Utilizing Inter-Subject Variability to Assess Performance and Brain Activity During Complex Task Learning

Pratasha Reddy (✉ ylr26@drexel.edu)

Drexel University <https://orcid.org/0000-0001-6320-0729>

Patricia A Shewokis

Drexel University

Kurtulus Izzetoglu

Drexel University

Research

Keywords: Human performance, training, inter-subject variability, functional brain imaging, near infrared spectroscopy, fNIRS.

Posted Date: October 12th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-958867/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Utilizing inter-subject variability to assess performance and brain activity during complex task learning

Pratusha Reddy^{a*} (ylr26@drexel.edu), Patricia A. Shewokis^{a,b,c} (pas38@drexel.edu), and Kurtulus Izzetoglu^{a,c} (ki25@drexel.edu)

^a School of Biomedical Engineering, Science and Health Systems, Drexel University. 3508 Market St Suite 100, Philadelphia, PA 19104 USA.

^b Nutrition Sciences Department - College of Nursing and Health Professions, Drexel University. 1601 Cherry St Free Parkway, Philadelphia, PA 19102 USA.

^c School of Education. 3401 Market Street 3rd Floor Suite 3000, Philadelphia, PA 19104 USA.

* Corresponding author

Abstract

With tasks becoming more mentally focused and operators being required to conduct multiple tasks simultaneously, it is important to not only acquire direct measurements from the brain, but also account for changes in performance and brain activity as a function of intersubject variability and task demands. Such methodology is particularly important when evaluating skill acquisition and transfer during training on a complex and ecologically valid task. To evaluate the aforementioned factors, we implemented a search and surveillance task (scanning an assigned area and identifying targets) using a high-fidelity Unmanned Aerial System operator training simulator, acquired brain activity changes via a portable functional near infrared spectroscopy (fNIRS) sensor array, and had novice participants (N = 13) undergo three sessions of easy difficulty followed by two harder difficulty sessions. Behavioral performance results indicated no significant change in scan or target find performance across easy sessions when intersubject variability was not accounted for. However, accounting for intersubject variability indicated that some individuals improved their scan performance, and they deteriorated their target find performance (Attention-focused group), while others deteriorated their scan performance, and they improved their target find performance (Accuracy-focused group). fNIRS results displayed that both groups exhibited a decrease in brain activity across easy sessions within the left dorsolateral prefrontal cortex (LDLPFC) and right anterior medial PFC (RAMPFC), while activity in left anterior medial prefrontal cortex (LAMPFC) increased in the Attention-focused group and decreased in the Accuracy-focused group. In both groups, transitioning to hard sessions resulted in a decrease in performance. The Attention-focused group displayed an increase in brain activity within LDLPFC, RAMPFC and LAMPFC, while the Accuracy-focused group displayed an increase in brain activity within LDLPFC, no change within RAMPFC and a decrease within LAMPFC. These results suggest that the Attention-focused group was able to acquire and transfer the skills needed to efficiently complete the scan task, while remaining engaged in a target find task. Alternatively, the Accuracy-focused group was engaged only on acquiring the skills needed to efficiently complete the target find task. In conclusion, these results suggest that utilizing intersubject variability as relevant information rather than noise improves assessments of skill acquisition and transfer during training on a complex task.

Keywords: Human performance, training, inter-subject variability, functional brain imaging, near infrared spectroscopy, fNIRS.

42 1. Introduction

43 With recent advances in autonomy capability, we expect human–autonomy systems to be safe and efficient.
44 However, the necessary and expected requirements in safety and efficiency have not yet been met. In fact, reports
45 over the last decade in the fields of transportation, aviation, health care, and ergonomics have indicated human error
46 as the largest contributing factor behind many severe accidents [1]. Increased information processing load and
47 decision-making demands placed on the human operator because of the highly complex systems and task objectives
48 have often been cited as primary reasons behind these accidents.

49 A common solution to ensure that an operator can handle such loads and demands is to have them undergo an
50 effective, theory-based training program (i.e., cognitive load theory), which encourages schema construction and
51 automation, so that the use of working memory and attentional resources while executing day-to-day tasks are not
52 only reduced but freed up to allow for attentiveness to new stimuli [2, 3]. However, due to differences in intrinsic
53 (i.e., age, handedness, etc.), contextual (i.e., prior knowledge regarding the task, stress, etc.) and strategic factors, the
54 process of schema construction and automation is largely variable across individuals [4]. This intrasubject variability
55 can be further exacerbated when an operator is required to learn complex tasks or attend to multiple stimuli
56 simultaneously [5]. Furthermore, current evaluation techniques are primarily dependent on subjective rating
57 methods, performance data using primary or secondary task techniques and physiological sensors. Such
58 methodologies do not enable direct measurement of cognitive resources and real-time monitoring of the mental state
59 of an individual while they are executing a task [6]. Therefore, there is a need to assess skill acquisition and transfer
60 during training on a complex task using methods that measure cognitive resources directly and intrasubject
61 variability as meaningful information rather than as error.

62 Emergence of an interdisciplinary field known as neuroergonomics has made it possible to identify methods
63 needed to objectively assess expertise development. This field is focused on understanding, evaluating, predicting,
64 and improving elements of human performance (such as workload, training, stress, and fatigue) in everyday settings
65 via wearable brain-based technologies such as functional near infrared spectroscopy (fNIRS) [7]. fNIRS is a
66 noninvasive and portable neuroimaging modality capable of continuously measuring correlates of brain activity
67 from the cortex in ecologically valid environments. fNIRS functions under three principles (1) increased neural

activity leads to an increase in metabolic demands, which results in an increase in oxyhemoglobin (HbO) and deoxyhemoglobin (HbR) concentrations; (2) these hemoglobin chromophores have unique optical properties within the 700 to 900 nm wavelength; (3) by examining the manner in which light passes through cortical tissue, concentrations of HbO and HbR can be calculated [8–10].

Over the last decade, fNIRS has been used extensively to assess workload, quantify mental capacity, and track training in both laboratory and field settings [6, 10–16]. Majority of these studies have focused on quantifying changes in brain activity within the prefrontal cortex (PFC), which is responsible for executive functions such as working memory, attention, problem solving, decision making, response inhibition, planning, conflict resolution, mental flexibility, and others [17]. In summary, these studies have indicated three important take-aways: (1) additional task load leads to an increase in brain activity during both standard and complex tasks; (2) these brain activity changes acquired by fNIRS are complementary to behavioral metrics; (3) task practice decreases the extent or intensity of brain activity changes, particularly in the attentional and control areas while maintaining higher behavioral or outcome performance.

In this paper, we evaluated skill acquisition and transfer during training on a complex ecologically valid task using behavioral and fNIRS measures. We selected an Unmanned Aerial System (UAS) operators' search and surveillance task as the complex task since the task requires an active role of attention and spatial working memory to ensure complete scanning of an assigned area and high accuracy of identifying and tracking targets. The complexity or duality of the task allowed for variation in strategy, therefore enabling modelling of intrasubject variability. Based on intrasubject variability as a factor, we hypothesize that performance will vary across individuals and these changes will be elicited in the prefrontal cortex and acquired by brain activation (fNIRS) measures. We had each participant undergo three sessions of easy difficulty, followed by two sessions of harder difficulty to evaluate skill acquisition and transfer, respectively. Based on previous investigation using a similar task, we hypothesize that those participants who demonstrate an increase in their behavioral performance with practice across easy sessions, will display a decrease in brain activity within task specific regions [18, 19]. Additionally, we hypothesize that those who reveal improvement in performance during easy sessions will presumably show transfer of skills during the hard sessions.

