Does What We See Shape History? Examining Workload History as a Function of Performance and Ambient/Focal Visual Attention

SHANNON P. DEVLIN, University of Virginia
JENNIFER K. BYHAM, Clemson University
SARA LU RIGGS, University of Virginia

Changes in task demands can have delayed adverse impacts on performance. This phenomenon, known as the workload history effect, is especially of concern in dynamic work domains where operators manage fluctuating task demands. The existing workload history literature does not depict a consistent picture regarding how these effects manifest, prompting research to consider measures that are informative on the operator's process. One promising measure is visual attention patterns, due to its informativeness on various cognitive processes. To explore its ability to explain workload history effects, participants completed a task in an unmanned aerial vehicle command and control testbed where workload transitioned gradually and suddenly. The participants' performance and visual attention patterns were studied over time to identify workload history effects. The eye-tracking analysis consisted of using a recently developed eye-tracking metric called coefficient *K*, as it indicates whether visual attention is more focal or ambient. The performance results found workload history effects, but it depended on the workload level, time elapsed, and performance measure. The eye-tracking analysis suggested performance suffered when focal attention was deployed during low workload, which was an unexpected finding. When synthesizing these results, they suggest unexpected visual attention patterns can impact performance immediately over time. Further research is needed; however, this work shows the value of including a real-time visual attention measure, such as coefficient *K*, as a means to understand how the operator manages varying task demands in complex work environments.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**; • **Applied computing →** *Psychology*;

Additional Key Words and Phrases: Workload history, eye tracking, UAV

ACM Reference format:
Shannon P. Devlin, Jennifer K. Byham, and Sara Lu Riggs. 2021. Does What We See Shape History? Examining Workload History as a Function of Performance and Ambient/Focal Visual Attention. *ACM Trans. Appl. Percept.* 18, 2, Article 8 (May 2021), 17 pages.
https://doi.org/10.1145/3449066

Shannon P. Devlin, Jennifer K. Byham, and Sara Lu Riggs are contributed equally to this research.
This study was supported in part by the National Science Foundation (NSF Grant No. 1566346, Program Manager: Dr. Ephraim Glinert; NSF Grant No. 1750850, Program Manager: Dr. Andrew Kerne).
Authors' addresses: S. P. Devlin and S. L. Riggs, University of Virginia, 151 Engineer's Way, Charlottesville, Virginia, 22904; emails: {spd7t, sriggs}@virginia.edu; J. K. Byham, Clemson University, 211 Fernow Street, Clemson, South Carolina, 29634; email: jbyham@g.clemson.edu.

1544-3558/2021/05-ART8 $15.00
https://doi.org/10.1145/3449066

1 INTRODUCTION

Workload transitions—i.e., changes in the number and/or complexity of task demands—are inherently embedded in dynamic work domains such as healthcare, aviation, and military operations. For example, operators overseeing the command and control of **unmanned aerial vehicles (UAVs)** are subject to varying workload levels as they manage various responsibilities, task demands, and automation levels [61]. For instance, UAV operators are predisposed to more frequent and drastic workload transitions as they complete tasks for multiple UAVs [32]. Previous research has shown that changes in task demands can affect performance later, which is known as a workload history effect [6, 10, 21, 26, 48]. For example, even if an operator can manage a workload transition, it may have ramifications on operator performance at a later period in time [33].

To date, workload history is not completely understood as there is a lack of consensus across studies. Workload history effects have been found to be dependent on the direction of the transition (i.e., transitions from low to high versus high to low workload), type of task (audio versus visual tasks, single versus dual tasks [8]), and measures of performance (errors of commission versus omission [6]). Analyzing subjective [20, 30, 36, 50], psychophysiological [2, 19], and personality measures [7, 9] post hoc have shed more light on workload history effects, but the investigation is not complete. For example, there is little work exploring relevant environmental factors, such as transition rate (i.e., transitioning slowly or quickly), or specifics about the operator's response to managing workload levels over time, like his/her visual attention patterns.

The goal of the present work is to determine whether eye-tracking analysis, namely, visual attention type, can be a non-invasive and quantitative means to further understand how operators are impacted by workload transitions over time. The eye-tracking metric called coefficient K can discriminate between two types of visual attention: focal (i.e., focused) and ambient (i.e., global) attention. Discerning between these two attention types can be an indication of how the operator is managing workload transitions—for instance, does the operator consistently concentrate their visual attention or disperse it across the entire display? Knowing this information can potentially inform, and ideally be incorporated in, future technologies to thwart adverse effects due to workload history.

