UAS Operator Workload Assessment During Search and Surveillance Tasks Through Simulated Fluctuations in Environmental Visibility

Jaime Kerr[1](✉), Pratusha Reddy[1], Shahar Kosti[2], and Kurtulus Izzetoglu[1]

[1] School of Biomedical Engineering, Science and Health Systems, Drexel University, Philadelphia, PA 19104, USA
jkk57@drexel.edu
[2] Simlat, Miamisburg, OH 45342, USA

Abstract. Unmanned aircraft system (UAS) sensor operators must maintain performance while tasked with multiple operations and objectives, yet are often subject to boredom and consequences of the prevalence-effect during area scanning and target identification tasks. Adapting training scenarios to accurately reflect real-world scenarios can help prepare sensor operators for their duty. Furthermore, integration of objective measures of cognitive workload and performance, through evaluation of functional near infrared spectroscopy (fNIRS) as a non-invasive measurement tool for monitor of higher-level cognitive functioning, can allow for quantitative assessment of human performance. This study sought to advance previous work regarding the assessment of cognitive and task performance in UAS sensor operators to evaluate expertise development and responsive changes in mental workload.

Keywords: Unmanned aircraft systems · Sensor operator · fNIRS · Cognitive workload

1 Introduction

Sensor operators are tasked with controlling unmanned aircraft systems (UAS) by maintaining flight systems and controlling visualization equipment onboard, typically in unpredictable and demanding environments regarding reconnaissance missions. Sensor operators face two main issues in the completion of the two aforementioned tasks, including: (1) high amounts of concurrent subtasks to complete, and (2) the nature of UAS ground control station (GCS) system operations removes operators from the physical environment, impairing their ability to assess contextual cues. As a consequence, the information-processing load and decision-making demands are increased for UAS operators, which in turn, impacts mission effectiveness, increases human error and safety concerns [1]. To ideally minimize human error among advancing levels of automation in

D. D. Schmorrow and C. M. Fidopiastis (Eds.): HCII 2019, LNAI 11580, pp. 394–406, 2019.
https://doi.org/10.1007/978-3-030-22419-6_28

UAS, development of operator-specific training protocols can aid in maximizing sensor operators' expertise development and enhance human performance [2].

Utilizing objective performance measures can support UAS operators by facilitating development and reinforcement of personalized training and real-time mental workload monitoring; thus, decreasing error and improving operator performance, with utility for training and in-field applications. UAS sensor operator roles can be arduous yet are often subject to boredom. Previous studies have analyzed lack of focus during detection tasks due to vigilance-related error and object prevalence, otherwise referred to as the prevalence effect, highlighting proportionality among detection miss-rates and target prevalence [3, 4]. However, such studies exhibit scope limited to assessments using subjective measures. Further examination regarding prevalence impacts on detection task performance have focused on subjective feature-dependent performance capabilities, including reported correlation with working memory capacity [5, 6].

Working memory (WM) is a cognitive system that temporarily stores limited information for processing availability with executive control domains; hence, WM is imperative for reasoning, decision-making and attentional control [7–9]. Although facets of theorized functional and mechanistic models of working memory are disputed, attentional control and cognitive load—individuals' capacity for targeting attention while attenuating unnecessary information, and the proportion of time or effort associated with task processing, respectively—are aspects involved in WM that are pertinent to learning [8]. Furthermore, adoption of appropriate cognitive resources taxed via highly realistic practice scenarios can aid in training attentional focus [2], and neurophysiological measures provide methods for monitoring higher-level cognitive functioning for assessing working memory and mental workload. To advance previous studies, we are examining changes in mental workload through neurophysiological data measures in response to the participants' completion of scanning and target identification tasks, using Simlat's C-STAR GCS training simulator [10]. Evidenced from prior studies, we hypothesized that increased level of task difficulty will result in mental workload increases. Task difficulty has been manipulated by administering realistic UAS flight operations under different daylight conditions and mental workload was assessed by behavioral performance measures and functional near-infrared spectroscopy (fNIRS).