2. Materials and Methods

2.1. *Participants*

Thirteen participants between the ages of 19 to 40 (22.92 ± 5.88 years) voluntarily consented to participate in an Institutional Review Board (IRB) approved study. Out of the 13 participants, nine were male and four were female. Recruited participants had no learning disability or sleep disorders, had either normal or corrected to normal vision, and had no prior experience with the simulator used in this study. However, participants had varying levels of overall experience playing three-dimensional (4.27 ± 6.07 hours) and first-player three-dimensional (3.23 ± 4.42 hours) games. Lastly, all participants were assessed as right-handed via use of the Edinburgh Handedness assessment (laterality index: 75.33 ± 20.05 ; and decile: 5.00 ± 3.58) [20].

2.2. *UAS Training Simulator*

A UAS simulator training setup (C-STAR, Simlat Inc., Miamisburg, Ohio) was used in this study, as it allowed for close implementation of a real operator work environment and presented a realistic representation of the daily task. This simulator is currently being used to support over 80 UAS training centers across 30 countries and has been used in previous studies [19]. The simulator allows for a two trainee and one instructor set-up, training on sensor operation and pilot tasks, and for the instructor to manually or automatically preset ‘emergency’ situations that the operator(s) might encounter (e.g., such as developing inclement weather conditions, equipment failure, etc.). In the present study, the simulator was used in a single instructor and trainee configuration (see Figure 1A), where the flight was auto-piloted, sensor operator’s search and surveillance tasks were implemented. The trainee’s simulator’s screen is partitioned into a map and pay-load portion. The map portion displays the flight path and assigned region on the landscape where the tasks need to be executed (blue line and shaded blue region in Figure 1B), provides real-time feedback on the location of the aircraft relative to the flight path, and indicates what portion of the overall map or shaded blue region is under the camera’s field of view (FOV; green polygon in Figure 1B). The sensor operator’s screen displays the actual landscape under the camera’s FOV, and the zoom level associated with the FOV. Additional scopes regarding the engine and the flight are provided below the payload screen, however these functionalities were disabled in this study. The simulator also consists of an embedded Performance Analysis & Evaluation module, which collects, evaluates, and summarizes trainee’s performance. Specifically, the

module records cameras zoom level, a logical index true or false representing when target is in FOV or is not in FOV, respectively and the coordinates of the FOV polygon every microsecond.

Figure 1. Unmanned aerial system training simulator. A. The simulator allows for two trainees and one instructor. B. The trainee screen is divided into map screen on the left and payload or sensor screen on the right. The map screen displays the route that the aircraft will flying along (1), the area where the scan and target find task are assigned (2), and what region the sensor screen is capturing (3). The payload screen displays real time visual of the landscape being looked at with feedback regarding the zoom level (4). This screenshot also shows how the target (5), a red civilian bus, looks like from a distance and when it is being tracked at a zoom angle of 3° .

2.3. *fNIRS Instrumentation*

Hemodynamic changes from PFC were monitored using the fNIR Imager 1200 (fNIR Devices LLC, Potomac, MD) (see Figure 2A). The system operates at a sampling frequency of 10 Hz and measures light intensity during ambient (when no light is shone), 750 and 830 nm wavelengths. The sensor has four surface mount light emitting diodes (red circles in Figure 2B), and twelve silicone photodiodes with integrated trans-impedance preamp (yellow circles in Figure 2B). Ten of the twelve detectors are located 2.5 cm away from each source and enable measurement of cerebral activity from 16 different locations (white circles labelled 1 through 16 in Figure 2B). The remaining two detectors are located 1 cm away from the middle two sources and allow measurement of extracerebral activity from two different locations (white circles labelled 17 and 18 in Figure 2B).

Figure 2. Functional near infrared spectroscopy system used to collect relative changes in oxyhemoglobin and deoxyhemoglobin concentrations. A. The data acquisition and sensor used. The sensor allowed for measurement of cerebral and extracerebral activities using long (2.5 cm) and short (1 cm) source detector separation (SDS) channels. B. The location of sources or light emitting diodes (red circles), detectors or photodiodes (yellow circles), and channels (white circles) on the prefrontal cortex. Channels 1 through 16 represent cerebral measurements, while channels 17 and 18.

2.4. *Experimental Protocol*

Each participant underwent a tutorial session, followed by three easy sessions and two hard sessions (see Figure 3). The tutorial session lasted five minutes, during which the participants were shown how to utilize the joystick and the computer mouse to navigate across map and payload screens, how to lock and track a target, and what constituted as proper scanning and target find behaviors, which were defined as completely scanning the assigned region and tracking the target (red civilian bus as shown on the screen in Figure 1B) at or below a zoom level of 15 for at least 3 seconds. After instructions were given, participants were allowed to practice utilizing the screens and equipment to execute the tasks for the remainder of the tutorial session. The easy and hard sessions were approximately 12.5 minutes in duration, and all had unique flight paths with inimitable target placements. A 15-minute break was given between the easy and hard sessions, during which time the fNIRS was taken off. The primary difference between easy and hard sessions was that the easy sessions occurred at a simulation time of 11:00AM, while the hard sessions occurred at 8:00PM and 6:00AM. To make sure there was no bias, easy, and hard scenarios were randomized during their specific time periods. Each session consisted of six subareas, which each lasted 2 minutes on average. A 10 second gap was programmed between each subarea, allowing participants to re-adjust their camera settings.

Figure 3. Experimental protocol began with a tutorial session followed by easy and hard sessions, respectively. Easy sessions consisted of three similar scenarios that occurred at 11:00 AM (simulator time) and were randomly administered. Hard sessions consisted of two different scenarios that occurred at 8:00 PM or 6:00 AM and were also randomly administered. Each scenario was approximately 12 minutes long and consisted of six subareas that each lasted 2 minutes. Within each of these subareas' participants were required to scan the assigned area and find a target.

2.5. Preprocessing of fNIRS Data

fNIRS signals are often confounded by factors such as motion artifacts, head movement, systemic physiological changes, instrumentation, and environment noises. To extract neural-activity related signals, the following methods were applied. Channels that were saturated (> 4500), had high dark current values (> 200), or had high correlations between wavelength and ambient measurements ($r > 0.7$) were removed from further analysis [21, 22]. Abrupt spikes were removed via wavelet-based motion artifact removal [23]. Low frequency drifts and high frequency noise associated with respiration and cardiac functions were removed via high- and low-pass finite impulse response filters with cut-off frequencies at 0.005 and 0.1 Hz [24]. Optical density data was then converted into the relative concentration changes of HbO and HbR using modified Beer Lambert law [9].

2.6. Dependent Variables

Scanned, not scanned, and over scanned measures were calculated using equations 1, 2 and 3, respectively. Within the equations represents FOV polygon and represents assigned ROI area. An example of these measures extracted from a subarea of an individual is shown in Supplemental Figure 1. Accuracy was set as '1' if the target was in FOV during a particular scan, and if scan occurred at a zoom level at or below 15. In subareas that did not have targets, accuracy was set to '1'. An adaptive target find score was calculated by dividing target find score by the number of the subarea.

$$\text{Scan} = (U_i P_i) \cap A \quad (1)$$

$$\text{Not scan} = A - S \quad (2)$$

$$\overline{\text{Over scan}} = (\sum_i P_i) - A \quad (3)$$

In alignment with the approach previously reported by Izzetoglu et. al., average HbO and HbR measures between 15 seconds after the start of the sub-area and 15 seconds before the end of the sub-area were extracted from each channel [25]. This was performed as the tasks here followed each other continuously to maintain ecological validity, i.e., no resting periods in between, and to wash out any effect from the previous task- hemodynamic response elicited by the preceding task which could be carried over to the present task.

