Research Article

Development of Non-Invasive Measurements of Cognitive Demand Using Standard Computer Input Devices

Richard Lamb[1*], Surbhi Rathore[2], William Elm[3], Christine Brugh[4], Lori Wachter[4], Stephen Shauger[4], John Light[4] and Elizabeth Richerson[4]

[1]*University of Georgia, College of Veterinary Medicine and College of Pharmacy, Neurocognition Science Laboratory, United States*

[2]*University of Rhode Island, United States*

[3]*Resilient Cognitive Solutions, United States*

[4]*Laboratory for Analytic Sciences, United States*

[*]**Corresponding Author**
Richard Lamb, University of Georgia, College of Veterinary Medicine and College of Pharmacy, Neurocognition Science Laboratory, United States.

Abstract

The objective of this research is to develop and align non-invasive biokinetic measurement tools such as mouse micromovements to currently existing measurement methods to understand underlying analyst cognitive states (such as attention or cognitive demand). Biokinetic measurement is defined as measurements which link movements within an environment to underlying cognitive states. The primary purpose of this project is to examine the alignment of mouse micromovements with neurological signals to detect cognitive demand differences as a means to center cognitive and affective factors and states in the development of analyst support tools. Based upon this purpose this work seeks to answer the following question. To what degree do mouse micromovement profiles provide align with known levels of cognitive demand as measured through neurological means. The authors recruited a total of 33 participants. Collection of cognitive demand data associated with analyst reading, graphical interpretation, and responses were also tracked using a functional near infrared spectrometer (fNIRS). A significant regression was found and mouse micromovements explained approximate 74% of the variance in hemodynamic response. This not only demonstrates a direct predictive link but also highlights a potential intervention point to improve analyst performance.

Keywords: Cognition, fNIRS, Cognitive Augmentation, Measurement

Highlights
- Examine how to use standard input devices to measure cognitive demand.
- Illustrate the measurement of cognitive demand using biokinetics.
- Establish a use case using the intelligence community as an example.
- Suggest applications of biokinetics in information processing.
- Identify the fundamental role of biokinetics and cognitive demand in context.

1. Introduction

The objective of this research work is to develop non-invasive biokinetic measurements of cognitive states using standard computer input tools using micromovements. Specifically, this work will align currently existing neurological measurement methods to mouse micromovements to understand underlying analyst cognitive states (such as cognitive demand and attention). Biokinetic measurement is defined as measurements which link movements within an environment [digital] which reflect underlying cognitive actions of interest [1,2]. For example, computer mouse micromovement tracking and keystroke actions have both been linked to a number of cognitive and emotional

state dynamics to include cognitive demand, attention, cognitive load, anxiety, depression, and others [3,4]. Micromovements are defined as the small involuntary movements of a person that are tied to unconscious responses to the environment [5]. The ability to assess and specifically understand analyst cognitive and affective state provides a potential identifier for intervention points to enhance tradecraft, task completion, and workflow efficiency. Primarily these interventions and enhancements can occur through application of environmental cognitive augmentation using the biokinetic derived cognitive state. Environmental cognitive augmentation is the development of environmental supports improving the efficacy and efficiency of cognition within the analyst work environments [6]. This approach is specifically designed to improve analyst workflow and task completion in human machine teaming contexts. A critical component needed to intervene in states of impasse and to more fully developing human-machine teaming i.e., joint cognitive processing is the ability to measure and quantify an analyst's cognitive state without disrupting the analyst, workflow, and the environment.

By linking cognitive state measurements such as cognitive demand to standard computer input device, such as a mouse micromovements will allow the detection of not just task cognitive demand and task attentional dynamics but provide in the moment i.e., real-time understanding of how the analyst is responding to tasks and workflow [7,8]. Cognitive demand is defined as the amount of demand resulting from task specific features [9]. Cognitive demand plays a critical role in task completion and overall success as analysts process information [10]. By using a standard mouse and keyboard inputs to measure cognitive demand we can leverage tools already present in sensitive working spaces helping to provide measurement without the introduction of new tools, technologies, and disruption of existing analyst tasks and workflows.

Biokinetic measurements of cognitive demand using a mouse and keyboard can not only provide insights into analyst capabilities which may improve through augmentation, but also be used to understand how and when to integrate existing tools and supports, order recommendations, and evaluate summarization quality at the individual user level [11]. The primary purpose of this work is to align mouse micromovements with neurological signals to detect cognitive demand differences. This will allow developers to "center" cognitive and affective factors in cognitive augmentation. Incorporating understandings of cognitive demand into the development of support tools, recommendation ordering, assessment of summarization quality, and in the provision of real-time data for feedback will allow increased competency and performance. A secondary purpose of this work is to understand what cognitive and attentional factors, within specific analyst tasks, impact the levels of completion and end-user product quality when using cognitive demand derived from biokinetic measurement. Based upon these purposes, this work seeks to answer the following question. To what degree do mouse micromovement profiles provide insight and align with known levels of cognitive demand and attentional dynamics as measured through neurological means?

Justification. With the dramatic increase in information available via the internet and through other open and restricted sources, intelligence analysts are continually overwhelmed with the amount of data and information available to them [12]. As information availability continues to expand exponentially, analysts and the larger intelligence community are seeking more effective ways to provide personalized recommendations, personalized summaries, and ultimately personalized cognitive support tools in the form of environmental augmentations. However, there is significant disagreement as to which factors should drive and be used to develop and evaluate quality support tools, cognitive augmentations, recommendations, and summaries [13]. While not all information that analysts work with are textual, this work focuses on text-based tasks as a demonstration and test case for the development of novel measurement approaches using standard computer input devices.