1.1 Functional Near Infrared Spectroscopy

Emerging wearable functional brain activity monitoring technologies can help evaluate the cognitive status and capacities of sensor operators in GCS settings. Regarding brain activity monitoring of WM and cognitive load, functional magnetic resonance imaging (fMRI), electroencephalography (EEG) and fNIRS have been widely utilized. Various studies have discussed benefits and challenges of different monitoring technologies involved in measuring cognitive activity, particularly within the prefrontal cortex (PFC) and other associated brain areas correlated with WM assessment efficacy [2, 9, 11–15]. fNIRS and fMRI are frequently subject of such assessments, but fNIRS balances portability with easy-to-engineer design while providing relatively higher temporal resolution compared to fMRI [11]. Additionally, fNIRS is easy to deploy and its shorter calibration time makes it suitable for field settings.

fNIRS utilizes near-infrared light to monitor changes in hemodynamic responses, i.e., oxygenated and de-oxygenated hemoglobin from the PFC area. Changes in blood oxygenation are associated with neuronal activation, as neurons require oxygen to metabolize glucose for activation—oxygen is carried to the site in the form of oxy-hemoglobin (HbO) and converted to deoxyhemoglobin (HbR) once used. HbO and HbR absorb photons at different wavelengths of light, the magnitude of which is captured by source detectors at varying wavelengths within the optical window (i.e., 700 nm to 900 nm) [1, 14]. Relative concentrations of HbO and HbR can be quantified through the application of a modified Beer-Lambert law [1]. Previous studies have analyzed the application of fNIRS as a non-invasive measurement tool to evaluate key aspects of higher-level cognitive functioning, task performance, expertise development and cognitive workload through monitoring hemodynamic response from the prefrontal cortex (PFC) areas collated with behavioral measures [1, 2, 14–16].

2 Methods

2.1 Participants

To date, five participants between the ages of 18 and 42 participated in the IRB approved protocol, including 3 males and 2 females. All participants consented and compiled with inclusion criteria. Participants had no previous experience with flight simulators and verified fulfillment of inclusion and exclusion criteria by completing an Attention Deficiency questionnaire, Edinburgh Handedness survey (LQ > 0 acceptance for right-hand dominance) and video gaming experience survey.

2.2 Experimental Protocol

High-fidelity task representation within sensor operator training can assist in the trainee's adoption of appropriate cognitive styles to perform their role effectively [2]. Simlat's C-STAR GCS (Simlat Inc., Miamisburg, Ohio) simulator is a commercially available training simulation system for UAS pilots and utilized in previous proof of concept studies [1, 2, 16] (Figs. 1 and 2). Simlat's system collects behavioral data regarding the trainee's behavioral task performance by calculating evaluation metrics including scanned, not-scanned, and over-scan percent.

Participants with no previous UAS piloting experience completed three training sessions and two probative sessions, during which, they were required to navigate the UAS sensor over six sub-areas while engaged in route scanning and target identification tasks. For all five sessions the dimensions of each sub-area increased by fifteen percent consecutively along the flight path, and environmental variables for all tests were established from system sunlight and weather data for Mallorca (Majorca), Spain, on September 1, 2018 to administer different levels of task-load conditions. All simulated scenarios utilized a generic tactical unmanned vehicle (G-TAC UAV) system that maintained a fixed route, following the same scan areas and UAV initial state

Fig. 1. Simlat C-STAR instructor screen

Fig. 2. Simlat C-STAR trainee screen

(Altitude: 2000 ft; Speed: 60 kts; Heading: 70°.). Differences among training sessions and probative sessions were established by manipulating the task load via simulated start time; three consecutive training sessions were set for a simulated time lapse from 11:00 to 11:30, while probative sessions one and two were set for simulated time lapses from 06:00 to 06:30 and 20:00 to 20:30, respectively.

Participants were instructed to complete two synchronous tasks: (1) operate the UAS sensor camera to scan along the designated flight path for each consecutive subarea, and (2) to identify and track at least one threat target within each scan area, marked by a red civilian bus. Participants were informed that a zoom angle lower than 15° and three-second target lock were classified as successful scans and target identification, respectively. The duration of each testing scenario was 24.12 min, and subjects were allotted breaks as requested.