To simultaneously evaluate behavioral and hemodynamic measures, relative efficiency and relative involvement measures were calculated using equations 4 and 5 [16, 26]. In the equations, P represents standardized performance score (scan, not scan, and over scan), while M represents standardized mental effort score (HbO and HbR). To fit within the efficiency definition set forth by Paas and van Merriënboer, HbR measures were multiplied by -1. Unlike HbO, HbR measures are expected to be negative and increase with practice.

$$\overline{\text{Relative Efficiency}} = P - M/\sqrt{2} \quad (5)$$

$$\overline{\text{Relative Involvement}} = P + M/\sqrt{2} \quad (6)$$

2.7. Statistics

Subjects were classified as either Attention-focused or Accuracy-focused based on whether their scan measures improved or declined from easy sessions 1 to 3. The subject-by-subject classification of individuals into their associated group is shown in Supplemental Figure 2. Due to a hierarchical nesting structure and the presence of missing data (determined to be missing at random), linear mixed effects regression (LMER) modelling was used. Models generated investigated the main and interaction effects of Group (Attention-focused vs Accuracy-focused), Session (Easy 1 – E1, Easy 2 – E2, Easy 3 – E3, and Hard 1 – H1 and Hard 2 – H2) on behavioral (3 measures: Scan, Not scan, and Over scan), mean fNIRS (32 measures: HbO and HbR), relative efficiency (96 measures: e.g., HbO – Scan) and relative involvement (96 measures: e.g., HbO – Scan) from all cerebral channels (1 through 16). Since those who performed well on scan tasks did not show improvements in target find tasks and vice versa, interactions between Group and Adaptive Target Find Score were incorporated as a fixed effect. Equation 7 describes the model investigated for behavioral measures. A random slope term based on short source detector separation measurements ($0 + \text{Short} | \text{ID}$) was added to equation 7 when evaluating mean fNIRS measures. This additional random term allowed for separation of task-induced extracerebral activity from that related to cerebral

activity [27, 28]. When evaluating relative efficiency and relative involvement measures, equation 7 without the “Group : Adaptive Target Find Score” term was used. A random slope factor accounting for extracerebral relative efficiency and relative involvement were added to the model.

$$DV \sim 1 + \text{Group} + \text{Group} : \text{Session} + \text{Group} : \text{Session} : \text{Adaptive Target Find Score} + (1 | \text{ID}) \quad (7)$$

Significance of fixed effect terms were evaluated using likelihood ratio tests, where the full effects model was compared against a model without the effect in question. Maximum likelihood estimation was used to conduct likelihood ratio tests, while restricted maximum likelihood was used to evaluate post hoc comparisons. If an interaction term was significant, then planned comparisons were performed between the same sessions of different groups (Attention-focused vs Accuracy-focused of E1, E2, E3, H1, and H2) and between sessions of the same group (e.g., Attention-focused: E1 vs E3, E3 vs H1, and H1 vs H2). A total of 10 comparisons were conducted per each dependent variable. Homogeneity of variance, and normality of residuals and random effects were conducted using visual inspection. If model predictions showed heteroscedasticity or non-normal distribution, then log10 transformations were performed on the response variables. Satterthwaite approximation of degrees of freedom was used in post hoc analyses (Friston, 2003). For all statistical analyses, the level of significance was set at $\sqrt{\alpha} = 0.05$. Adjustments using false discovery rate (FDR) were made on p-values to account for Type I error inflation per dependent variable. Cohen’s *d* was used to examine post hoc effects (Westfall et al., 2014). *d* of 0.2 is considered a small effect, while 0.5 and 0.8 represent medium and large effects, respectively. All statistical analyses were conducted in R (R Core Team, 2019) using *lme4*, *lmerTest* and *emmeans* functions [29–31].

3. Results

3.1. *Effect of incorporating intersubject variability as a fixed factor while evaluating behavioral performance measures*

Effect of the session was significant for scan ($\chi^2(4) = 12.00, p = 0.017$) and not scan ($\chi^2(4) = 12.33, p = 0.015$) measures when the group was not added to the model. However, post hoc analysis indicated that none of the

comparisons were significant (E1 vs E3; E3 vs H1; H1 vs H2) for both behavioral measures. Comparison of models with the Group term, indicates significant improvement in goodness of fit as shown in Table 1.

Table 1. Comparison of goodness of fit between models with and without group as fixed factor for behavioral measures.

Dependent Variable	Model	Parameters	Comparison against null model			Comparison between models	
			Log Likelihood	χ^2	p value	χ^2	p value
Scan	1 + Session + (1 ID)	7	-526.10	12.00	0.017	33.40	<0.001
	1 + Group + Group : Session + (1 ID)	12	-509.41	45.39	<0.001		
Not Scan	1 + Session + (1 ID)	7	-502.56	12.33	0.015	14.86	0.011
	1 + Group + Group : Session + (1 ID)	12	-495.13	26.32	0.001		
Over Scan	1 + Session + (1 ID)	7	-524.71	1.73	0.786	19.50	0.002
	1 + Group + Group : Session + (1 ID)	12	-514.96	20.09	0.010		

3.2. *Effect of Group, Session and Adaptive Target Find on behavioral performance measures*

Interaction between Group and Session was significant for scan ($\chi^2(8) = 45.39, p < 0.001$), not scan ($\chi^2(8) = 26.32, p = 0.001$) and over scan ($\chi^2(8) = 20.09, p = 0.010$) measures (see Figure. 4). Post hoc testing for interaction between Group and Session, revealed significant differences only for scan measures. Specifically, comparisons between Groups per Session revealed significant differences only in easy session 1 (adj.p = 0.022, $d = -1.07$), where the Accuracy-focused group had superior scanning than the Attention-focused group. Pairwise comparisons between sessions within the Attention-focused group indicated significant increases in scanning from easy session 1 to easy session 3 (adj.p = 0.002, $d = -0.84$). Comparisons within the Accuracy-focused group indicated significant decreases in scanning from easy session 1 to easy session 3 (adj.p = 0.001, $d = 0.92$). Interaction between group and adaptive target find score was significant only for only over scan ($\chi^2(10) = 23.16, p = 0.010$) measures. However, post hoc comparisons revealed no significant differences between Groups per Session and between Sessions per Group.

3.3. Effect of Group, Session and Adaptive Target Find on mean fNIRS measures

Interaction between Group and Session was significant across most channels, with only channels 3, 4, 15 and 16

displaying no significant effects (see Figure. 5A). Largest effects were observed in channel 12 (HbO: $\chi^2(8) = 30.11$, $p < 0.001$; HbR: $\chi^2(8) = 109.29$, $p < 0.001$), followed by channels 2 (HbO: $\chi^2(8) = 54.28$, $p < 0.001$; HbR: $\chi^2(8) = 77.16$, $p < 0.001$), 14 (HbO: $\chi^2(8) = 38.46$, $p < 0.001$; HbR: $\chi^2(8) = 86.43$, $p < 0.001$), 7 (HbO: $\chi^2(8) = 78.83$, $p < 0.001$; HbR: $\chi^2(8) = 43.63$, $p < 0.001$), and 11 (HbO: $\chi^2(8) = 23.50$, $p = 0.003$; HbR: $\chi^2(8) = 68.99$, $p < 0.001$).

Figure 4. Changes in behavioral measures as a function of Session and Adaptive target find score per performance group. Attention-focused performers (N = 6), and Accuracy-focused performers (N = 7). Plotted points reflect mean. Easy 1 – E1, Easy 2 – E2, Easy 3 – E3, and Hard 1 – H1 and Hard 2 – H2.