2 BACKGROUND

The **National Research Council (NRC)** was tasked in 1988 to examine the effects of having extended periods of underload follow intense periods of overload to prepare the U.S. Army for its reduction of crew size and increase in automation reliance [33]. They found performance measures did not recover over time even though the ultimate level of load decreased to low workload. As a result, the NRC initiated research to study workload history—i.e., how fluctuations in task load affects performance over time—given its relevance for a variety of data-rich environments [33]. Previous research helped provide insights into workload history effects, but there was still a lack of consensus [10, 26, 46, 48]. For example, performance after a workload transition has been found to improve [40, 48], deteriorate [6], neither improve nor deteriorate [30, 36, 40, 48], and/or can change over time [24]. Often, workload history effects depend on which performance metric was measured [6, 40, 50]. Subsequently, theoretical explanations of workload history are not widely agreed upon [6]. More recent work [2, 12, 36] supports the effort regulation explanation, which suggests people can mitigate workload history effects if they can accurately assess the environment's dynamics and develop effective management strategies [28, 31].

Even with the continued support of the effort regulation explanation, workload history is still an open research topic [29]. For example, the growing dynamism in complex domains [1] may compound currently observed workload history effects. However, continuous, real-time, psychophysiological measures [2, 20] offer as a promising means to address some of these challenges. For example, when performance suffers after a workload transition from low to high workload, **electroencephalograms (EEG)** have shown an increase in the electrical activity in the areas of the brain that are associated with the mental processes of "sustained attention, conflict resolution, and rapid updating of working memory.'' This suggests operators produce a physiological response to workload transitions and it can be reliably tracked. However, the authors also found unexplained inconsistencies with some measures and suggest it may be necessary to rely on multiple EEG measures to detect workload history effects [2]. Similarly, Kim et al. [40] built upon this work by examining EEG correlates of the low workload periods in a low-high-low workload transition. Minimal differences between performance were observed, but they found the area of the brain that was responsible for "cognitive behavior and decision making" was higher for the second low workload period than the first. The authors stated that these findings indicated a hysteresis effect was present, even though performance was unaffected. Although these findings are interesting, it is worth investigating if eye tracking can serve as a less invasive and less complex psychophysiological measure to capture the impacts of workload history in real-time. Of interest is to see whether eye tracking can be more of informative on the process operators use to manage workload transitions and, in turn, how that process impacts performance. Analyzing eye-tracking data has been found to "add another dimension to traditional speed/performance analysis" [16], making it a potential method to address this current need in the workload history literature.

2.1 Using Eye-tracking Technology to Understand Workload History

Eye-tracking technology typically relies on the corneal-reflection technique, which shines an infrared light in to a person's eye to create and track a single reflection point on the cornea. The resulting data consists of timestamped Cartesional coordinates, which are based on the eye tracker's sampling rate and the display's resolution. These coordinates are then used to distinguish between fixations and saccades. Fixations are when the eye is relatively stationary and characterize about 90% of viewing behavior. Fixations are when information processing occurs. Saccades are the ballistic movements between fixations [16, 55]. Fixations and saccades are the basis of most eye-tracking metrics, i.e., the way in which visual attention patterns are quantified [55]. Examples of metrics include the amount of time a given fixation lasts (i.e., fixation duration) or the size of a saccade (i.e., saccade amplitude; [60]). These metrics aim to capture the, "objective and quantitative evidence of the user's visual, overt attentional processes, based on the user's scan patterns" [15]. To date, eye-tracking metrics have seldomly been used in workload history research.

Often, metrics are aggregated across experimental conditions for comparison [11, 25, 37, 44]; however this type of analysis often overlooks changes in visual attention patterns over time. For example, a study conducted by Jiang et al. [38] found that as participants completed a web search task, their performance and scan patterns were inversely proportional to the workload changes over time. In other words, as participants completed more search tasks (i.e., increase in workload), both the likelihood of selecting the correct search result and mean fixation duration decreased. This work demonstrates the importance of capturing how visual attention may evolve over time. However, the interpretation of a certain metric's trend over time can be convoluted [53]. New metrics that account for the order in which certain types of fixations and saccades occur over time, such as coefficient K, have the potential to capture and accurately depict the dynamism of visual attention [42].