Neurophysiological and behavioral performance data were simultaneously captured for elapsed time of participant sessions using a 16-channel fNIRS device and C-STAR integrated Performance Analysis & Evaluation Module (PANEL) logs, respectively. Further analysis was completed to evaluate cognitive performance during task completion, and potential effects of task-load changes in environmental visibility, using observed levels of oxygenated hemoglobin from participants' PFC area. Behavioral task performance was analyzed from system evaluations including scanned, not-scanned, over-scanned percentage and target detected indicator.

2.3 Behavioral Data Processing

The C-STAR PANEL module captured simulator performance metrics for each subject trial and area-specific scan task, including number of scans, field of view (FOV) polygons, region of interest (RoI) scanned percentage, RoI not-scanned, and RoI over-scan percent. Regarding the target identification tasks, the PANEL module captured trial and individual target identification task data, including trial start time, elapsed time, FOV polygon zoom level, target within FOV polygon (TRUE, FALSE) and confirmed target identification (Found = 1, Not Found = 0).

2.4 fNIRS Signal Processing

For the purpose of this study, we utilized a 16-channel fNIRS system with a sampling frequency of 2 Hz., consistent with previous studies [1, 2, 16] (see Fig. 3). fNIRS signals are susceptible to artifacts resulting from instrument noise, physiological noise and motion artifacts. Several signal pre-processing techniques were applied for artifact removal and to improve sensitivity and spatial specificity of brain activity measures [16].

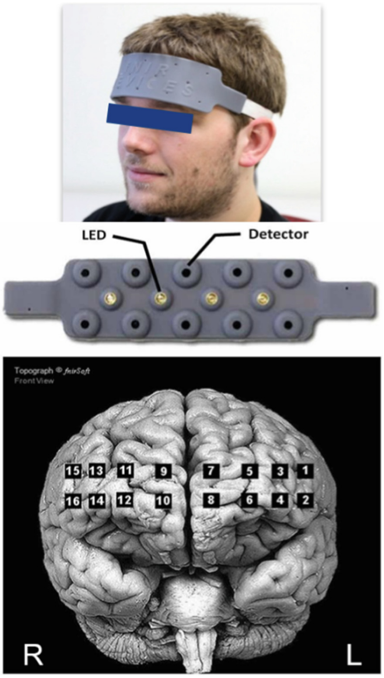

Fig. 3. 16-Channel fNIRS system, node and quadrant location diagram of PFC areas (quadrant 1: 1, 2, 3, 4) (quadrant 2: 5, 6, 7, 8) (quadrant 3: 9, 10, 11, 12) (quadrant 4: 13, 14, 15, 16)

Signal quality analysis and channel rejection were determined from saturation effects and high noise levels. Pre-processing was completed through application of a low pass filter with finite impulse response and cutoff frequency of 0.014 Hz for each channel, to tease instrumental noise, cardiac output, and respiration. To mediate low frequency drift, a linear detrending algorithm was applied. Additionally, due to DC shifts in amplitude for some channels, a temporal derivative distribution repair (TDDR) motion correction method was applied [17]. Following the completion of artifact removal and extraction of processed optical density data, a modified Beer-Lambert Law (MBLL) was applied for calculations of channel-specific changes in HbO and HbR. HbO and HbR measures were utilized for further derivation of channel-specific oxygenation (Hb-diff) and total hemoglobin (Hb_Tot). Finally, sampled outliers were categorized from real-time data as being greater than three standard deviations above the expected values and removed from subsequent analyses.