Textual summaries are texts designed to convey specific key information contained in a reference text and are often used to allow analysts to scan large numbers of recommendations [14]. Summary texts are also significantly shorter than the original text and thus poor summaries potentially result in poor information, poor inferences, and potentially analytic failure. In addition to loss of information, development and interpretation of recommendations and summaries are often time consuming and difficult for a person to complete due to the necessary levels of cognitive demand and attentional dynamics associated with these tasks. This is because significant cognitive and attentional resources must be spent when identifying and assessing the necessary aspects of the reference sources and recommendations [15]. In developing summaries and recommendations, human and automated recommendation systems tend to rely on prior history, bias, selection behaviors, and other data driven non-cognitive factors. When little consideration is given to the use of cognitive factors which can influence the use of support tools, augmentations, recommendations, summaries, and the resultant inferences are not well aligned and have reduced utility. Despite the relatively clear identification of the non-cognitive factors that act as a driver, the identification and leveraging of cognitive and affective factors is not as clear [16]. However, there is a consistent and robust body of literature which clearly points to a person's underlying cognitive and affective characteristics as the critical core driver in decisions, intervention points, products, and ultimately analyst performance. The lack of accounting of cognitive and affective factors has resulted in the development of tasks and workflows which are often unclear, time consuming, and potentially unreproducible from an internal cognitive state perspective [17]. As a result of these characteristics (lack of clarity, time, and inability to reproduce) current tools, interventions, and augmentations including recommendations and summaries lack transparency and explainability. This lack of

transparency and explanablity results in loss of trust in the system, the system products, and in more extreme cases analyst resistance to the use of proposed systems [18]. To address these concerns the larger intelligence community and specifically members of the U.S. federal intelligence agencies have sought to develop automated support tools, summarizers, and recommenders toward the goal of cognitive augmentation to improve analyst performance. These agencies have begun to consider how to best account for analysist cognitive and affective states when completing tasks and integrating tools into analyst workflows.

Automatic text summarization is the process of developing a brief and coherent summary of a reference text which retains key elements to include overall reference document meaning without human input. Recommendation in this context is defined as an automated process which suggests specific content to the analyst. When appropriately tuned to meet and analyst's specific cognitive and affective needs the effect is to increase analyst performance e.g., increased processing speed or endurance among other traits [8,19,20]. These outcomes essentially are cognitively augmenting the analyst. However, useful and high-quality automatic recommendations and summarization and the ability to subsequently identify intervention points which support analyst cognition is still extremely difficult to achieve. Despite the difficulties in sufficiently parametrizing human cognitive and affective factors for use in augmenting analyst cognitive processes, there are several promising approaches. The approaches showing promise arise in areas examining the development of cognitive and affective computing specifically using underlying human neurological data and data from the human autonomic nervous system to shape the tools, interventions, and supports in HCI and HMT contexts [21]. However, these promising results have not been specifically applied to work within the intelligence community and with analysts to improve automated tool use and interaction.

Importantly, many of the existing and proposed analyst tools do not have fully developed mechanisms to effectively incorporate evaluation and measurement of key human cognitive and affective factors related to support, intervention, recommendation, and summary. In part this is due to the specialized equipment need to assess cognitive and affective states. In addition, there is resistance to using this type of equipment in sensitive information areas. However, the use of standard computer input devices would allow greater integration into existing intelligence platforms and workflows.

The lack of fully developed processes to effectively measure analyst cognitive and affect state and account for the related factors results in tools, interventions, recommendations, and summaries which are not necessarily aligned with analyst needs, are not fully trusted, explainable, and transparent. This creates a critical gap in the development of automated support tools, interventions, augmentations, recommendations, and

text summaries because to optimize the underlying models it is necessary to access and evaluate the quality of these tools from a cognitive and affective perspective. More importantly, when support tools, recommendations, and summaries are aligned with a person's cognitive states the results are more easily explained, processes appear more transparent, ultimately more trusted, and better support individual analysts [22].

Evaluation of tools, intervention points for augmentation, generated recommendations and summaries is difficult because there is little to no agreement of what needs should be addressed. There are also questions related to what cognitive and affective factors influence analyst behaviors and how they impact outcomes regardless of source modality [23]. Even when humans attempt to evaluate a series of interventions, tools, recommendations, and summaries there is little to no interrater agreement and reliability. Additionally, human raters are often unable to fully evaluate resultant documents and are inaccurate in their evaluation and measurement of preferences, their trust levels, and in their explanations of their selections. This has led to the use of several (ensemble) metrics which lack standardization and do not consider analyst individual differences. This lack of standardization has only exacerbated the problems the team has outlined and has further reduced their [interventions, automatic support tools, recommendations, and summaries] utility to human users [24].

In contrast, neurocognitive data makes use of autonomic nervous system responses. These responses, are not under conscious control and have been found to be more accurate, uniform, and are more easily able to be standardized [25]. Importantly, these [ANS] responses more directly lend themselves to identification of difficulties in understanding, levels of demand, and are able to provide real-time individualized data as an analyst engages in tasks throughout their workflow. Neurocognitive data has also been shown to be sensitive to individual differences related to textual features, modality, task requirements, and the quality of tool design [26].

To address this gap, a tool which is capable of measuring neurocognitive data non-invasively using standard computer input devices is warranted. Specifically, the research team suggests the use of an approach which makes use of neurocognitive data to align mouse micromovements with standardized neurocognitive measures. This allows mouse micromovement profiles to act as a proxy for the standard neurocognitive measures of cognitive demand. Thus, the profiles can ultimately be used to replace the use of neurotechnology to evaluate support tools, interventions, recommendations, and text summaries. These measures can help to understand the role of underlying critical cognitive factors which improve an analyst's ability to complete specific tasks.