Interaction between Group and Adaptive target find score was found to be significant in channels 1, 2, 3, 5, 6, 7, 8, 10 and 13, with channel 3 (HbO: $\chi^2(2) = 39.09$, $p < 0.001$) displaying largest effect, followed by channels 6 (HbO: $\chi^2(2) = 36.92$, $p < 0.001$), 1 (HbO: $\chi^2(2) = 29.42$, $p = 0.001$), 2 (HbO: $\chi^2(2) = 20.55$, $p = 0.024$; HbR: $\chi^2(2) = 27.85$, $p = 0.002$) and 10 (HbO: $\chi^2(2) = 27.30$, $p = 0.002$) (see Figure. 5B).

Evaluation of post hoc comparisons between Groups per Session for “Group : Session” term revealed significant activity in channels 5, 7, 9, 12, 13 and 15. No significant differences were found in any channels for easy session 1. In easy session 3, significant differences were observed in channels 7 (HbO: adj.p = 0.004, $d = 1.62$), 9

(HbO: adj.p = 0.007, $d = 1.12$), 12 (HbR: adj.p = 0.013, $d = 1.11$), 13 (HbO: adj.p = 0.002, $d = -1.82$) and 15 (HbR: adj.p = 0.045, $d = 1.20$). Specifically, besides channel 13 Attention-focused performers had greater activity than Accuracy-focused performers. In hard session 1, activity was significant only in channel 12 (HbR: adj.p = 0.036, $d = -0.88$), with activity being dominant in the Accuracy-focused group. Lastly, in hard session 2, activity was significant in channel 5 (HbR: adj.p = 0.006, $d = 1.50$), which was greater in the Attention-focused group.

Evaluation of pairwise comparisons between Sessions per Group for “Group : Session” term indicated significant activity across multiple comparisons within channels 2, 5, 7, 9, 11, 12, 13 and 14 (see Figure 6 and Supplemental Table 1). Specifically, in Attention-focused group, activity from (i) easy session 1 to 3 increased in channel 7, and decreased in channel 2, 11, 12, 13 and 14; (ii) easy session 3 to hard session 1 decreased in channel 7 and increased in channels 11, 12 and 13; (iii) hard session 1 to hard session 2 increased in channels 12 and 14 and decreased in

Figure 5. Interactions effects between Group, Session and Adaptive target find score for each fNIRS measures across all channels. Channels with overlaid coloring indicate significance of the associated term at α less than 0.05.

channel 5. Alternatively, in Accuracy-focused group, activity from (i) easy session 1 to 3 decreased in channels 2, 7, 9, 11, 12, 13 and 14; (ii) easy session 3 to hard session 1 increased in channels 2, 5, 7 and 9, and decreased in channel 12; (iii) hard session 1 to hard session 2 decreased in channels 2 and increased in channels 5 and 12.

Post hoc analysis associated with “Group : Session : AdpTF” indicated significant comparisons within channels 1, 2, 6 and 10. Only channel 1 displayed significant changes in relationship between AdpTF and brain activity across Groups per Session. Specifically, a significant difference was observed during easy session 1 (HbO: adj.p = 0.037, $d = 4.157$), where Accuracy-focused performers displayed a negative relationship, while Attention-focused performers displayed a positive relationship. No comparisons across sessions per Attention-focused were significant.

Alternatively, in Accuracy-focused group significant changes in relationship between AdpTF and brain activity were observed (i) from easy session 1 to easy session 3 in channels 1 (HbO: adj.p < 0.001, $d = -5.18$), 2 (HbR: adj.p = 0.001, $d = 3.97$) and 10 (HbR: adj.p = 0.010, $d = 3.65$), where relationship moved from negative to positive in channels 1 and 2, and from positive to negative in channel 10; (ii) from easy session 3 to hard session 1 in channels 2 (HbR: adj.p = 0.001, $d = -4.61$) and 6 (HbO: adj.p = 0.006, $d = -4.60$), where relationship moved from positive to negative and negative to positive, respectively; (iii) from hard session 1 to hard session 2 in channel 2 (HbR: adj.p = 0.001, $d = -4.61$), where relationship moved from negative to positive.

Figure 6. Post hoc comparisons between Sessions per Group across all channels and fNIRS measures. Comparisons consisted of easy session 1 – easy session 3 (across easy), easy session 3 – hard session 1 (between easy and hard), and hard session 1 and hard session 2 (across hard). Only effects (Cohen's $\bar{d}$) associated with significant ($\bar{\alpha} < 0.05$) comparisons were plotted. Cohen's $\bar{d}$ of 0.2 is considered a small effect, while 0.5 and 0.8 represent medium and large effects, respectively. If Cohen's $\bar{d}$ is negative for HbO and positive for HbR, then this indicates that the activity increased in the second term of the

comparison. For example, in Attention-focused performers, channel 13 displayed higher activity in easy session 1 than easy session 3 (HbO $d = 1.61$; HbR $d = -0.87$), while channel 7 displayed higher activity in easy session 3 than easy session 1 (HbO $d = -0.74$; HbR $d = 0.92$).

3.4. *Effect of Group and Session on relative efficiency and relative involvement measures*

Significant interaction effects of Group and Session were observed on relative efficiency (RE) and relative involvement (RI) measures extracted from behavioral (scan) and fNIRS measures from channels 2 (RE - HbO:

$\chi^2(1) = 41.61, p < 0.001$; HbR: $\chi^2(1) = 52.08, p < 0.001$; RI - HbO: $\chi^2(1) = 51.65, p < 0.001$; HbR: $\chi^2(1) = 61.36,$

$p < 0.001$), 7 (RE - HbO: $\chi^2(1) = 66.81, p < 0.001$; HbR: $\chi^2(1) = 26.20, p = 0.001$; RI - HbO: $\chi^2(1) = 79.61, p <$

0.001 ; HbR: $\chi^2(1) = 56.57, p < 0.001$), and 12 (RE - HbO: $\chi^2(1) = 30.58, p < 0.001$; HbR: $\chi^2(1) = 80.29, p < 0.001$)

(see. Figure 7). In channel 2, both Groups displayed increased relative efficiency (Attention-focused: adj.p < 0.001 , $d = -1.57$; Accuracy-focused: adj.p < 0.001 , $d = -1.09$) and decreased relative involvement (Attention-focused: adj.p < 0.001 , $d = 1.05$; Accuracy-focused: adj.p < 0.001 , $d = 1.80$). No differences in relative efficiency were observed

transitioning from easy to hard sessions and across hard sessions by either Group, but significant increases in relative involvement by the Accuracy-focused group was observed when transitioning from easy to hard sessions (adj.p = 0.006, $d = -0.83$). In channel 7, Attention-focused performers decreased efficiency (adj.p = 0.001, $d = 1.14$)

and increased relative involvement (adj.p < 0.001 , $d = -1.30$) across easy sessions, while Accuracy-focused performers had no change in efficiency (adj.p = 0.298, $d = -0.33$) and decreased relative involvement (adj.p < 0.001 , $d = -1.14$). Attention-focused performers increased in relative efficiency (adj.p = 0.001, $d = -1.00$) and decreased in

level of relative involvement (adj.p = 0.006, $d = 0.81$) while transitioning from easy to hard sessions. In contrast, Accuracy-focused performers decreased in relative efficiency (adj.p = 0.001, $d = 0.94$), while remaining highly involved (adj.p < 0.001 , $d = -1.38$). No differences in relative efficiency were seen for neither Attention-focused nor

Accuracy-focused performers across hard sessions. In channel 12, Attention-focused performers displayed increased relative efficiency (adj.p < 0.001 , $d = -1.48$) and decreased relative involvement (adj.p < 0.001 , $d = 1.11$) across easy sessions, while Accuracy-focused performers had no change in relative efficiency (adj.p = 0.251, $d = -0.31$) and

decreased relative involvement ($\text{adj.p} = 0.001$, $d = 1.10$). Attention-focused performers decreased in relative efficiency ($\text{adj.p} = 0.001$, $d = 0.96$) and increased in relative involvement ($\text{adj.p} = 0.007$, $d = -0.78$) from easy to hard session, while Accuracy-focused performers increased in relative efficiency ($\text{adj.p} = 0.001$, $d = -0.94$) and decreased in relative involvement ($\text{adj.p} = 0.016$, $d = 0.63$). No differences in relative efficiency were observed across hard sessions for neither group of performers, but both groups of performers increased their level of relative involvement (Attention-focused: $\text{adj.p} = 0.008$, $d = -0.73$; Accuracy-focused: $\text{adj.p} = 0.020$, $d = -0.58$).