3 Results

Behavioral performance changes were most clearly determined from measurements for over-scan, rather than scanned, given that participants had no immediate indication of the RoI on their map screen. As an example, Fig. 4 shows the standard PANEL scan task evaluation: the RoI (blue) boundary indicates the region where the scanning task is assigned, but not shown to the operator. The polygon (orange) within field-of-view (FOV) shows a scan overstepping the RoI (blue) boundary, which is classified as over-scan. The polygon (green) encapsulated by the RoI is classified as a proper scan. Scan percentage is extrapolated as the intersection of the RoI area (blue) with the union of all FOV polygons during the elapsed time of the scan task. Alternatively, over-scan percentage was computed by subtracting the area of the RoI polygon from the union of all FOV polygons (black outline).

Provided the protocol design, the initial three training sessions were purposed to facilitate participants' sufficient familiarity of using the system under stable conditions regarding visibility. Upon fulfillment of these sessions, the goal was to study the effects of varying the task difficulty on mental workload, with the assumption all participants have no similar task familiarity and training time. Therefore, we focused on behavioral performance and neurophysiological measurement results of probative sessions only. Workload (i.e., probative) session one started under dark lighting conditions for scan-task one, followed by increasingly improved lighting conditions up to scan-task six—hence, we classified this session as decreasing in difficulty level with each consecutive scanning and target identification task. Figure 5 shows changes in participants' over-scan measures for scan-tasks one and six according to workload, noted as hard to easy, respectively. Figure 5 exhibits a general trend for subjects 1, 3, 6 and 11 of decreased performance associated with increasing workload conditions.

According to the hypothesis, workload conditions are positively related to oxygenation levels, meaning that as task difficulty increases, oxygenation levels should increase provided that the subjects stayed on task. fNIRS channels were grouped by quadrants (i.e., Q1 = Channels 1 to 4; Q2 = Channels 3 to 8; Q3 = Channels 9 to 12; Q4 = Channels 13 to 16) and Hb-diff levels were extracted and averaged per

Fig. 4. Simlat PANEL FOV polygon classification (Color figure online)

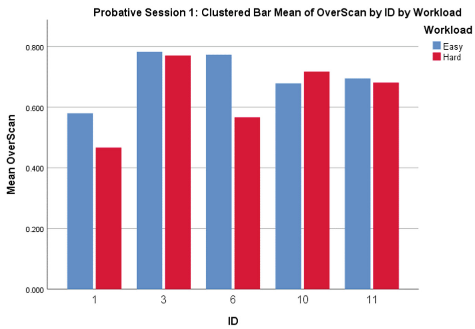

Fig. 5. Probative session 1: clustered means comparison of over-scan percentage sorted from easy workload conditions (blue) to hard (red) (Color figure online)

participant, workload condition and task. Figure 6 presents averages of oxygenation levels for each subject, for Q1 (a) to Q4 (d). Figure 6(a, b) Q1 and Q2 data present trends for subjects 1, 3, 6 and 11 consistent with our hypothesis of decreased oxygenation levels during the lower workload conditions; subject 10 did not present any significant interaction with the workload changes. Figure 6(c) shows Q3 trends among participants 1, 3, 10 and 11 consistent with our hypothesis, but subject 6 did not have apparent trends; whereas, Fig. 6(d) demonstrates that Q4 had hypothesis-consistent trends for subjects 6, 10 and 11, with no similar trend for 1 and higher oxygenation levels for lower workload conditions in 3.