1.1. Theoretical Framework
The Brain Microstate Framework (BMF) describes the relationship between the autonomic nervous system response (ANS) i.e., blood

flow or hemodynamic response, due to neuronal tissue metabolic demand related to task or stimuli demands such as examining a summary or recommendation [27,28]. Neuronal tissue when engaged due to task specific demands, such as actions associated with reading or viewing a complex graphic, results in a consistent time-based response of .5 to 3 seconds which can be measured using functional near infrared spectroscopy (fNIRS) [29]. The neural tissue when used in task processing, demands oxygenated blood which is then deoxygenated to support tissue metabolism. The blood movement due to neural tissue as it is demanded is referred to as the hemodynamic response. The combination of the triggering stimulus and the hemodynamic response is known as the stimulus-response complex and is the basis of the standardized measurement used to align the mouse micromovements to cognitive demand. Standardization of the hemodynamic response occurs by calculating the ratio of oxygenated blood to deoxygenated blood. This ratio is then used to quantify the level of demand on the person related to the specific task. The level of task demand measured via hemodynamic response is the quantification of cognitive demand. The consistency of the stimulus-response complex latency is 0.5 to 3 seconds as the neuronal tissue demand changes from no demand to task demand. fNIRS is responsive to changes in blood flow for timeframes within 0.1 seconds and as a result is capable of capturing the shifting changes in hemodynamic response over the span of the task in real-time. Using this framework (Brain Microstate) it is possible to capture and analyze, via fNIRS, an intelligence analyst's fluctuation in cognitive state and the overall demand of the task [30]. It is this demand that is then aligned with mouse micromovement profiles.

1.2. Cognitive Demand Versus Cognitive Load
A key concern when exploring, developing and deploying approaches to human cognitive augmentation is that the choices and actions taken should reflect the most current understandings of the architecture of human cognitive systems. The underlying assumption when considering measurement of cognitive demand is that the interaction of cognition with neurological systems associated with human information processing results in changes in cognitive demand [31]. Cognitive demand is defined as the mental effort needed to respond to and continuously interact with a specific task [32]. Effort is understood as the demand for controlled information processing and the ability to control related behavioral outcomes. Thus, as the features of a task increase in complexity, change modality, or change other features the level of cognitive actions must necessarily increase (change), resulting in increases (changes) in cognitive demand. These changes will manifest neurologically as increases (changes) in hemodynamic response. The measurement of hemodynamic response allows cognitive demand, within the context of the task, to be a more direct measure of analyst cognitive state making the use of cognitive demand well suited for this work [27]. One assumption when contextualizing cognitive demand is that the person from an autonomic nervous

system perspective will seek to minimize the use of metabolic resources as they complete the task in an effort to free metabolic resources for other processing tasks as they arise [33].

Cognitive demand differs from cognitive load; cognitive load is the total amount of working memory resources available to a person which can be used address all of the internal (intrinsic cognitive load) and external (extrinsic cognitive load) working memory demands as needed to acquire and automate schemata in long-term memory (germane load) [34]. While cognitive load provides an overall view of the memory landscape for a person, it is difficult to isolate specific quantified aspects of cognitive load (i.e., intrinsic, extrinsic, and germane) to a task without accounting for the contribution of all aspects of the environment, internal mental states, and the task itself. At the core of this concern, is that the current methodological approaches associated with the measurement of cognitive load may not provide a sufficient level of consistency and are not likely to differentiate the sources of the cognitive load [35]. To this point even when directly asked, participants are often unable to accurately assess not only the level of cognitive load but also the source of the cognitive load making retrospective protocols unreliable [36]. More importantly, even using talk aloud protocols, the ability to engage meta-cognitive process derails the primary cognitive processes we seek to explore. This makes retrospective and talk-aloud approach less reliable and more difficult to implement.

In contrast to the construct of cognitive load, the construct of cognitive demand is specifically tied to the task and the features of the task. In this light cognitive demand is a relationship defined by the hemodynamic response as a function of the stimulus-response complex. This bypasses many of the classification, retrospective analysis, and cognitive interruption concerns associated with cognitive load and its measurement. Cognitive demand specifically isolates the task and neurological response in much the same way an evoke potential is used within EEG studies to tie together stimulus and EEG responses.

1.3. Cognitive Augmentation
Technology tools, design choices, and the arrangement of information in an effort to reduce cognitive demand during information processing is at the core of human cognitive augmentation [37]. The use of tools to improve performance by reducing errors and increasing speed, i.e., increasing the efficacy and efficiency of human cognition is a central function of technologies such as artificial intelligence and machine learning. These tools are intended to helped promote augmentation by facilitating recording, storing, and exchanging information. There are several forms of cognitive augmentation ranging from mechanism targeting neural circuits (internal) to environmental modification (external). Figure 1 illustrates a summarization of some of the available techniques associated with cognitive augmentation rated by invasiveness.
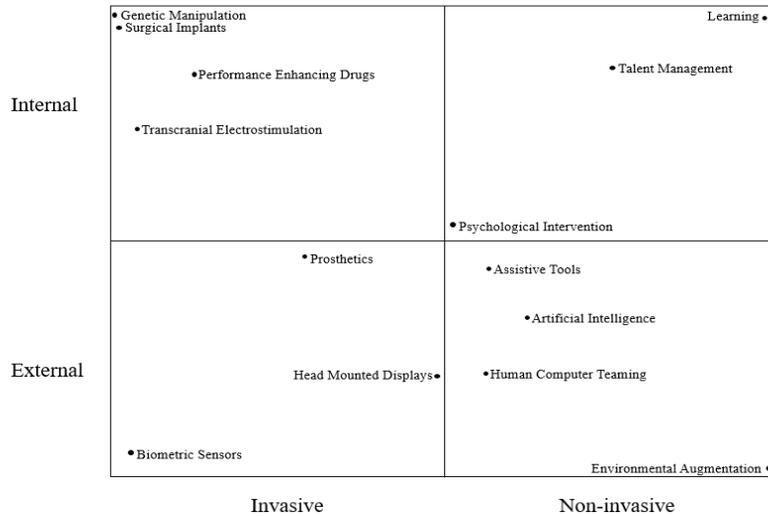
Figure 1: Summary of Cognitive Augmentation Techniques

While all cognitive augmentations seek to improve the efficacy and efficiency of cognition. Approaches differ based upon the targeted function and means of implementation. Implementations for cognitive augmentation can range from modifications which are very invasive and placed inside the body such as surgical cognitive augmentation to augmentations which are outside of the body and modify the environment [6]. This work specifically examines measurement approaches to develop cognitive augmentation via the measurement of mouse micromovement. These micromovements when expressed as a profile, provide a potential driving factor for the quantification of factors which contribute to the development of cognitive augmentation interventions. Environmental cognitive augmentation is the structuring and adaption of a person's environment and tools with specific aim toward improving their abilities to process information or promote other cognitive activities.