Figure 7. Changes in efficiency and involvement based on scan performance and fNIRS measures from channels 2, 7 and 12 per Group and Session. The four quadrants generated by the efficiency ($E = 0$) and involvement ($I = 0$) lines represent combination of high efficiency (HE) or low efficiency (LE) and high involvement (HI) or low involvement (LI). Circles on the graphs reflect mean, while error bars represent standard deviation.

4. Discussion

With tasks becoming denser for use of mental resources in safety critical domains such as aviation and medicine, there is a need to measure brain activity in conjunction with behavioral performance to evaluate skill acquisition and

transfer during training programs. Furthermore, due to complexity of the tasks being performed and since the effect of training varies across individuals due to intrinsic (e.g., age) or extrinsic factors (e.g., cognitive strategy or prior knowledge), there is a concomitant need to evaluate skill acquisition and transfer as a function of intersubject variability. Therefore, in this study we utilized intersubject variability as the data (i.e., signal of interest) rather than as a confounding factor or noise to evaluate skill acquisition and transfer in novice operators during performance of a complex and realistic task via behavioral and fNIRS measures.

4.1. *Intrasubject variability offers an improved assessment of performance*

Investigators have pressed to include intrasubject variability as a factor while evaluating task effects [4, 32–35]. Like these studies, our results indicate that inclusion of an intra-subject factor - Group, not only improved goodness of fit, but also indicated significant interactions between Group and Session and significant post hoc comparisons.

Intersubject variability can arise due to intrinsic, contextual, or strategic factors [4]. Firstly, intrinsic variability arises from differences in age, gender, handedness, along with other participant-related attributes and experiences. Subjects recruited in this study were all right-handed, were of similar age and were equally distributed within each group, therefore we assume that the variability we observed was not a result of these intrinsic factors. Secondly, contextual variability is driven by familiarity with the task and the environment. Prior to engaging with the task, all individuals indicated that they had limited simulator experience and had no knowledge regarding UAS operator tasks. In addition to being at a similar expertise level, all individuals underwent the same tasks, therefore further supporting the fact that contextual variability may not be the reason behind the observed intersubject variability. According to cognitive load theory, intrinsic and contextual factors interact to make up the “intrinsic load” [2]. Lastly, strategic variability, also termed ‘germane load’ according to cognitive load theory, emerges when an individual can adopt a different strategy that best matches their expectations, prior knowledge, and prior experiences to drive the construction and automation of schemas. Our data strongly indicates the presence of strategic variability [2]. Specifically, evaluation of accuracy related metrics across easy sessions or during skill acquisition displayed that some individuals improved, while others worsened. Furthermore, those who improved in their accuracy metrics showed decreased scan performance across easy sessions. This finding indicates that individuals prioritized different tasks, which is expected as subjects were not informed which task to prioritize. Intersubject variability in behavioral performance due to differences in strategy has been observed in other studies [34, 36–39]. Like our study, other

reports have incorporated intersubject variability as random or fixed factors in mixed effects models and have demonstrated improved model fitness [4, 35].

4.2. *Prefrontal cortex's involvement during UAS operator's search and surveillance task*

Previous studies have repetitively indicated that brain activity during complex tasks is not localized to one PFC area [10, 12, 15, 18, 40, 41]. Furthermore, studies have also shown increased intersubject variability when studying a complex task in comparison to a standard or simple task (i.e., Stroop, etc.) [5]. Our results indicated significant brain activity changes within most channels for both HbO and HbR biomarkers (see Figure 5.A). This global effect is likely due to the nature of the task requiring execution and coordination among multiple cognitive processes. However, post hoc results indicated significant differences across numerous comparisons within channels 2, 7, 11, 12 and 14. These results indicate that even though most of the PFC was recruited to perform the task, that stronger activity was observed in task-relevant areas. Specifically, fNIRS studies have shown that channel 2 is approximately measuring from the left dorsolateral prefrontal cortex (LDLPFC), which is reported to be involved in spatial working memory or recognizing specific features and task setting [17, 24, 32, 42, 43]. Alternatively, channels 11, 12, and 14 are the measures from the right anterior medial PFC (RAMPFC), which is known to be involved with attentional control [17, 24, 32]. Activity within these regions implies that the search and surveillance task employed in this study taxes attention and spatial working memory processes [44]. Furthermore, activations within these channels and regions are in line with other similar fNIRS and fMRI studies evaluating activity during spatial navigation tasks [12, 21, 37, 45, 46]. Lastly, channel 7 is overlayed on top of the left anterior medial PFC (LAMPFC) and has been shown by fMRI studies to be involved with task switching [36, 39]. Activity within this region likely indicates executive control needed to engage in scan and target find tasks simultaneously.

4.3. *Attention-focused group were efficient at acquiring scan skills and remained involved in both tasks, while Accuracy-focused were efficient and involved in only acquiring target-find skills*

Practice is effective in decreasing brain activity intensity within attentional and control areas from an overly active one to one that is nearly automatic [13, 14, 37, 45, 47]. Both Attention-focused and Accuracy-focused groups displayed a decrease in brain activity within RAMPFC and LDLPFC areas. Although these results support neural plasticity and practice theory, the level of neural activation must be interpreted considering behavioral measures [10,

15, 16]. Behavioral results indicated that the Attention-focused group improved in scan performance and declined in target find performance, while Accuracy-focused group decreased in scan performance and increased in target find performance. Combinations of behavioral and brain activity results (relative efficiency and relative involvement measures) indicated that the Attention-focused group prioritized or were more efficient at the scan task, while the Accuracy-focused group prioritized the target find task. However, activity within LAMPFC increased in the Attention-focused group, while it decreased in the Accuracy-focused group. These results indicate that although the Attention-focused group did not demonstrate improvement in target detection, that they were engaged in both tasks. These engagement results are further supported by the results investigating the effect of adaptive target find score on fNIRS measures from LDLPFC, where over scan measures increased with target find during easy session 1 and decreased during easy session 3 (see. Supplemental Figure 3A), while brain activity increased in easy session 1 and 3 when target was found (see. Supplemental Figure 3B). Alternatively, analysis of this task also supports the fact that the Accuracy-focused group was only engaged in the target find task. Specifically, the results showed that the over scan measures decreased with target find during easy session 1 and increased during easy session 3 (see. Supplemental Figure 3A). This shift from negative to positive association, suggests that either the performers stopped scanning after they found the target or were aimlessly wandering. The association between adaptive target find score and fNIRS measures from LDLPFC went from a positive to negative indicating that the subjects went from using more to fewer resources with practice, while finding targets (see. Supplemental Figure 3B). Such associations were not observed in the Attention-focused group, indicating that even though they were task switching they needed additional practice to improve target find performance.