2.2 Coefficient K: A Dynamic Measure of Ambient and Focal Visual Attention

Coefficient K accounts for changes in the magnitude and sequence of each fixation and saccade in a scan pattern. It has been found to effectively distinguish between the two types of visual attention used during visual search tasks: ambient and focal [42]. Ambient visual attention occurs when people are gaining spatial orientation (i.e., "getting a sense" of the environment), whereas focal visual attention occurs when people are processing the details of the environment [3, 56]. Ambient visual attention usually consists of a pattern where sequences of short fixations are followed by long saccades. Oppositely, focal visual attention usually consists of sequences of long fixations followed by short saccades [3, 63]. The interaction between ambient and focal visual attention in scene perception is dynamic [56, 63]. For example, when a scene is initially being examined, there is typically more ambient visual attention (i.e., shorter fixations and longer saccades), but as objects are identified, there is more focal visual attention (i.e., fixation durations increase and saccades decrease [34, 52]).

When it comes to studying visual attention, most analyses aggregate the duration and magnitude of the fixations and saccades over time. There has been limited work analyzing how eye-tracking metrics change over time and whether these changes exhibit particular changes in visual attention patterns. Pannasch et al. [54] compared fixation durations and saccade amplitudes during the early and late phases of scene perception to better understand the relationship between ambient and focal visual attention over time. However, their analysis aggregated fixations and saccades over blocks of time so it did not capture how the sequence of fixations and saccades impacts the evolution of visual attention patterns over time. However, Krejtz et al. [42] developed a metric that distinguishes between ambient and focal visual attention each time a fixation or saccade occurs. Equation (1) shows how coefficient K is calculated. Essentially, it is the difference between standardized values (Z-score) of each fixation duration (d_i) and its following saccade amplitude (a_{i+1}):

$$K_i = \frac{d_i - \mu_d}{\sigma_d} - \frac{a_{i+1} - \mu_a}{\sigma_a} \text{ such that } \frac{1}{n}\sum K_i = 1. \tag{1}$$

Here, μ_d is the mean fixation duration, μ_a is the mean saccade amplitude, σ_d and σ_a are the respective standard deviations of fixation duration and saccade amplitude for the entire data set to account for any bias. These parameters is then calculated for each sequential fixation and then averaged over all fixations *(n)* [42]. Coefficient K is a measure of standard deviation; a value of 1 indicates that "the duration of the current fixation is beyond 1 standard deviation longer than the subsequent saccade amplitude," whereas a value of −1 indicates that "a saccade is more than 1 standard deviation longer than the preceding fixation duration" [42]. Positive coefficient K values are an indicator of focal attention as they occur when long fixations are followed by short saccades. Negative coefficient K values are an indicator of ambient visual attention as they occur when short fixations are followed by long saccades [63]. A coefficient K value that approaches zero suggests that fixations and subsequent saccades are relatively equal with their respective means and is not an indicator of either attention type. *(*Note: This occurrence is rare on an individual level, although may occur when averaging).*

Coefficient K was empirically validated by comparing the values during serial and parallel search [42]. It was found that the most focal eye movements (i.e., positive coefficient K values) occurred during serial search and the most ambient eye movements (i.e., negative coefficient K values) occurred during parallel search. These results align with what is expected during serial and parallel searches, confirming the validity of coefficient K, i.e., its ability to distinguish between the two types of visual attention. Previous research has used it to distinguish the visual attention patterns when viewing artwork [43], completing cartographic tasks [41], and between socially anxious and non-anxious viewers [45]. In all cases, coefficient K has been more informative on the details of the focal-ambient viewing dynamics, which further explains the viewer's overall visual attention patterns in these environments. Coefficient K may further shed light on the effects of workload history by quantifying the visual attention patterns in dynamic environments.

2.3 Motivation

This work aims to further understand the performance effects of workload history by using the coefficient K eye-tracking metric to assess how operators change their visual attention with different workload transition rates, workload levels, and over time. An eye-tracking metric that is sensitive to how visual attention evolves with each fixation and saccade can provide as a process-based measure of the operator [16]. This could potentially lead to a more in depth understanding of the workload history effect—a large gap in the knowledgebase—while also potentially informing the design of intelligent technology for domains where workload transitions are prevalent. We are exploring the potential of coefficient K, because it characterizes real-time visual attention in a straightforward format and allows for direct interpretation. The present study examines (a) whether a workload history effect is present when workload transitions from low to high at different rates and (b) whether visual attention type is informative of any of the present workload history effects. Specifically, the research questions are:

(1) **Is there a workload history effect present in the performance results (i.e., a decrease in response accuracy and increase in response time over time) and does this effect differ for different transition rates and workload levels?**
   (a) Expectations: We expect response time and accuracy to degrade with workload transitions [6, 8, 27, 50] and it to be more pronounced for gradual transitions compared to sudden ones [51]. However, we expect performance to partially recover and then plateau for low workload periods as this is consistent with the previous research that studies workload transitions in dynamic environments like the one used in the present study [17, 36, 39, 50].
(2) **Is the coefficient K sensitive to the differences between low and high workload periods?**
   (a) Expectations: There has been limited work to establish whether coefficient K can be a measure of mental workload, let alone workload transitions (tangentially related work [14]). However, it is expected that coefficient K will be sensitive to low and high workload as the inputs for coefficient K (i.e., fixation duration and saccade amplitude) are sensitive to different levels of workload [5, 35]. Specifically, it is predicted that high workload will lead to more focal attention (i.e., higher positive coefficient K values), as increases in workload often cause operators to have increased fixation durations [5] and to narrow their visual attention [62], which in turn decreases saccade amplitude. These two changes in fixations and saccades lead to coefficient K being positive (Equation (1)). For low workload, it is expected that attention will be more ambient as visual attention is usually less direct and more exploratory (i.e., larger saccades and shorter fixations) [13, 49].
(3) **Is coefficient K informative of workload history effects?**
   (a) Expectations: Based on its previous success [41, 45], we expect coefficient K to serve as a quantitative measure of the theoretical explanation that people can mitigate workload history effects if they can assess the current needs of the environment and then develop effective management strategies based on that assessment [28, 31, 36]. If this is the case, then participants will first survey the overall environment (i.e., use ambient visual attention and have negative coefficient K values) to "get a sense" and evaluate its dynamics [31]. During this evaluation, performance may suffer if ambient visual attention is not ideal for the current environment. However, once participants develop a management strategy to account for varying workloads, coefficient K values will most likely be positive during high workload periods (i.e., focal visual attention) and negative during low workload periods (i.e., ambient visual attention).

The context of this experiment is UAV command and control, since workload transitions are inherent in their operations and are expected to continue as the number of UAVs an operator is responsible for increases [64]. The broader impact of the research questions is addressing the research gaps in the workload history literature as a means to potentially inform the design of intelligent technology, such as adaptive displays, in complex and dynamic work environments.

3 METHOD

This work uses the same methodology and experimental testbed from References [12, 49], but focuses on the performance results and visual attention patterns over time.

3.1 Participants

Twenty-one undergraduate students participated in this study (13 male; mean age = 20.9 years, SD = 1.5). Participants self-reported normal or corrected to normal vision and were compensated $10/h for their participation. Two participants' data was excluded from analysis, due to a malfunction with the experimental testbed during the study and poor performance in the primary task, respectively. The study was approved by the Institutional Review Board at Clemson University (IRB2015-217) and all participants provided informed consent.