2. Methods

The authors of this work seek to identify how mouse micromovements and neurological response align to produce unique profiles associated with high and low demand cognitive states. The authors believe that an understanding of this alignment will allow the establishment of measurements which can be further refined to tune artificial intelligence-based tools to improve analyst productivity through environmental cognitive augmentation. This work uses a mixed block-event design with counter-balanced stimuli [8]. This approach consists of two parts. The first part, is an event related response in which a stimulus-response complex is examined, in this case, the presentation of text, pictures, and other information commonly examined by analysts at varying levels of cognitive demand. The second part of the design is a counter balance approach in which the analyst is randomly assigned to a condition of high or low cognitive demand and differing modality.

This approach is used to ensure ordering and practice effects are reduced or not present. The primary conditions of interest are the manner in which the information is presented i.e., modality and levels of cognitive demand. These areas are of primary interest because of their association with the analyst environment and are less difficult to manipulate when considering how best to alter the environment to improve cognition.

2.1. Participants

The authors recruited 24 participants, 12 male and 12 female, ages 20 through 25 years old. In addition, the team worked with eight participants taking part in a summer conference in applied data science. All participants had a significant background (i.e., undergraduate or graduate degree) in the fields of data analytics, statistics, computer science or other related fields. The eight participants selected from the summer conference were in the past or currently intelligence analysts. The participants in the group of 24 were asked to self-identify thier ethnic / racial backgrounds categories which are taken from the United State Census designations. Ten participants self-identify as White, 10 participants self-identify as Black, 3 participants self-identify as Asian, 1 participant self-identified as Hispanic / Latino. Twenty-one of the participants taking part in the work are from suburban areas, three are from urban areas, and zero were from rural areas. Of the 33 participants the majority were in the process of or did complete a master's degree (75%), 5% were or did complete a terminal degree (Ph.D.), and the remaining 20% were in the process of completing a bachelor's degree. All participants met a priori inclusion criteria i.e., speak, read, write, and understand English, and be within a qualifying technical background which would allow the participant to be considered for an intelligence analyst position. Please note specific demographic and other information from the participants taking part in the summer conference is

unavailable.

2.2. Task Presentation

Participants who met the inclusion requirements were invited to the work location and met by the research team. The participants were asked a series of screening questions to ensure there were no concerns related to language, reading, or writing. Once screening was complete the participants were seated at a computer with a standard monitor, keyboard, and mouse. Twenty-four of the participants were fitted with a whole head fNIRS systems with a sampling rate of 10Hz. Once comfortable with the fNIRS headset, the participant was given a complex text (300-words) with graphical information to read. These texts and graphics were selected as the population was not familiar with them and the content was outside of their expertise areas. The participants were given as much time as needed to read the paragraph. Once the participant completed reading the paragraph and viewed the graphic the participants were asked to complete a series of questions using mouse and keyboard inputs which were tracked on a frame-by-frame basis (120 fps). Participants responded to the varied difficulty questions using the mouse. The eight participants at the summer conference also completed a brief interview after completing the test to discuss the aspects of the tasks which were most demanding. Their interview and mouse micromovement data were used to verify the extracted mouse micromovement profiles from the first 24 participants.

2.3. Instruments and Measurements

Collection of cognitive demand data associated with analyst reading, graphical interpretation, and responses were also tracked using a functional near infrared spectrometer (fNIRS), in addition to mouse tracking. A fNIRS is a non-invasive functional neuroimaging device capable of measuring changes in the oxygenation levels of blood up to 10 mm subcortical within the human brain. A functional near infrared spectrometer consists of two parts. The first component is an infrared emitter, which produces infrared light between 600 nm and 900 nm. The second part of the fNIRS system consists of a detector (known as an optode) which is responsible for receiving the reflected infrared light. Light is initially emitted at 695 nm, oxygenated and deoxygenated blood (hemoglobin) differentially absorb light within this range. Through this differential in absorption, it is possible to estimate the ratio of the concentration between the two forms of hemoglobin. This ratio changes based upon the metabolic demands on the neuronal tissue (hemodynamic response) due to task demands. Thus, the ratio of oxygenated blood to deoxygenated blood is the quantification of the cognitive demand over the span of the task. Larger differences between the volume of oxygenated and deoxygenated blood are specifically due to greater demands on the neural tissue. The greater demand on the neural tissue is due to task difficulty. This increase results in a larger ratio between the oxygenation and deoxygenated blood manifesting due to greater cognitive demand. In this light, tasks which have greater difficulty will illustrate greater hemodynamic response and subsequently cognitive demand.

2.4. Data Analysis

To conduct an analysis of hemodynamic response data, extensive data processing is essential. Initially, it is necessary to remove artifacts caused by gross movements, respiration, and heart pulsation [38]. The research team employed a 0.14 Hz cutoff low-pass filter to achieve this [39]. However, filtering the data with a 0.14 Hz low-band filter resulted in a 7% data loss. Additionally, the team separated the extracranial and extracerebral contributions to the fNIRS signal using regression analysis across each optode, which led to a further 2% data loss.The ratios of oxygenated and deoxygenated blood concentrations were then converted to standardized Z-scores relative to the initial Baseline 1 measurement (i.e., no task present). Using a baseline measurement facilitates comparisons within and between tasks and across different analysts. A moving mean was calculated for each participant based on their hemodynamic response during the task and subsequent inference about the content. This moving mean statistically smoothed short-term hemodynamic response spikes and filtered out signal noise, ensuring that large variations in hemodynamic response did not disproportionately affect the analysis. Finally, a mixed model analysis of variance (MX ANOVA) and post-hoc planned comparisons were conducted using R 4.4.1.