4.4. *Attention-focused group were involved in transferring their scan skills, while the Accuracy-focused group were involved in transferring their target find skills.*

Our behavioral results indicate that introduction to a task of higher load resulted in maintenance of scan performance by both Attention-focused and Accuracy-focused groups. fNIRS results from Attention-focused group displayed increased activity within RAMPFC, decreased activity within LAMPFC, and no change in LDLPFC, while Accuracy-focused group demonstrated decreased activity within RAMPFC and increased activity within LDLPFC and LAMPFC. Although these results indicate that an increase in task load led to recruitment of more neural resources to maintain similar behavioral outcomes as that observed during performance of the easy tasks, the

recruitment of neural resources varied across PFC region and Group. These results indicate the Attention-focused group prioritized the scan task, while Accuracy-focused group prioritized the target find task. However, assessment of relative efficiency and involvement measures provide further insight. In particular, the Attention-focused group decreased in efficiency and remained involved in scan tasks, while ignoring the target find task. These results validate that the Attention-focused group needed more practice on the target-find task during easy conditions before being able to transfer the skills to the hard condition. Additionally, the Attention-focused group displayed a decrease in relative efficiency across hard sessions, while their relative involvement remained high. This further supports the different prioritization strategies used by the two groups. Unlike the Attention-focused group, the Accuracy-focused group was relatively efficient in the scan task, and they were not when performing the target find task. However, they were relatively involved in the target-find task and not relatively involved in the scan task. As previously described the target find task is a secondary task to the primary scan task, which means that improvement in scan task performance should enable improvement in the target find task. Based on this presumed connection, the Accuracy-focused group could be zooming in further to accommodate for the lack of visibility in the hard condition, therefore they may have utilized scan task performance as way of completing the target find task. This could be the reason why the Accuracy-focused group had increased activity in LAMPFC and LDLPFC, but not in RAMPFC. Lastly, the Accuracy-focused group showed no change in relative efficiency or relative involvement across hard conditions. A possible reason for this could be that they quit.

4.5. *Limitations and future directions*

Despite the promising methodology and results, the study results are subject to a few limitations. The study recruited a limited sample size. Specifically, the Attention-focused group had $N = 6$ individuals, while Accuracy-focused group had $N = 7$. Therefore, the findings reported here are preliminary in nature and future studies with a larger sample cohort is needed, especially when investigating intersubject variability. fNIRS signals are influenced by extracerebral and systemic activity. We utilized “(0 + ShortSDS | ID)” to account for global extracerebral and systemic activity. Although the inclusion of this factor significantly improved model fitness (see. Supplemental Table 2), the factor does not remove task-evoked extracerebral and systemic activity [27, 28, 48]. Therefore, future studies will need to incorporate signal processing techniques such as least squares adaptive filters, Kalman filter and state-space model-based methods to improve removal of task-evoked and non-evoked extracerebral and systemic

activities [49]. The brain activity results must be interpreted with caution, as not all areas of the brain that are involved with the UAS operator search and surveillance task could be measured with fNIRS technology employed in this study. Mean fNIRS measures were calculated from fifteen seconds after onset and before end of a subarea, leading to an average over 90 seconds. Studies have indicated that averaging over trials longer than 60 seconds may include unrelated cortical activity and have suggested parsing of the fNIRS time series into small time segments before averaging [50]. Future studies will incorporate analysis of the temporal changes in brain activity while performing the task of interest. Such analyses will allow for examination of whether scan performance remained similar or depreciated after finding a target and in turn validate the results regarding intersubject variability indicated in this study. In addition to task familiarity, contextual variability or intrinsic variability may result from changes in mood states that affect cooperation, motivation, habituation, awareness, and stress [4].

4.6. *Conclusion*

Our study provides a unique insight into intersubject variability through neural and behavioral measures while we analyze human performance in a complex and ecologically valid task. We demonstrated that including intersubject variability as a factor can enhance assessment of skill acquisition and transfer. The study contributes to the existing literature and reports that brain activity changes acquired via fNIRS are sensitive to changes in task demands and that complex task execution elicits recruitment of resources within multiple regions of the PFC. We posit that the changes in cortical activity, particularly within left anterior medial prefrontal cortex region, could be associated with task switching. Our results support previous findings that task practice results in an improvement in behavioral performance metrics and a reduction in the level of brain activity changes. Lastly, our results highlight that integrated behavioral performance and brain activation assessments of relative efficiency and relative involvement are improved metrics for describing skill acquisition and transfer.

Availability of data and materials

The data that supports the findings of this study are available from the corresponding authors, P.R and K.I, upon request.

Competing interests

fNIRS Devices, LLC., manufactures the optical brain imaging instrument which was utilized in this study. K.I. was involved in the technological development and thus offered a minor share in the startup firm, fNIRS Devices, LLC that licensed IP from Drexel University. The remaining authors declare no conflicts of interest.

Funding

The study received no external funding.

Authors' contribution

P.R., P.S. and K.I. developed the study concept and experimental task protocols. Testing and data collection were performed by P.R. P.R performed the data analysis and interpretation under the supervision of P.S and K.I. P.R. drafted the manuscript, and P.S, and K.I helped write the final draft and provided critical revisions. All authors approved the final version of the manuscripts for submission.

Acknowledgements

We graciously thank Shahar Kosti and Simlat, Inc., of Miamisburg, Ohio, USA, for providing access, licensing, and data extraction from the C-STAR simulator, which made this study possible. We also thank Jaime Kerr, for helping in the data collection process.

References

1. Zhu Y, Rodriguez-Paras C, Rhee J, Mehta RK (2020) Methodological Approaches and Recommendations for Functional Near-Infrared Spectroscopy Applications in HF/E Research. *Hum Factors* 62:613–642. <https://doi.org/10.1177/0018720819845275>
2. Sweller J, Van Merriënboer JJG, Paas FGWC (1998) Cognitive Architecture and Instructional Design. *Educ Psychol Rev* 10:251–296. <https://doi.org/10.1023/A:1022193728205>
3. Wickens CD (2008) Multiple Resources and Mental Workload. *Hum Factors J Hum Factors Ergon Soc* 50:. <https://doi.org/10.1518/001872008X288394>
4. Seghier ML, Price CJ (2018) Interpreting and Utilising Intersubject Variability in Brain Function. *Trends Cogn Sci* 22:517–530. <https://doi.org/10.1016/j.tics.2018.03.003>