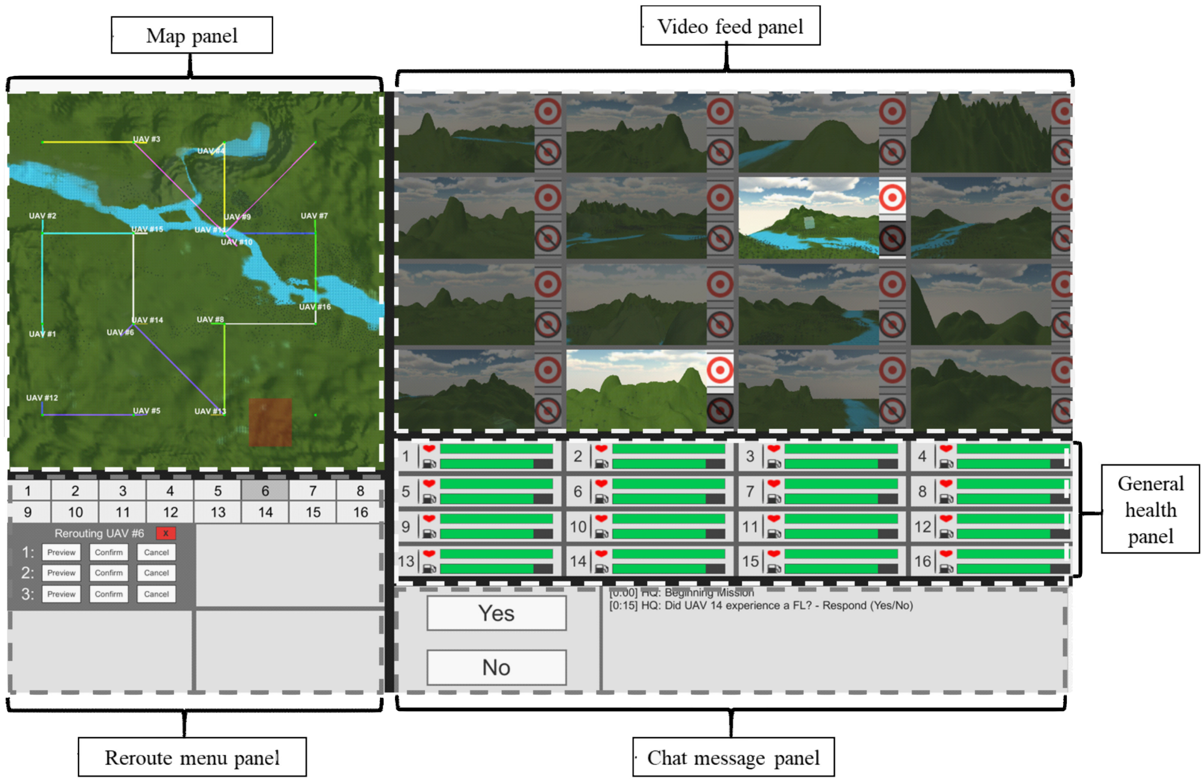

Fig. 1. The interface of the UAV testbed.

3.2 Apparatus

Participants completed the testbed scenarios on a desktop computer with a 32 in. monitor (2,560 × 1,600 resolution) and using a standard mouse. A desktop-mounted corneal reflection FOVIO eye tracker (Seeing Machines platform) with a sampling rate of 60 Hz was used to collect gaze data. Participants sat 71–78 cm from the eye tracker, which was placed 2 cm below the bottom edge of the monitor. The mean degree of error for the FOVIO eye tracker is 0.78° (SD = 0.59°; [18]). Participants completed a 5-point calibration procedure before each recording and the accuracy of the calibration was verified by the experimenter before proceeding.

3.3 Experimental Setup

The UAV command and control testbed was based on the **Vigilant Spirit Control Station (VSCS)** used by the United States Air Force [23]. The testbed tasked participants with four tasks: (1) detecting targets, (2) rerouting UAVs, (3) stopping fuel leaks, and (4) responding to chat messages. Figure 1 is a screenshot of the UAV testbed with labelled panels and Table 1 shows how each task mapped to each panel.

3.4 Testbed Tasks

3.4.1 Primary Task: Target Detection Task. The participants' primary task was called the target detection task. The video feed of each UAV could become "active" during the scenario, meaning the specific UAV video feed would be illuminated (Figure 2). Participants were tasked with searching for a target, i.e., a semi-transparent cube, in the active UAV video feeds. If the participant saw a target, then they were instructed to select the "target

Table 1. Tasks and Corresponding Panels

Task	Panels
Target detection task	Video feed panel
Reroute task	Reroute menu panel and Map panel
Fuel leak task	General health panel
Chat message task	Chat message panel