The mixed model analysis of variance (MX ANOVA) enabled the research team to identify statistically significant differences between means across baseline and treatment condition [40]. Specifically, the MX ANOVA was employed to examine measures at each time point: Baseline 1 (A1-condition), Stimulus (B-condition), and Baseline 2 (A2-condition) across conditions. The A-B-A design facilitated the examination of which optodes exhibited greater hemodynamic responses compared to Baseline 1 and 2. Upon identifying the optodes of interest, partial eta-squared was calculated to determine the effects of the condition on the hemodynamics for specific optodes. This process allowed the researchers to pinpoint specific optodes of interest and exclude those that did not contribute to the signal. The MX ANOVA is robust to unbalanced repeated measures, which is necessary due to the hierarchical clustered nature of the data across time. Post-hoc Tukey HSD analysis was conducted to determine which aspects of the task elicited the greatest hemodynamic response. A second post-hoc Tukey HSD analysis was used to identify which optodes exhibited hemodynamic responses greater than each of the baselines (1 and 2) for each task, thereby identifying optodes corresponding to task-related activations.

In addition to hemodynamic response analysis, the authors of this work conducted a time-series regression analysis to establish predictive connection between the mouse micromovements, cognitive dynamics, and analytic outcomes using a Partial Least Squares Procedure (PLS) [41]. Specifically, the regression consisted of examination of mouse micromovements (coordinates) over time as a predictor of hemodynamic responses related to cognitive demand. This allows the team to isolate specific mouse micromovements corresponding to specific hemodynamic

responses. PLS is used to estimate models for variables which have been indirectly observed by multiple indicators such as hemodynamics. The PLS approach makes use of estimates for each construct derived from the observed variables. Iterative model development i.e., examination of rival models, occurs until values for each variable are stable. PLS has been shown to be particularly useful in the development of causal-predictive analysis of model variables. Once multiple models are identified, fit statistics are used to identify the model which best fits the data.
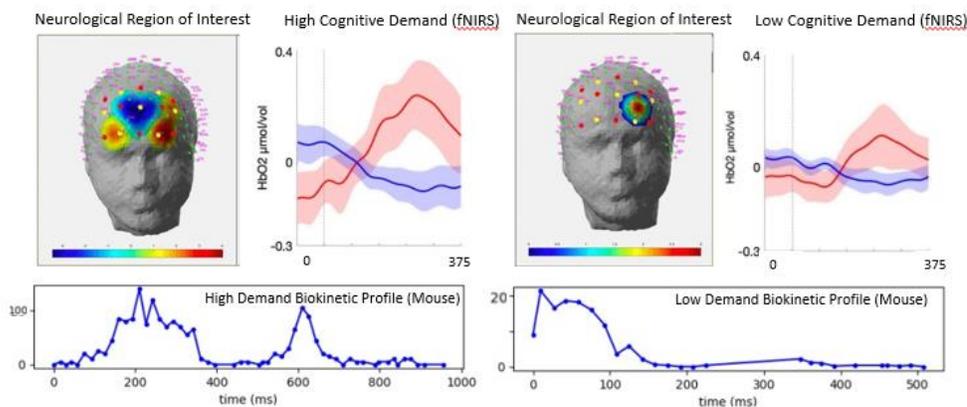
3. Results

To answer the research question: what is the relationship of cognitive demand as measured via hemodynamic response to specific mouse movements; the research team made use of PLS time series regression modeling framework to develop the specific profile corresponding to high and low levels of hemodynamic response. Analysis of rival models shown in Table 1 illustrates that the selected models have the best fit to the data. The selected model shown in Figure 2 illustrates that the predicative relationship between the variables of interest.

Model	AIC	AIC Corrected	BIC
1	921.22	913.78	932.77
2	920.59	907.47	927.44
3	889.37	871.14	899.21
4	874.58	820.19	809.10
Selected Model	867.14	807.54	799.87

Table 1: Rival Model Comparisons

Resultant composite mouse profiles and composite hemodynamic response visualizations are shown in Figure 1. Review of the fit indices (Table 1) illustrates that the selected models have adequate fit to the data. Closer examination of the selected models illustrates that mouse movements are predictive of high and low levels of hemodynamic response. A significant regression was found *(F(2,31)=3.97, p=.029.* The R2 value was .74 indicating that mouse movement explained approximate 74% of the variance in hemodynamic response. That is for every one unit of mouse movement there was a corresponding change of .29 in the hemodynamic response ratio. This is verified by analyst's responses indicating that the mouse micromovements identified when high-level tasks were more demanding. This also held true for the mouse movements corresponding to low level demand on the tasks. Overall, this suggests that the unique profiles extracted from the mouse movements are predictive of levels of cognitive demand.

Figure 2: Results of SCADS 2024 biokinetic profiles resulting from mouse movements in high and low demand analyst tasks when reviewing multiple recommendations and summaries at differing levels of complexity and density (n=8 analysts / n=24 proxies)

4. Discussion

The purpose of this work is to examine the feasibility of extracting mouse micromovement profiles corresponding to levels of cognitive demand. Development of analyst capability and augmentation within the United States has traditionally focused on the development of a collection apparatus and to a lesser degree more rigorous and coherent training approaches without consideration of individual differences in cognition and

affect. Despite this investment there are still significant analytic failures which occur due to misalignment of tools with underlying cognitive and affective factors as analysts work. In order to develop and maintain a competitive advantage, reduce cost, increase individual analysis efficacy and efficiency, attention must be given to approaches which can measure cognitive state. This is particularly true is we are seeking to augment analyst information processing capabilities (i.e., cognition).

