All participants had normal or corrected vision and had no history of brain damage or trauma.

After providing consent, physiological sensors were placed on the participant and configured, and participants began a training period for using MIST. Being automated, participants followed the MIST instructions on the computer screen. After training, participants experienced a 2-minute Rest condition, before experiencing both the Control (MWL) and Experimental (MWL+Stress) conditions. The duration of both task conditions was 4 minutes, and the order was counter-balanced. After each condition, participants completed the SSSQ to assess their stress levels. Participants rested for 2 minutes between conditions to reset physiological baselines. Video recordings were made to assist analysis. After the study was finished, a short debriefing interview was conducted with participants. The study protocol was approved by the school's ethics board, and participants received £10 of Amazon vouchers for remuneration.

Measurement Instruments

Functional Near-Infrared Spectroscopy: fNIRS is a relatively non-invasive imaging method that uses near-infrared light to monitor hemodynamic changes. F.F. Jobsis introduced fNIRS in 1977 and was later used for functional brain imaging [17]. Recent work has shown fNIRS to be more tolerant of natural conditions than other brain sensors [10, 16], and has been used to evaluate usability [8], and even musical performance [19]. In this study, we used the fNIRS300 device and the associated Cognitive Optical Brain Imaging Studio

hardware integrated software platform provided by Biopac Systems Inc. (see Figure 1).

Empatica: Empatica is a wristband of sensors designed to measure heart-rate variability, blood-volume pulse, skin temperature, and electrodermal activity. This physiological data can be used to measure stress responses [2].

Short Stress State Questionnaire: The SSSQ is a self-report assessment that was introduced in 2004, as a shorter version of the Dundee Stress State Questionnaire DSSQ [12]. The SSSQ scale uses 24 questions to assess three primary stress factors (Engagement, Worry, and Distress), and has been used in several studies to effectively measure stress [4].

Results

Overall, while our data showed differences between the Control MWL condition and the Experimental MWL+ Stress condition, the data from our Rest condition was not as would be expected.

Performance: As shown in Figure 2, there was a clear difference in MIST performance between the Control and Experimental conditions. Participants' performance was the worst in stress condition, which confirms that MIST worked as expected. The t-test showed that participants performed considerably worse in the stress condition. The Timeout in stress condition score was higher than the control condition score.

Subjective Stress: Figure 3 shows the differences between conditions in the SSSQ scales: worry, distress and engagement, and highlights the unexpected rest results. Confirming that MIST worked as expected, the Distress (z =2.637, p < 0.01) and Worry (z =2.373, p < 0.05) scores were significantly different between
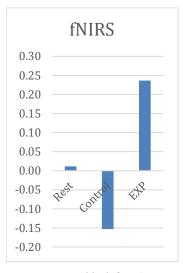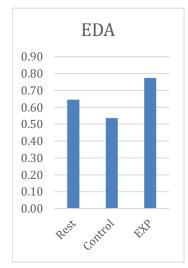
Figure 4: Mean block fNIRS Data.

Figure 5: Mean block EDA Data.

MWL and Stress conditions, whilst engagement was not significantly different. The scores for worry and distress, however, were comparable between the Rest and Control conditions. We expected these, however, to be significantly lower in the rest condition.

fNIRS: After examining data loss, the fNIRS data was analysed for 18 participants. There was no significant difference between conditions for Oxy across the whole pre-frontal cortex ($F_{(2,51)} = 0.400$, $p>0.05$), nor for DeOxy, nor within the left and right regions alone. Noticeable in Figure 4, however, is that the data was higher in Rest than the Control MWL condition, where it should be the lowest in Left and Right of PFC. Again, these results were counter to our expectations, where Rest data should be lower than during high mental workload tasks like arithmetic.

Physiological Signals from Empatica: Unfortunately, after severe data loss, the Empatica recordings were analysed for only the final four participants. With only four participants, the power was not sufficient for finding significant differences in the data. Like the fNIRS data, however, EDA (Figure 5) was the highest in Experimental condition while the lowest in Control condition; again, the rest condition has higher number than control condition. More in line with our expectations, Skin Temperature in the Rest condition was the lowest (32.6°C), whereas Experiment condition has the highest number (34.3°C). Against expectations, however, Heart Rate was highest during Rest (82.9bpm), and lowest in the Experimental condition 80.6bpm).