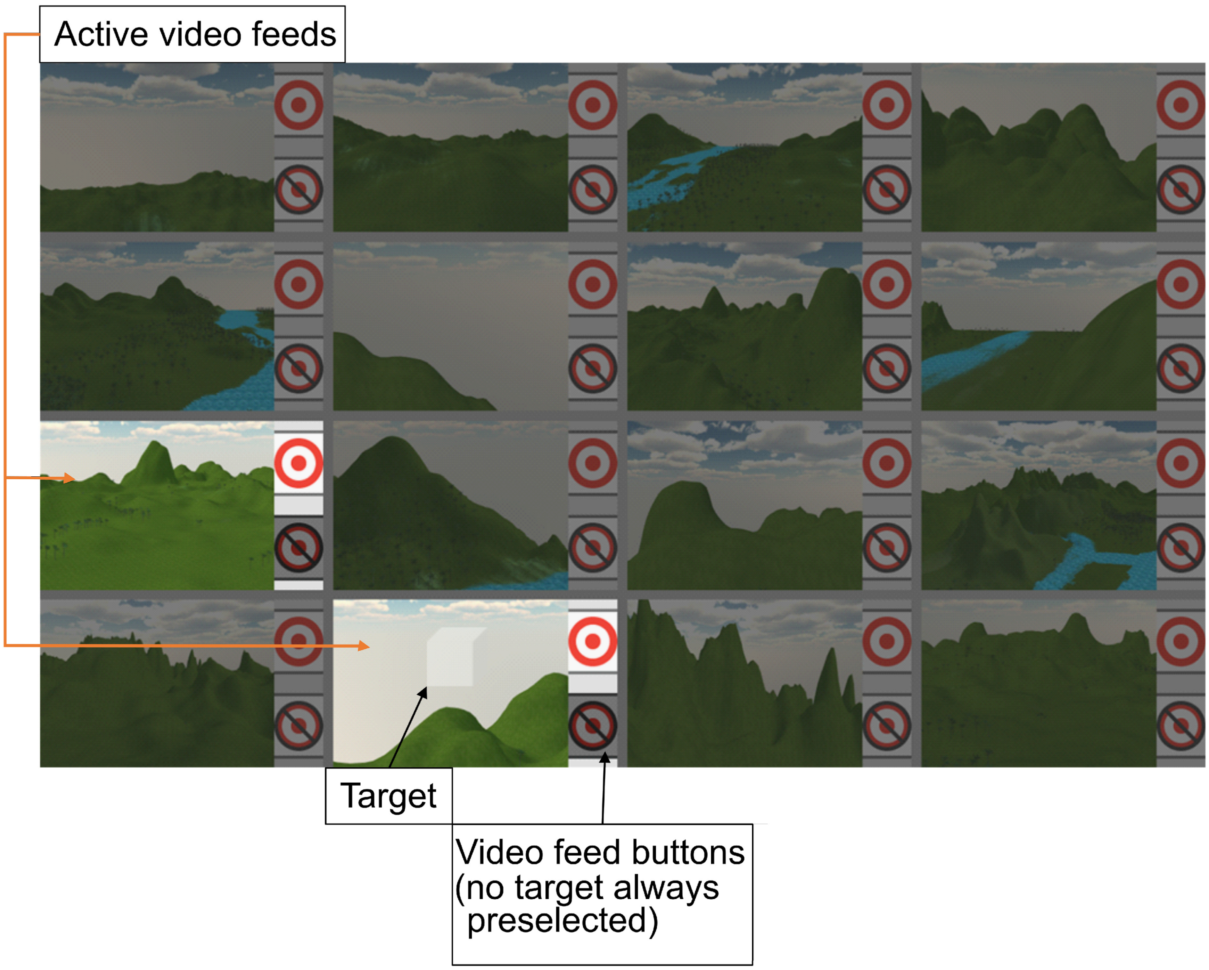

Fig. 2. Example of active and inactive UAVs on the Video Feed panel and how to detect a target.

button" in the top right-hand corner of the UAV video feed. Participants had 10 seconds to respond to each target detection task. If there was no target, then no action was necessary given the "no target button" (which was right below the "target button") was automatically selected for each active UAV video feed. On average, targets were present in 20% of the active video feeds. The number of simultaneously active UAVs varied from three to sixteen and determined the workload (see Section 3.5).

3.4.2 Secondary Tasks. While completing the primary target detection task, participants were also tasked to attend to three secondary tasks: (1) reroute task, (2) fuel leak task, and (3) chat message task. Participants were instructed to respond quickly and accurately to these secondary tasks.

(1) **Reroute task.** This required the participant to be vigilant of when a UAV was going to enter a no-fly-zone, i.e., the red square on the Map panel. If this was the case, then the participant would need to select an alternative route to avoid the no-fly-zone for that UAV in the Reroute menu panel. To reroute a UAV, participants clicked on the UAV's numbered button on the Reroute menu panel and selected from three alternatives. If the UAV was not successfully rerouted in time (i.e., UAV entered the no-fly-zone), then that UAV would be unable to participate in the target detection task (primary task) for the remainder of the scenario.
(2) **Fuel leak task.** This required the participant to monitor the fuel status of each UAV in the General health panel. If a UAV was experiencing a fuel leak, then the fuel level dropped three times faster than the normal rate, which was displayed in the fuel level bar (the bar next to gas pump icon on the General health panel). The health status bar (the bar next to the heart icon on the General health panel) would turn yellow and a "FIX FUEL LEAK" warning would appear. To stop a fuel leak, participants had to click on the health status bar so fuel usage would return to normal. Failure to do so would lead to a "FATAL FUEL LEAK" condition for the remainder of the scenario.
(3) **Chat message task.** Participants had to respond to chat messages as quickly and accurately as possible. Every time a new chat message appeared, participants could choose from one of two options that appeared in the Chat message panel.