The first step toward the development of the capabilities and approaches which augment analyst cognitive action is the development of non-invasive metrics which can assess analyst cognitive and affective states in real-time. The use of mouse micromovements to measure cognitive demand provides an initial step toward this goal. The results of this work support other works illustrating the similar conclusions related to the use of biokinetic and autonomic nervous system data to assess underlying states [42,43]. Specifically, this work focuses on the development of measure to assist in the creation and automatic adaption of environmental cognitive augmentations or analyst support tools.

As we as a nation and an intelligence community collectively move into a more dynamic international environment the environment has become significantly more dangerous, perversive, and elusive [44]. This increased complexity in international dynamics is accompanied by a simultaneous increase in complex unstructured information from which sense making must occur. From the National Intelligence Council's Global Trends 2030, (p. iv, 2012):

"Enabled by communications technologies, power will shift toward multifaceted and amorphous [information] networks that will form to influence state and global actions. Those countries with some of the strongest fundamentals—GDP, population size, etc.—will not be able to punch their weight unless they also learn to operate in networks and coalitions in a multipolar world [processing critical unstructured information]."

This statement by the National Intelligence Council clearly points to the need to radically reconsider the tools, environments, and manner in which an intelligence analyst works and how we understand the factors which contribute to analyst success and failure. It is incumbent upon intelligence agencies to move beyond education and simply developing new tools and to consider how we meet the cognitive, affective, and behavioral needs of the analyst to maximize performance. It is clear that there is significant need for non-invasive neurological measurement, approaches which provide insights into analyst cognitive and affective state. This work illustrates the significant and non-trivial role that non-invasive measures of analyst cognition can play in potentially supporting their ability to produce actionable outcomes and to tune environmental supports using artificial intelligence.

The relationship between cognitive demand and mouse movements within the context of specific analyst tasks is intricately linked to the hemodynamic responses observed in analysts and provides evidence of a leveraging point for adaptive technologies. The hemodynamic response, which reflects changes in blood flow within the brain, serves as a measurable indicator of cognitive activity and is a reasonable metric for efficacy and efficacy of adaption for analyst performance. However, the collection of this neurological data using standard nuerotechnologies is often time consuming, invasive, and not well suited for use in sensitive environments. Thus, using computer mouse data which results from a standard and ubiquitous device (mouse or keyboard) in sensitive environments enable the collection of additional data for classification, scoring, and feedback in a wide range of applications to include recommendations, summaries, and other activities across the analyst workflow.

The results presented in Figure 1 suggests there is a predictive relationship between mouse movements and cognitive demand. Specifically, it indicates that a per unit increase in cognitive control, as measured by hemodynamic response, is associated with a corresponding increase in cognitive demand. To quantify this relationship, the model shows that for every mouse movement during a specific analyst task there is a .29 unit increase in cognitive demand related to hemodynamic response. This not only demonstrates a direct predictive link but also highlights a potential intervention point to improve analyst performance. Using this measurement approach, it becomes possible to understand at which points the level of cognitive demand during a task and within the workflow an analyst loses productive cognitive action or attention. This can allow the analyst to reconsider conclusions and actions they may have taken during these times. The quantification of cognitive demand also allows classification though applications of machine learning which can then adjust aspects of the presented information to assist the analyst in processing texts, graphics, or other sources.

For example, using mouse micromovements it becomes possible to use machine learning classifications to identify neurological changes in demand and control which can then be used to adjust textual features in real-time as analysts engage with content. For example, by adjusting the textual characteristics of information presented to analysts, it may be possible to reduce or change levels of cognitive demand, cognitive control, and, consequently, reduce instances of analytic failure due to erroneous inferencing. Importantly, all of this can be done in real-time as the analyst works with their sources. Understanding how cognitive systems behave in response to analytical tasks is crucial for enhancing the efficiency and accuracy of the analytic process and creating the conditions for successful environmental cognitive augmentation. It opens up avenues for the development of automated systems that can adapt not just content but other aspects of the analyst workflow based upon individual cognitive profiles. Such systems would leverage cognitive data to tailor information presentation and tool supports, thereby minimizing cognitive overload and reducing the likelihood of analytic errors. This approach to system design

underscores the importance of cognitive ergonomics as a means to create cognitive augmentation tools that align with the cognitive capabilities and limitations of users, ultimately fostering a more effective analytical environment. By developing analytic tools that adapt to the user's cognitive needs, we can significantly enhance the data analysis process.

These tools would be capable of presenting information at an optimal level of complexity and density, tailored to the user's current cognitive state, thus preventing information overload and improving decision-making accuracy. Personalization of these systems is another key benefit, allowing for adjustments based on individual cognitive profiles in almost real-time ~250 ms. This could lead to a more intuitive user experience, reducing the risk of cognitive fatigue and increasing overall job success. Furthermore, the insights from this research can be leveraged to improve training programs for analysts. By focusing on enhancing cognitive control abilities, analysts can be better equipped to handle complex data without an increase in error rates. Additionally, the work suggests that by carefully modifying the textual and modal characteristics of the information presented, it is possible to reduce analytic failures. Organizations can use these findings to create guidelines and analyst centered adaption that helps analysts process information more efficiently, leading to more accurate inferences.

Lastly, this research emphasizes the importance of cognitive ergonomics in the design of analytic systems. By creating tools that are in harmony with the cognitive strengths and limitations of users, we can foster an analytic environment that not only supports the analysts in their work but also promotes a more effective and error-free analytical process. This approach to system design is crucial for the development of future analytical tools that are both "user-friendly" and "highly functional". However, for success in these areas we must be able to measure underlying cognitive and affective states.