5. Hedge C, Powell G, Sumner P (2018) The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behav Res Methods* 50:1166. <https://doi.org/10.3758/S13428-017-0935-1>
6. Curtin A, Ayaz H (2018) The Age of Neuroergonomics: Towards Ubiquitous and Continuous Measurement of Brain Function with fNIRS. *Jpn Psychol Res* 60:374–386. <https://doi.org/10.1111/jpr.12227>
7. Parasuraman R (2008) *Neuroergonomics: Brain, Cognition, and Performance at Work*. <https://doi.org/10.1177/0963721411409176>
8. Jobsis FF (1977) Noninvasive, Infrared Monitoring of Cerebral and Myocardial Oxygen Sufficiency and Circulatory Parameters. *Science* (80-) 198:1264–1267
9. Villringer A, Chance B (1997) Non invasive optical spectroscopy and imaging of human brain function. *Trends Neurosci* 20:435–442. [https://doi.org/10.1016/S0166-2236\(97\)01132-6](https://doi.org/10.1016/S0166-2236(97)01132-6)
10. Izzetoglu K, Bunce S, Onaral B, et al (2004) Functional Optical Brain Imaging Using Near-Infrared During Cognitive Tasks. *Int J Human–Computer Interact* 17:211–227. https://doi.org/10.1207/s15327590ijhc1702_6
11. Ayaz H, Onaral B, Izzetoglu K, et al (2013) 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. <https://doi.org/10.3389/fnhum.2013.00871>
12. Galoyan T, Betts K, Abramian H, et al (2021) Examining Mental Workload in a Spatial Navigation Transfer Game via Functional near Infrared Spectroscopy. *Brain Sci* 11:45. <https://doi.org/10.3390/brainsci11010045>
13. Gentili RJ, Shewokis PA, Ayaz H, Contreras-Vidal JL (2013) Functional near-infrared spectroscopy-based correlates of prefrontal cortical dynamics during a cognitive-motor executive adaptation task. *Front Hum Neurosci* 7:. <https://doi.org/10.3389/fnhum.2013.00277>
14. Prakash RS, De Leon AA, Mourany L, et al (2012) Examining neural correlates of skill acquisition in a complex videogame training program. *Front Hum Neurosci* 6:. <https://doi.org/10.3389/fnhum.2012.00115>
15. Shewokis PA, Shariff FU, Liu Y, et al (2017) Acquisition, retention and transfer of simulated laparoscopic tasks using fNIR and a contextual interference paradigm. *Am J Surg* 213:336–345. <https://doi.org/10.1016/j.amjsurg.2016.11.043>
16. Shewokis PA, Ayaz H, Panait L, et al (2015) Brain-in-the-loop learning using fNIR and simulated virtual reality surgical tasks: Hemodynamic and behavioral effects. In: *Lecture Notes in Computer Science*

- (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer Verlag, pp 324–335
17. Cabeza R, Kingstone A (2006) Handbook of Functional Neuroimaging of Cognition, 2nd ed. The MIT Press, Cambridge, Massachusetts
 18. Izzetoglu K, Richards D (2020) Human Performance Assessment: Evaluation of Wearable Sensors for Monitoring Brain Activity. In: Improving Aviation Performance through Applying Engineering Psychology. CRC Press, pp 163–180
 19. Reddy P, Richards D, Izzetoglu K (2019) Evaluation of UAS Operator Training During Search and Surveillance Tasks. In: 20th International Symposium on Aviation Psychology
 20. Oldfield RC (1971) THE ASSESSMENT AND ANALYSIS OF HANDEDNESS: THE EDINBURGH INVENTORY. *Neuropsychologia* 9:97–113
 21. Ayaz H, Shewokis PA, Curtin A, et al (2011) Using MazeSuite and functional near infrared spectroscopy to study learning in spatial navigation. *J Vis Exp* 8:3443. <https://doi.org/10.3791/3443>
 22. Izzetoglu M, Izzetoglu K (2014) REAL TIME ARTIFACT REMOVAL. 1–9
 23. Molavi B, Dumont GA (2012) Wavelet-based motion artifact removal for functional near-infrared spectroscopy. *Physiol Meas* 33:259–270. <https://doi.org/10.1088/0967-3334/33/2/259>
 24. Izzetoglu M, Bunce SC, Izzetoglu K, et al (2007) Functional Brain Imaging Using Near-Infrared Technology. *IEEE Eng Med Biol Mag* 26:38–46
 25. Izzetoglu M, Shewokis PA, Tsai K, et al (2020) Short-term effects of meditation on sustained attention as measured by fNIRS. *Brain Sci* 10:1–16. <https://doi.org/10.3390/brainsci10090608>
 26. Paas F, Tuovinen JE, Tabbers H, Van Gerven PWM (2003) Cognitive load measurement as a means to advance cognitive load theory. *Educ Psychol* 38:63–71. https://doi.org/10.1207/S15326985EP3801_8
 27. Yücel MA, Selb J, Aasted CM, et al (2016) Mayer waves reduce the accuracy of estimated hemodynamic response functions in functional near-infrared spectroscopy. *Biomed Opt Express* 7:3078–3088. <https://doi.org/10.1364/boe.7.003078>
 28. Yücel MA, Selb J, Aasted CM, et al (2015) Short separation regression improves statistical significance and better localizes the hemodynamic response obtained by near-infrared spectroscopy for tasks with differing autonomic responses. *Neurophotonics* 2:035005. <https://doi.org/10.1117/1.nph.2.3.035005>

- 556 29. Bates D, Mächler M, Bolker BM, Walker SC (2015) Fitting linear mixed-effects models using lme4. J Stat
557 Softw 67:1–48. <https://doi.org/10.18637/jss.v067.i01>
- 558 30. Kuznetsova A, Brockhoff PB, Christensen RHB (2017) lmerTest Package: Tests in Linear Mixed Effects
559 Models lmerTest Package: Tests in Linear Mixed Effects Models. J Stat Softw 82:.
560 <https://doi.org/10.18637/jss.v082.i13>
- 561 31. Lenth R (2020) emmeans: Estimated Marginal Means, aka LeastSquares Means
- 562 32. Friedman NP, Miyake A (2017) Unity and Diversity of Executive Functions: Individual Differences as a
563 Window on Cognitive Structure. Cortex 86:186. <https://doi.org/10.1016/J.CORTEX.2016.04.023>
- 564 33. Ganis G, Thompson WL, Kosslyn SM (2005) Understanding the effects of task-specific practice in the
565 brain: Insights from individual-differences analyses. Cogn Affect Behav Neurosci 2005 52 5:235–245.
566 <https://doi.org/10.3758/CABN.5.2.235>
- 567 34. Parasuraman R, Jiang Y (2012) Individual differences in cognition, affect, and performance: Behavioral,
568 neuroimaging, and molecular genetic approaches. Neuroimage 59:70.
569 <https://doi.org/10.1016/J.NEUROIMAGE.2011.04.040>
- 570 35. Thirion B, Pinel P, Mériaux S, et al (2007) Analysis of a large fMRI cohort: Statistical and methodological
571 issues for group analyses. Neuroimage 35:105–120. <https://doi.org/10.1016/J.NEUROIMAGE.2006.11.054>
- 572 36. Armbruster-Genç DJN, Ueltzhöffer K, Fiebach CJ (2016) Brain Signal Variability Differentially Affects
573 Cognitive Flexibility and Cognitive Stability. J Neurosci 36:3978.
574 <https://doi.org/10.1523/JNEUROSCI.2517-14.2016>
- 575 37. Iaria G, Petrides M, Dagher A, et al (2003) Cognitive Strategies Dependent on the Hippocampus and
576 Caudate Nucleus in Human Navigation: Variability and Change with Practice. J Neurosci 23:5945.
577 <https://doi.org/10.1523/JNEUROSCI.23-13-05945.2003>
- 578 38. Speer NK, Jacoby LL, Braver TS (2003) Strategy-dependent changes in memory: Effects on behavior and
579 brain activity. Cogn Affect Behav Neurosci 2003 33 3:155–167. <https://doi.org/10.3758/CABN.3.3.155>
- 580 39. Wager TD, Jonides J, Smith EE, Nichols TE (2005) Toward a taxonomy of attention shifting: Individual
581 differences in fMRI during multiple shift types. Cogn Affect Behav Neurosci 2005 52 5:127–143.
582 <https://doi.org/10.3758/CABN.5.2.127>
- 583 40. Ayaz H, Shewokis PA, Bunce S, et al (2012) Optical brain monitoring for operator training and mental