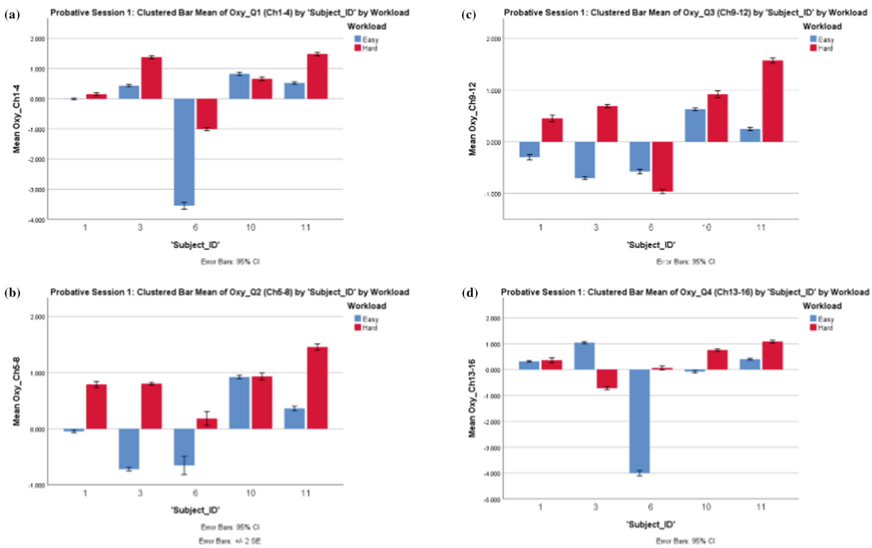

Fig. 6. Probative session 1 fNIRS quadrant Hb-diff Levels comparison between easy (blue) workload conditions and hard (red) for: (a) quadrant 1 (Ch 1, 2, 3, 4), (b) quadrant 2 (Ch 5, 6, 7, 8), (c) quadrant 3 (Ch 9, 10, 11, 12) and (d) quadrant 4 (Ch 13, 14, 15, 16) (Color figure online)

Consequently, for probative session two, participants' over-scan performance was expected to decrease from scan-task one to six, classified from easy to hard, respectively; over-scan performance trends identified in Fig. 7 are consistent with this prediction for subjects 1 and 6. However, improved scan performance was observed for subjects 3, 10 and 11. Figure 8(a–d) demonstrate Q1–Q4 consistency with our hypothesis of oxygenation level dynamics for subject 3. Figure 8(a) presents Q1 trends opposing our hypothesis, with lower levels of oxygenation for higher workload conditions in participants 1, 6, 10 and 11. Figure 8(b, c) show Q2 and Q3 consistency regarding opposing trends in 1, 10 and 11, while (b) presents no trend for 6 and (c) demonstrates an oxygenation level increase for 6. Fig. 8(d) demonstrates Q4 hypothesis consistent trends for subjects 6 and 10, yet opposing trends were observed in subjects 1 and 11.

To further examine the relationship of oxygenation level dynamics with workload conditions, the most difficult scan-task workload condition was compared to the least-difficult workload condition, taken as probative session two scan-task six and probative session one scan-task six, respectively. Difficulty level determinations were obtained from comparisons of visibility levels for probative scan-tasks to that of the training sessions, with the least difficult condition being similar to conditions of training tasks, and the most difficult being the task with the lowest lighting condition. Figure 9 demonstrates the comparison between over-scan measurements from the aforementioned *easiest* and *hardest* tasks. Fig. 9 indicates decreases in scan performance during the higher workload condition for subjects 1 and 6, and increases for 3, 10 and 11. Figure 10(a–c) present Q1–Q3 oxygenation level movements in line with our

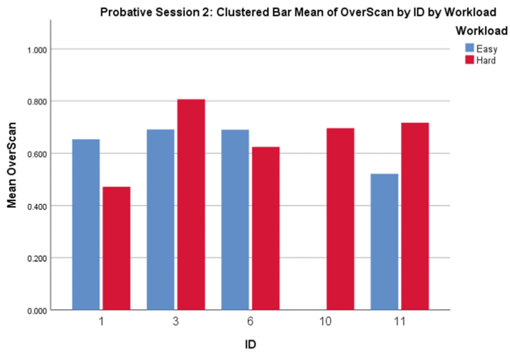

Fig. 7. Probative session 2: clustered means comparison of over-scan percentage sorted from easy workload conditions (blue) to hard (red) (Color figure online)