To practically implement the findings from this work on cognitive demand, organizations can take several steps. They can develop adaptive analytic systems that adjust information complexity based on real-time cognitive demand assessments. Training programs can be customized to enhance cognitive abilities, using simulations that replicate real-world data analysis scenarios. Organizations can also set information design standards beyond readability to optimize the presentation of content preventing excessive cognitive demand. Collecting and analyzing cognitive data from analysts during their tasks will also help in creating individualized support strategies. Conducting user experience research and human factors research as opposed to user design approaches is essential to understand how analysts interact with systems and to inform design improvements. Finally, feedback mechanisms can be introduced for analysts to communicate their cognitive needs, allowing for system adjustments. These measures can lead to a supportive work environment that promotes cognitive well-being and improves productivity and accuracy in analytical tasks. The success of these implementations however depends on continuous refinement through feedback and performance metrics based upon the measures.

The key to minimizing analytical errors lies not merely in enhancing data gathering capabilities, but rather in crafting and implementing systems that are inherently adaptive to individual needs. Such systems are designed to furnish analysts with optimally tailored cognitive environmental frameworks that serve to externally augment their analytical processes. This augmentation is instrumental in diminishing cognitive demand while simultaneously bolstering the analysts' capacity for information processing and endurance when processing large amounts of information. It's important to note that this augmentation does not inherently imply a diminution, condensation, or any form of scaling down of the information at hand—although such approaches are not entirely dismissed. Instead, the primary goal is to achieve augmentation by reconfiguring the environmental context in a manner that capitalizes on and enhances the analysts' pre-existing cognitive functions. By reorganizing the environment, we can alleviate the cognitive burden and amplify the analysts' proficiency in exercising higher levels of information processing.

4.1. Limitations
While there is considerable application for neurocognitive data in the development and assessment of cognitive demand and cognitive control there are limitations within this work. fNIRS is limited in its ability to obtain neurological data in other regions of interest the brain which are subcortical and located away from the sensors. These regions might be responsible for and capable of compensating for the processing of other cognitive systems, thereby indicating or alleviating demand that may not be accounted for. A second limitation is the limited sample size of 33 participants. The a priori analysis of power shows a .95 probability of detecting a small effect, though the sample may not be completely representative of the varied types of analysts that are employed across the intelligence community and the tasks they are asked to complete.

5. Conclusion
Ultimately the reduction of analytic failure rests not in greater collection capability, but in the development of systems and which adapt and provide the analyst with appropriate cognitive environmental structures providing external augmentation. These augmentative structures assist in the reductions in cognitive demand and increases in information processing capability [45].

References
1. Moreau, C., Rouaud, T., Grabli, D., Benatru, I., Remy, P., Marques, A. R., ... & Fabbri, M. (2023). Overview on wearable sensors for the management of Parkinson's disease. *npj Parkinson's Disease, 9*(1), 153.
2. Katerina, T., & Nicolaos, P. (2018). Mouse behavioral patterns and keystroke dynamics in End-User Development: What can

they tell us about users' behavioral attributes?. *Computers in Human Behavior, 83,* 288-305.

3. Whisenand, T. G., & Emurian, H. H. (1999). Analysis of cursor movements with a mouse. *Computers in Human Behavior, 15*(1), 85-103.

4. Gupta, S., Maple, C., Crispo, B., Raja, K., Yautsiukhin, A., & Martinelli, F. (2023). A survey of human-computer interaction (HCI) & natural habits-based behavioural biometric modalities for user recognition schemes. *Pattern Recognition, 139*, 109453.

5. Torres, E. B., Brincker, M., Isenhower, R. W., Yanovich, P., Stigler, K. A., Nurnberger, J. I., ... & José, J. V. (2013). Autism: the micro-movement perspective. *Frontiers in integrative neuroscience, 7*, 32.

6. Prather, E. A., Badr, A. S., Simões, B., & De Amicis, R. (2020). A systematic literature review on dynamic cognitive augmentation through immersive reality: challenges and perspectives. *Virtual, Augmented, and Mixed Reality (XR) Technology for Multi-Domain Operations, 11426*, 74-93.

7. Lamb, R. L., Annetta, L., Firestone, J., & Etopio, E. (2018). A meta-analysis with examination of moderators of student cognition, affect, and learning outcomes while using serious educational games, serious games, and simulations. *Computers in Human Behavior, 80*, 158-167.

8. Lamb, R., Neumann, K., & Linder, K. A. (2022). Real-time prediction of science student learning outcomes using machine learning classification of hemodynamics during virtual reality and online learning sessions. *Computers and Education: Artificial Intelligence, 3*, 100078.

9. Lamb, R., Akmal, T., & Petrie, K. (2015). Development of a cognition-priming model describing learning in a STEM classroom. *Journal of Research in Science Teaching, 52*(3), 410-437.

10. Kudyba, S., Fjermestad, J., & Davenport, T. (2020). A research model for identifying factors that drive effective decision-making and the future of work. *Journal of Intellectual Capital, 21*(6), 835-851.

11. Lim, L., Bannert, M., van der Graaf, J., Singh, S., Fan, Y., Surendrannair, S., ... & Gašević, D. (2023). Effects of real-time analytics-based personalized scaffolds on students' self-regulated learning. *Computers in Human Behavior, 139*, 107547.

12. Bawden, D., & Robinson, L. (2020). Information overload: An introduction. In *Oxford research encyclopedia of politics*.

13. Zamboni, K., Baker, U., Tyagi, M., Schellenberg, J., Hill, Z., & Hanson, C. (2020). How and under what circumstances do quality improvement collaboratives lead to better outcomes? A systematic review. *Implementation Science, 15*(1), 27.

14. Hickman, L., Thapa, S., Tay, L., Cao, M., & Srinivasan, P. (2022). Text preprocessing for text mining in organizational research: Review and recommendations. *Organizational Research Methods, 25*(1), 114-146.

15. Wei, X., Saab, N., & Admiraal, W. (2021). Assessment of cognitive, behavioral, and affective learning outcomes in massive open online courses: A systematic literature review. *Computers & Education, 163*, 104097.

16. Frantz, J., Cupido-Masters, J., Moosajee, F., & Smith, M. R. (2022). Non-cognitive support for postgraduate studies: A systematic review. *Frontiers in Psychology, 12*, 773910.