- workload assessment. *Neuroimage*. <https://doi.org/10.1016/j.neuroimage.2011.06.023>
41. Fraga RP, Reddy P, Kang Z, Izzetoglu K (2020) Multimodal Analysis Using Neuroimaging and Eye Movements to Assess Cognitive Workload. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer, pp 50–63
 42. Curtin A, Ayaz H, Tang Y, et al (2019) Enhancing neural efficiency of cognitive processing speed via training and neurostimulation: An fNIRS and TMS study. *Neuroimage* 198:73–82. <https://doi.org/10.1016/j.neuroimage.2019.05.020>
 43. Shewokis PA, Ayaz H, Curtin A, et al (2013) Brain in the Loop Learning Using Functional Near Infrared Spectroscopy. In: *LNAI*. pp 381–389
 44. Izzetoglu K, Ayaz H, Hing JT, et al (2015) Uav operators workload assessment by optical brain imaging technology (fnir). In: *Handbook of Unmanned Aerial Vehicles*. Springer Netherlands, pp 2475–2500
 45. Little DM, Thulborn KR (2005) Correlations of cortical activation and behavior during the application of newly learned categories. *Cogn Brain Res* 25:33–47. <https://doi.org/10.1016/J.COGBRAINRES.2005.04.015>
 46. Sanfratello L, Caprihan A, Stephen JM, et al (2014) Same task, different strategies: How brain networks can be influenced by memory strategy. *Hum Brain Mapp* 35:5127. <https://doi.org/10.1002/HBM.22538>
 47. Kelly AMC, Garavan H (2005) Human Functional Neuroimaging of Brain Changes Associated with Practice. *Cereb Cortex* 15:1089–1102. <https://doi.org/10.1093/CERCOR/BHI005>
 48. Erdoğan SB, Yücel MA, Akın A (2014) Analysis of task-evoked systemic interference in fNIRS measurements: Insights from fMRI. *Neuroimage* 87:490–504. <https://doi.org/10.1016/j.neuroimage.2013.10.024>
 49. Scholkmann F, Kleiser S, Metz AJ, et al (2014) A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *Neuroimage* 85:6–27. <https://doi.org/10.1016/j.neuroimage.2013.05.004>
 50. Yücel MA, Lüthmann A v., Scholkmann F, et al (2021) Best practices for fNIRS publications. *Neurophotonics* 8:012101. <https://doi.org/10.1117/1.NPH.8.1.012101>

Supplemental Material

Supplemental Figure 1. Example scan and target find performance from a particular Subject, Session and Subarea. A. Raw FOV polygons overlaid on task area. **B.** FOV polygons that had Bottom Max Size less than 750 and FOV Area Ratio less than 0.50 result in scan, not scan, and over scan ratios of 0.57, 0.13, and 0.29, respectively. During this particular subarea the target was not found.

Supplemental Figure 2. Overall and subject-specific changes in normalized scanned performance across training sessions. A. Scanned performance decreased across training session (blue line); however large variability can be seen (black lines). **B.** Six individuals displayed an increasing scan performance, while seven displayed a decreasing behavior. Blue lines represent median of across subjects and subareas, while black lines represent median across subareas.

Supplemental Figure 3. Association between Adaptive target find score and behavioral or fNIRS measures across easy sessions per Group. Dark line represents smoothed conditional mean or regression line, while shaded regions represent confidence interval of 0.95.

Supplemental Table 1. Post Hoc comparisons between Sessions per Group. Cohen's $\bar{d}$ of 0.2 is considered a small effect, while 0.5 and 0.8 represent medium and large effects, respectively. If Cohen's $\bar{d}$ is negative for HbO and positive for HbR, then this indicates that the activity increased in the second term of the comparison. For example, in Attention-focused performers, channel 13 displayed higher activity in easy session 1 than easy session 3, while channel 7 displayed higher activity in easy session 3 than easy session 1.

Channel	Post Hoc Comparison	Attention-focused				Accuracy-focused			
		HbO		HbR		HbO		HbR	
		Adjusted p value	Cohen's $\bar{d}$	Adjusted p value	Cohen's $\bar{d}$	Adjusted p value	Cohen's $\bar{d}$	Adjusted p value	Cohen's $\bar{d}$
2	E1 - E3	0.900	-0.04	0.624	-0.17	0.181	0.49	<0.001	-1.87
	E3 - H1	0.217	0.43	<0.001	-1.71	0.001	1.21	0.014	0.81
	H1 - H2	0.051	-0.65	0.066	0.55	<0.001	-1.65	0.043	0.63
5	E1 - E3	0.655	-0.20	0.892	0.08	0.529	0.31	0.985	0.01
	E3 - H1	0.655	-0.18	0.504	-0.33	0.016	-1.10	0.136	-0.59
	H1 - H2	0.097	0.69	0.004	-1.06	0.105	0.68	0.004	0.97
7	E1 - E3	0.049	0.64	0.362	-0.34	<0.001	1.30	0.001	-1.13
	E3 - H1	0.029	-0.75	0.008	0.92	0.308	0.30	0.495	0.19
	H1 - H2	0.061	-0.55	0.054	-0.68	<0.001	-1.84	0.459	-0.23
9	E1 - E3	0.228	-0.48	0.935	-0.02	0.005	0.85	<0.001	-1.03
	E3 - H1	0.563	0.20	0.543	0.35	<0.001	-1.38	0.792	-0.12
	H1 - H2	0.563	-0.21	0.543	-0.32	0.563	0.14	0.935	0.02
11	E1 - E3	0.004	1.05	0.184	-0.49	0.525	0.21	<0.001	-1.45

	E3 - H1	0.004	-0.96	0.184	0.51	0.261	-0.48	0.732	-0.11
	H1 - H2	0.494	-0.24	0.352	0.27	0.810	0.09	0.352	0.24
12	E1 - E3	0.150	0.61	<0.001	1.03	0.278	0.43	<0.001	-0.98
	E3 - H1	0.856	-0.06	<0.001	-1.60	0.726	-0.21	<0.001	-0.97
	H1 - H2	0.002	-0.93	0.026	0.59	0.856	-0.07	0.002	0.78
13	E1 - E3	<0.001	1.61	0.026	-0.87	0.439	-0.22	0.047	-0.61
	E3 - H1	0.014	-0.91	0.298	0.39	0.204	0.46	0.178	-0.43
	H1 - H2	0.041	-0.72	0.276	0.41	0.439	0.24	0.848	0.06
14	E1 - E3	0.578	0.29	<0.001	-1.54	0.010	0.87	0.008	-0.80
	E3 - H1	0.128	0.61	0.363	0.31	0.838	0.06	0.190	-0.42
	H1 - H2	<0.001	-1.18	0.007	0.79	0.263	-0.43	0.606	-0.16

633

634 **Supplemental Table 2. Effect of systemic activity on fNIRS measures.** Comparing model with 1 + (0+ShortSDS|ID) term

635 against 1+ (1|ID).

Channel	HbO		HbR	
	$\chi^2(4)$	<i>p</i> value	$\chi^2(4)$	<i>p</i> value
1	102.72	<0.001	122.65	<0.001
2	96.10	<0.001	49.56	<0.001
3	183.42	<0.001	66.53	<0.001
4	115.49	<0.001	132.47	<0.001
5	80.35	<0.001	17.45	<0.001
6	46.74	<0.001	63.15	<0.001
7	133.87	<0.001	38.02	<0.001
8	141.76	<0.001	66.31	<0.001
9	207.79	<0.001	49.48	<0.001
10	190.45	<0.001	27.93	<0.001
11	217.47	<0.001	10.22	0.001
12	152.96	<0.001	52.60	<0.001
13	170.28	<0.001	12.56	<0.001
14	185.11	<0.001	70.46	<0.001
15	123.93	<0.001	57.79	<0.001
16	214.49	<0.001	61.79	<0.001

636