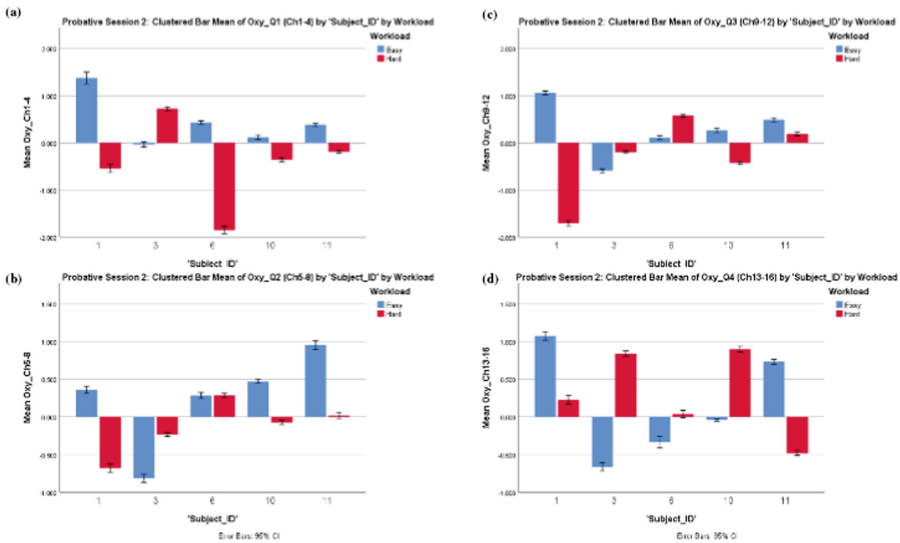

Fig. 8. Probative session 2 fNIRS quadrant Hb-diff Levels comparison between easy (blue) workload conditions and hard (red) for: (a) quadrant 1 (Ch 1, 2, 3, 4), (b) quadrant 2 (Ch 5, 6, 7, 8), (c) quadrant 3 (Ch 9, 10, 11, 12) and (d) quadrant 4 (Ch 13, 14, 15, 16) (Color figure online)

hypothesis for participants 3 and 6, with large disparity in Hb-diff levels between the hardest (red) and easiest(blue) workload conditions in 6, but conflicting Hb-diff level movements for subjects 1, 10 and 11—with significant difference between workload conditions in (b) and (c) for subject 1. Figure 10(d) demonstrates Hb-diff level movement in agreement with our hypothesis for subjects 6 and 10, no similar trend observed for 1, and opposing Hb-diff movement in 3 and 11.

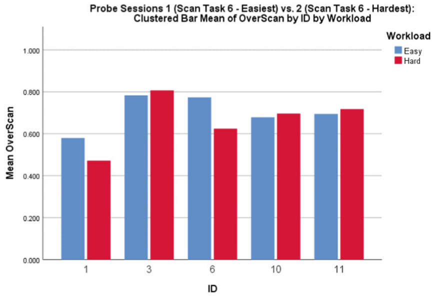

Fig. 9. Probative sessions over-scan comparison according to workload conditions: easiest (blue; probative session 1, scan-task 6), hardest (red; probative session 2, scan-task 6) (Color figure online)