17. da Silva, M. D., & Postma, M. (2020). Wandering minds, wandering mice: Computer mouse tracking as a method to detect mind wandering. *Computers in Human Behavior, 112*, 106453.

18. Lockey, S., Gillespie, N., Holm, D., & Someh, I. A. (2021). A review of trust in artificial intelligence: Challenges, vulnerabilities and future directions.

19. Lamb, R., Firestone, J., Kavner, A., Almusharraf, N., Choi, I., Owens, T., & Rodrigues, H. (2024). Machine learning prediction of mental health strategy selection in school aged children using neurocognitive data. *Computers in Human Behavior, 156,* 108197.

20. Samuel, J., Kashyap, R., Samuel, Y., & Pelaez, A. (2022). Adaptive cognitive fit: Artificial intelligence augmented management of information facets and representations. *International journal of information management, 65*, 102505.

21. Jangwan, N. S., Ashraf, G. M., Ram, V., Singh, V., Alghamdi, B. S., Abuzenadah, A. M., & Singh, M. F. (2022). Brain augmentation and neuroscience technologies: current applications, challenges, ethics and future prospects. *Frontiers in Systems Neuroscience, 16*, 1000495.

22. Davern, M., Shaft, T., & Te'eni, D. (2012). Cognition matters: Enduring questions in cognitive IS research. *Journal of the Association for Information Systems, 13*(4), 1.

23. Schmidt-Weigand, F., Kohnert, A., & Glowalla, U. (2010). A closer look at split visual attention in system-and self-paced instruction in multimedia learning. *Learning and instruction, 20*(2), 100-110.

24. Asan, O., & Choudhury, A. (2021). Research trends in artificial intelligence applications in human factors health care: mapping review. *JMIR human factors, 8*(2), e28236.

25. Kunde, W., Reuss, H., & Kiesel, A. (2012). Consciousness and cognitive control. *Advances in cognitive psychology, 8*(1), 9.

26. Turker, S., Seither-Preisler, A., & Reiterer, S. M. (2021). Examining individual differences in language learning: A neurocognitive model of language aptitude. *Neurobiology of Language, 2*(3), 389-415.

27. Lamb, R., Hoston, D., Lin, J., & Firestone, J. (2022). Psychological allostatic load: The cost of persistence in STEM disciplines. *Research in Science Education, 52*(4), 1187-1206.

28. Lamb, R., & Firestone, J. (2022). The moderating role of creativity and the effect of virtual reality on stress and cognitive demand during preservice teacher learning. *Computers & Education: X Reality, 1,* 100003.

29. Lamb, R., Cavagnetto, A., & Akmal, T. (2016). Examination of the nonlinear dynamic systems associated with science student cognition while engaging in science information processing. *International Journal of Science and Mathematics Education, 14*(Suppl 1), 187-205.

30. Lamb, R. L. (2013). *The application of cognitive diagnostic approaches via neural network analysis of serious educational games.* George Mason University.

31. Lamb, R., Firestone, J., Schmitter-Edgecombe, M., & Hand, B. (2019). A computational model of student cognitive processes while solving a critical thinking problem in science. *The Journal of Educational Research, 112*(2), 243-254.

32. Embrey, J. R., Donkin, C., & Newell, B. R. (2023). Is all mental effort equal? The role of cognitive demand-type on effort avoidance. *Cognition, 236*, 105440.

33. Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological review, 80*(4), 237.

34. Kalyuga, S. (2011). Cognitive load theory: How many types of load does it really need?. *Educational psychology review, 23*(1), 1-19.

35. Klepsch, M., & Seufert, T. (2020). Understanding instructional design effects by differentiated measurement of intrinsic, extraneous, and germane cognitive load. *Instructional Science, 48*(1), 45-77.

36. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big?. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623).

37. McLaughlin, A. C., & Byrne, V. E. (2020). A fundamental cognitive taxonomy for cognition aids. *Human Factors, 62*(6), 865-873.

38. Pinti, P., Tachtsidis, I., Hamilton, A., Hirsch, J., Aichelburg, C., Gilbert, S., & Burgess, P. W. (2020). The present and future use of functional near-infrared spectroscopy (fNIRS) for cognitive neuroscience. *Annals of the new York Academy of Sciences, 1464*(1), 5-29.

39. Nguyen, H. D., Yoo, S. H., Bhutta, M. R., & Hong, K. S. (2018). Adaptive filtering of physiological noises in fNIRS data. *Biomedical engineering online, 17*(1), 180.

40. Boisgontier, M. P., & Cheval, B. (2016). The anova to mixed model transition. *Neuroscience & Biobehavioral Reviews, 68*, 1004-1005.

41. Willaby, H. W., Costa, D. S., Burns, B. D., MacCann, C., & Roberts, R. D. (2015). Testing complex models with small sample sizes: A historical overview and empirical demonstration of what partial least squares (PLS) can offer differential psychology. *Personality and Individual Differences, 84*, 73-78.

42. Musick, G., O'Neill, T. A., Schelble, B. G., McNeese, N. J., & Henke, J. B. (2021). What happens when humans believe their teammate is an AI? An investigation into humans teaming with autonomy. *Computers in Human Behavior, 122,* 106852.

43. Hauptman, M., Rogers, M. L., Scarpaci, M., Morin, B., & Vivier, P. M. (2023). Neighborhood disparities and the burden of lead poisoning. *Pediatric research, 94*(2), 826-836.

44. Villa, D. L., Schostek, T., Govertsen, K., & Macmillan, M. (2023). A stochastic model of future extreme temperature events for infrastructure analysis. *Environmental Modelling & Software, 163,* 105663.

45. Embrey, J. R., Li, A. X., Liew, S. X., & Newell, B. R. (2024). The effect of noninstrumental information on reward learning. *Memory & Cognition, 52*(5), 1210-1227.