3.5 Workload Modulation

Each participant completed two 15-min scenarios in the UAV testbed. The scenarios differed by the rate workload transitioned (i.e., gradually or suddenly) from low to high. NASA-TLX data from pilot tests showed that 3–5 simultaneously active UAV video feeds impose low workload, whereas 13–16 simultaneously active UAV video feeds impose high workload. In the gradual transition scenario, workload transitioned from low to high steadily starting with a low workload period lasting 20 s, followed by a transition period to high workload of 40 s, and then a high workload period lasting two minutes. This cycle repeated five times during the 15-min scenario. For the sudden transition scenario, the low workload period lasted one minute and was followed by high workload period lasting two minutes. This low to high workload transition also repeated five times. Figure 3 depicts the workload levels over the course of each 15-min transition scenario. The transition scenarios, workload level, and workload periods were the basis of our repeated measures ANOVA analysis (more details in Section 4.2).

3.6 Procedure

Participants completed the experiment over the course of two consecutive days at approximately the same time each day to avoid fatigue effects. On Day 1, participants were given informed consent, briefed about the experimental goals, and then trained on the testbed's tasks. Training consisted of an interactive presentation with the experimenter and completing a five-minute scenario at slightly above low workload (4–6 UAVs active). After establishing proficiency in all the tasks, i.e., a minimum of 70% accuracy was achieved, participants completed two constant workload conditions, one consisting of constant low workload and one consisting of constant high workload, in a counterbalanced order. They then completed a demographic and debriefing questionnaire. On Day 2, participants completed the gradual and sudden transition scenario in a counterbalanced order, were debriefed again, and then compensated for their time. Scenario order, i.e., gradual and sudden, was counterbalanced across participants. The present work only analyzed data from Day 2 given it included the gradual and sudden transition scenarios, which was pertinent to the stated research questions. The results from Day 1 are reported in References [12, 49].

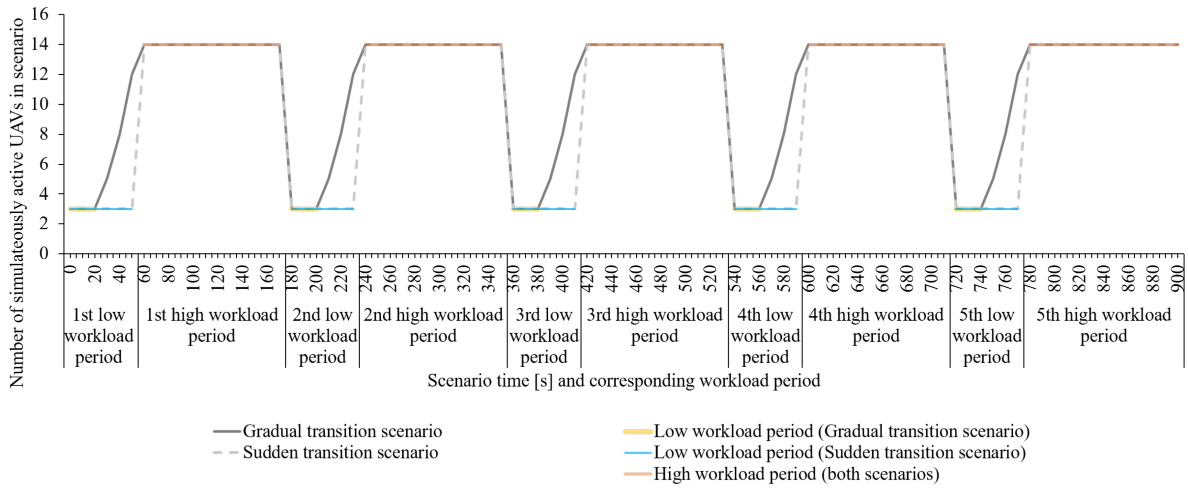

Fig. 3. The number of UAVs active throughout the gradual and sudden transition scenarios.