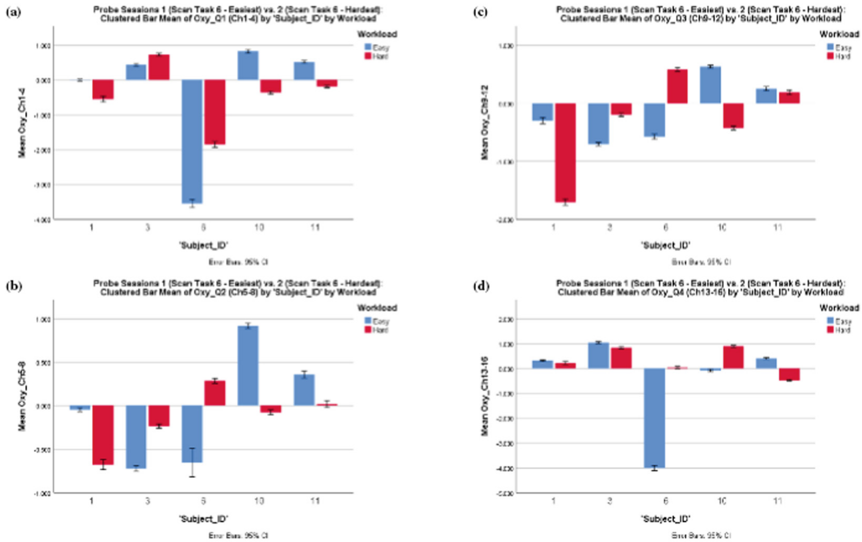

Fig. 10. Probative sessions fNIRS Hb-diff levels comparison between easiest (blue; probative session 1, scan-task 6) workload conditions, and hardest (red; probative session 2, scan-task 6) for: (a) quadrant 1 (Ch 1, 2, 3, 4), (b) quadrant 2 (Ch 5, 6, 7, 8), (c) quadrant 3 (Ch 9, 10, 11, 12) and (d) quadrant 4 (Ch 13, 14, 15, 16) (Color figure online)

4 Discussion

Sensor operators have a crucial role in the completion of mission objectives for UAS flight crew, particularly within the realm of reconnaissance missions. Increased information-processing load and decision-making demands require vast cognitive resource utility; however, such increases in attentional control and cognitive load have been associated with increases in human error, and thus escalating safety concerns.

Proper neurophysiological assessment with realistic training scenarios can facilitate personalized and adaptive training necessary to cope and prepare for such mental strain. Current evaluation criteria for UAS operator training is limited to subjective behavioral and performance evaluations, which may lack in-depth understanding of what factors may contribute to scanning and target-identification task performance. Use of neurophysiological measures for extraction of mental workload-associated parameters can provide objective performance measurements and contextual clues for evaluators to help sensor operators perform tasks more efficiently and effectively.

Since this is an ongoing project and due to limited number of participants, the scope of this manuscript includes only within-subject analysis. Evaluation of results from all subjects require understanding of engagement and training effect. There are four possible trends mainly observed between behavioral and fNIRS measurements in this preliminary study: (1) decreased scan performance from easy to hard workload conditions, with decreased fNIRS Hb-diff measurements; (2) decreased scan performance from easy to hard workload conditions, with increased Hb-diff levels; (3) increased scan performance from easy to hard workload conditions, with an associated decrease in Hb-diff; and (4) increased scan performance from easy to hard workload conditions, with concurrent increases in Hb-diff measures.

Each of these trends can be interpreted with respect to task engagement and training effect. For instance, when a subject follows trend (1) it may indicate lack of engagement for both easy and hard workload conditions. This phenomenon may be understood and validated contextually by the case with requisite movement of sensor camera to complete each scan task successfully, wherein the number of recorded scans increases, there by increasing over-scan percentage. However, if sensor camera movement is limited or absent, over-scan percentage is expected to decrease. fNIRS along with this behavioral measure can introduce a capability potential to detect the engagement, or lack thereof, through quantification of cognitive effort. Subject 1 is a suitable example for lack of engagement (i.e. disengagement) during the experiment, demonstrated by their fNIRS Hb-diff results from probative session two and comparison between easiest and hardest workload conditions (see Figs. 8 and 10, respectively). Subsequently, trend (2) is descriptive of our primary hypothesis, for which subject 6 data effectively demonstrates. That is, increased workload conditions lead to a rise in oxygenation changes acquired from the PFC region. This trend may suggest that if a subject is sufficiently trained on the task—which was achieved through completion of the three initial sessions with optimal visibility—they would likely have the ability to complete such tasks successfully with minimal cognitive effort. If the same trained participant continued active engagement, then administration of more difficult tasks would cause relative increases in cognitive load, revealing that user was able to stay on task and able to keep up with the challenge introduced. Subject data following trend (3) appears akin to trend (2) regarding easier workload conditions; however, improvement in behavioral over-scan percentage concurrent with decrease in brain activity assessed by Hb-diff measures. This preliminary finding suggests that the hardest workload conditions were too high for a given subject, during which (s)he was overloaded and failed to keep up with the task objectives, i.e., camera movement for route scanning and target identification. Finally, subjects outlining trend (4) exhibited proficient task engagement under high-workload conditions. Behavior of trend

(4) subjects are comparable to trend (3) under the easier workload conditions, yet their cognitive effort and behavioral performance rose with increased workload conditions.