Discussion
Overall, the meaning of our results was unclear. The Rest condition is typically used to provide a baseline against which high Mental Workload, for example, can be clearly differentiated. While MIST performance data and SSSQ data showed that the Experimental Stress condition was experienced correctly, the fNIRS and Empatica measurements were typically higher during Rest than during the high Mental Workload conditions.

From these results, we produced two interpretations: 1) that our Rest condition was not effectively executed or 2) that participants did, in fact, experience stress or anxiety in anticipation of the subsequent conditions. In our original protocol Rest was always after Training and before the Control and Experimental conditions. If our first interpretation was correct, then we were unable to draw clear conclusions from the results. Our second interpretation, however, was drawn from the SSSQ data for the Rest condition, which had higher levels of Worry and comparable levels of Distress. If correct, our second interpretation would mean that our high fNIRS and Empatica measurements during rest would support the differences found between Control and Experimental conditions. Unfortunately, the study protocol did not have sufficient control over these conditions to draw firm conclusions.

Revised Study Protocol
We concluded that our Rest condition may have been an anticipation anxiety condition. A revised study protocol (Figure 6) was designed to control for this potentially confounding variable. The pre-task Rest condition has been kept, but is considered an Anxiety Rest condition. An additional post-task Rest condition has been added to the end of the protocol, which should not include any anxiety. In addition, a subjective measure of anxiety has been added to the experiment.
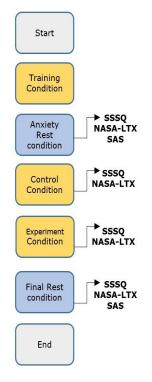
Figure 6: New Experiment Design

Given that the fNIRS data might be affected by Stress, we considered that MWL could be better measured in the revised study. We further added NASA TLX to confirm that Mental Workload was comparable in the Control and Experimental conditions.

Finally, we chose to increase the rigour associated with the two Rest conditions. Rather than asking participants simply to rest, the new protocol will utilize the common approach of asking participants to relax and focus on a small cross on a blank screen. More explicit guidance for the rest condition has also been developed. So far, pilot studies indicate that participants have trouble limiting their self-evaluations to the period of individual conditions, which are less than 5 minutes, rather than to the study participation as a whole.

Anxiety, Stress, and Mental Workload
Research shows there is a relationship between stress and anxiety [9]. Like Stress, much research has focused on defining long-term anxiety, as chronic disorders, and how it could affect mental health and well-being. Instruments have been developed to measure this long-term form of anxiety, such as the Depression Anxiety Stress Scales (DASS) [7]. However, as with stress in our study, we are interested in anticipation imminent forms of anxiety, such as performance anxiety. So far, we have examined the Sport Anxiety Scale Questionnaire (SAS) [13], which is focused on immediate anxiety about imminent performance. This questionnaire could capture participants' anxiety in the rest before the experiment, in comparison to after. This would either a) explain the reasons behind the high score of Rest Condition in the first study, b) exclude anxiety as a confounding variable, or c) confirm that our new protocol manages the rest condition more carefully.

One additional challenge in examining our hypotheses is that Stress, Anxiety, and Workload are each captured in different instruments in different forms [14]. NASA TLX, for example, mentions Stress in the description of its Frustration subscale. Similarly, both SAS and SSSQ have a factor labelled 'Worry' when measuring anxiety and stress, respectively. Evaluating these subjective instruments, in comparison to the recordings, will be conceptually interesting too, and could form the basis of an interesting secondary analysis.

Conclusions

The aim of this research was to identify whether stress has an impact on the measurements of MWL using fNIRS. MIST was used as an established tool for manipulating stress over MWL. Although MIST performance data and our other dependent variables recorded differences in our primary conditions, and support the hypothesis that Stress affects fNIRS measurements of MWL, our data from the Rest Condition were often higher than those during Mental Workload and Stress conditions. We concluded that participants were indeed experiencing stress during the rest period, in anticipation of the subsequent tasks, but the protocol was not able to provide clear evidence for this interpretation. Consequently, a revised protocol is being tested to better control for anxiety, stress, and mental workload across conditions.