The study reported here has ongoing recruitment and testing efforts; more data are being collected for further examination of correlations, trend classification and associated statistical significance. Data collection and analyses of the relationship between task engagement, task performance and correlational analysis with localized oxygenation changes in PFC shall be carried out across subjects in the following study. Although the sample population is limited, this study reported mainly within-subjects results with promising indication of trends and similar results reported in [1, 2, 14, 16].

Acknowledgements. The authors would like to thank Simlat, Inc for providing access to their UAS training simulator.

References

1. Richards, D., Izzetoglu, K., Shelton-Rayner, G.: UAV operator mental workload - a neurophysiological comparison of mental workload and vigilance. In: AIAA Modeling and Simulation Technologies Conference. American Institute of Aeronautics and Astronautics (2017)
2. Armstrong, J., Izzetoglu, K., Richards, D.: Using functional near infrared spectroscopy to assess cognitive performance of UAV sensor operators during route scanning. In: 11th International Joint Conference on Biomedical Engineering Systems and Technologies, vol. 3 (2018)
3. Wolfe, J.M., Horowitz, T.S., Kenner, N.M.: Rare items often missed in visual searches. Nature **435**, 439 (2005)
4. Wolfe, J.M., et al.: Low target prevalence is a stubborn source of errors in visual search tasks. J. Exp. Psychol. Gen. **136**(4), 623–638 (2007)
5. Schwark, J., Sandry, J., Dolgov, I.: Evidence for a positive relationship between working-memory capacity and detection of low-prevalence targets in visual search. Perception **42**(1), 112–114 (2013)
6. Peltier, C., Becker, M.W.: Individual differences predict low prevalence visual search performance. Cogn. Res. Princ. Implic. **2**(1), 5 (2017)
7. Chai, W.J., Abd Hamid, A.I., Abdullah, J.M.: Working memory from the psychological and neurosciences perspectives: a review. Front. Psychol. **9**, 401 (2018)
8. Cowan, N.: Working memory underpins cognitive development, learning, and education. Educ. Psychol. Rev. **26**(2), 197–223 (2014)
9. D'Esposito, M., Postle, B.R.: The cognitive neuroscience of working memory. Annu. Rev. Psychol. **66**(1), 115–142 (2015)
10. Products | Simlat, 04 March 2019. https://www.simlat.com/products
11. Ayaz, H., et al.: Continuous monitoring of brain dynamics with functional near infrared spectroscopy as a tool for neuroergonomic research: empirical examples and a technological development. Front. Hum. Neurosci. **7**, 871 (2013)
12. Berivanlou, N.H., Setarehdan, S.K., Noubari, H.A.: Quantifying mental workload of operators performing n-back working memory task: toward fNIRS based passive BCI system. In: 2016 23rd Iranian Conference on Biomedical Engineering and 2016 1st International Iranian Conference on Biomedical Engineering (ICBME) (2016)

13. Braver, T.S., et al.: A parametric study of prefrontal cortex involvement in human working memory. NeuroImage **5**(1), 49–62 (1997)
14. Hernandez-Meza, G., Slason, L., Ayaz, H., Craven, P., Oden, K., Izzetoglu, K.: Investigation of functional near infrared spectroscopy in evaluation of pilot expertise acquisition. In: Schmorrow, Dylan D., Fidopiastis, Cali M. (eds.) AC 2015. LNCS (LNAI), vol. 9183, pp. 232–243. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-20816-9_23
15. Hill, A.P., Bohil, C.J.: Applications of optical neuroimaging in usability research. Ergon. Des. **24**(2), 4–9 (2016)
16. Reddy, P., Richards, D., Izzetoglu, K.: Cognitive performance assessment of UAS sensor operators via neurophysiological measures. In: 2nd International Neuroergonomics Conference (2018)
17. Fishburn, F.A., et al.: Temporal derivative distribution repair (TDDR): a motion correction method for fNIRS. NeuroImage **184**, 171–179 (2019)