References
1. Dedovic, K., Renwick, R., Mahani, N., Engert, V.; Lupien, S.J. & Pruessner, J.C. 2005. The Montreal Imaging Stress Task: using functional imaging to investigate the effects of perceiving and processing psychosocial stress in the human brain. *J Psychiatry Neurosci*, 30, 5: 319-325

2. Gjoreski, M., Luštrek, M., G ams, M., & Gjoreski, H. 2017. Monitoring stress with a wrist device using context. *J. Biomed. Inform.*, *73*: 159-170.

3. Hart, S. G., and Staveland, L. E. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*. 52: 139–183.

4. Helton, W.S., and Näswall, K. 2014 Short Stress State Questionnaire Factor Structure and State Change Assessment. *European Journal of Psychological Assessment*. *31*(1), 20-30.

5. Johnson J. V., and Hall E. M. 1988. Job strain, work place social support, and cardiovascular disease: A cross-sectional study of a random sample of the Swedish working population. *American Journal of Public Health*. 78, 10: 1336–1342.

6. Kramer, A.F., 1991. Physiological metrics of mental workload: A review of recent progress. *Multiple-task performance*, pp.279-328.

7. Lovibond, P.F., and Lovibond, S.H. 1995. The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*. 33, 3: 335-343.

8. Lukanov, K. E., Maior, H. A., & Wilson, M. L. 2016. Using fNIRS in usability testing: understanding the effect of web form layout on mental workload. *In: Proc. CHI'16*. 4011-4016.

9. Maes, M., Song, C., Lin, A., De Jongh, R., Van Gastel, A., Kenis, G., Bosmans, E., De Meester, I., Benoy, I., Neels, H., Demedts, P., Janca, A., charpé, S. & Smith, RS. 1998. The Effects of Psychological Stress on Humans: Increased Production of Pro-Inflammatory Cytokines and Th1-Like Response In Stress-Induced Anxiety. *Cytokine*. 10, 4: 313-318.

10. Maior, H., Pike, M., Sharples, S., & Wilson, M. L. 2015. Examining the Reliability of Using fNIRS in Realistic HCI Settings for Spatial and Verbal Tasks. *In Proc. CHI'15.* 3039-3042.

11. Maior, H. A., Wilson, M. L., and Sharples, S., 2017. Workload Alerts - Using Physiological Measures of Mental Workload to Provide Feedback during Tasks. *ACM Trans. Comput.-Hum. Interact. (in Press)*, 30 pages.

12. Matthews, G., Szalma, J., Panganiban, A., Neubauer, C., and Warm, J. S., 2013. Profiling task stress with the dundee stress state questionnaire. *Psychology of stress: New research*. 49–90.

13. Ronald, S. E., Frank, S. L. & Robert, S. W. 1990. Measurement and correlates of sport-specific cognitive and somatic trait anxiety: The sport anxiety scale. *Anxiety Research.* 263-280.

14. Schvaneveldt, R.W., Gomez, R.L., & Reid, G.B. 1998. Modeling mental workload, *Int. J. Cogn. Technol.* 3, 1: 19–31.

15. Selye, H. 1936. A syndrome produced by diverse nocuous agents. *Nature*. 138 (1), 32

16. Solovey, E. T., Girouard, A., Chauncey, K., Hirshfield, L. M., Sassaroli, A., Zheng, F., Fantini, S., and Jacob, R. J. 2009. Using fNIRS brain sensing in realistic HCI settings: experiments and guidelines. *In Proc. UIST'09.* 157–166.

17. Villringer, A., Planck, J., Hock, C., Schleinkofer, S & Dirnagl. U. 1993. Near infrared spectroscopy (NIRS): a new tool to study hemodynamic changes during activation of brain function in human adults. *Neuroscience letters.* 154, 1: 101–104.

18. Yeh, Y., and Wickens, C. D. 1988. Dissociation of Performance and Subjective Measures of Workload. *Human Factors.* 30, 1: 111–120.

19. Yuksel, B. F., Oleson, K. B., Harrison, L., Peck, E. M., Afergan, D., Chang, R., & Jacob, R. J., 2016. Learn piano with BACh: An adaptive learning interface that adjusts task difficulty based on brain state. In *Proc. CHI'16.* 5372-5384.