4 RESULTS

4.1 Data Reduction

The raw gaze data for 19 participants was preprocessed by custom VBA scripts, where missing and invalid data (e.g., coordinates outside the screen and blinks) was removed. The mean data loss across all participants and trials was 11.9%, (SD = 11.2%). It was then smoothed with a second-order Butterworth filter, with a 60 Hz sampling and 6.15 Hz cutoff frequency. A velocity-based algorithm [60] was used to distinguish saccades from fixations. Saccades were any eye movement above the velocity threshold of 20°/s; otherwise, all other data points were classified as fixations. This procedure matched the one used to empirically validate coefficient K [42].

4.2 Experimental Design

A $2 \times 2 \times 5$ **repeated measures analysis of variance (RM ANOVA)** was used to analyze workload history effects of performance and visual attention type. There were two workload transition rates (gradual, sudden), two levels of workload (low, high), and five workload periods (1st, 2nd, . . . , 5th; refer to the highlighted sections of Figure 3 as those were the only sections included in the analysis). Specific main and interaction effects of the RM ANOVA were of interest to our research questions and are indicated throughout the results. However, we report the results of each main effect to give a general overview of each independent variable's impact. The dependent performance measures were based on the primary task (i.e., target detection), because workload was manipulated using this task. Response time for the primary target detection task was calculated from the appearance of the target to when participants clicked the "target button." Accuracy was calculated as the percentage of correct "target detected" responses that occurred within the 10 second time limit. "No target detected" responses were not included in either performance measure given it did not require an explicit response from the participant. Both measures were averaged for each workload period over time, resulting in five low workload periods and five high workload periods. The primary task varied the number of active UAVs depending on the workload period, meaning there were slight differences on the number of tasks to attend to across participants (e.g., if a UAV was lost to the no-fly-zone). However, it never deviated from the expected workload level. The secondary tasks occurred at a constant rate of one every 20 s. Coefficient K was calculated as described in Equation (1). Significance was set at $\alpha = 0.05$ and post-hoc tests with Tukey's adjustment method were used to determine significant differences between means. Violations of normality were assessed prior to analysis and Greenhouse-Geisser corrections were used when sphericity was violated. RStudio 1.2.1335 was used for all analyses [58]. General eta-squared (η_G^2) is reported as a measure of effect size for the omnibus test and values of 0.01, 0.06, and 0.14 are interpreted as small, medium, and large effect sizes, respectively [4]. For all posthoc pairwise comparisons, effect size was measured with Cohen's d for repeated measures (d_{rm}; [47]) and values of .2, .5, and .8 are interpreted as small, medium, and large effect sizes [4]. For all graphs, error bars represent the standard deviation.

4.3 Performance Results

4.3.1 Response Time. When comparing gradual to sudden workload transitions, there was a main effect of transition rate ($F(1,18) = 18.828$, $p < 0.001$, $\eta_G^2 = 0.047$). Response time for gradual transitions (M = 2.82 s, SD = 0.22 s) was significantly slower than sudden transitions (M = 2.60 s, SD = 0.14 s). When comparing low to high workload, there was a main effect of workload ($F(1,18) = 229.236$, $p < 0.0001$, $\eta_G^2 = 0.327$). Response times were faster for low workload (M = 2.38 s, SD = 0.22 s) than high workload (M = 3.05 s, SD = 0.09 s). Mauchly's test of sphericity indicated that the assumption of sphericity had been violated for period ($\chi^2(9) = 0.242$, $p = 0.006$). Using the Greenhouse-Geisser correction of $\epsilon = .589$, there was a significant effect of period ($F(2.356, 42.405) = 6.810$, $p < .0001$, $\eta_G^2 = 0.080$) indicating there was a workload history effect.

To understand how workload history impacted the primary task response time, we examine the three-way interaction effect between transition rate, workload level, and workload period. The results of this effect begins to address our first research question, i.e., is there a workload history effect present in the performance results and does this effect differ for different transition rates and workload levels? Mauchly's test of sphericity indicated that the assumption of sphericity had been violated for the three-way interaction effect ($\chi^2(9) = 0.292$, $p = 0.017$). Using the Greenhouse-Geisser correction of $\epsilon = 0.608$, there was a significant three-way interaction effect ($F(2.429, 43.718) = 5.123$, $p = 0.001$, $\eta_G^2 = 0.049$). Post-hoc tests showed that for gradual transitions, response times were faster in the 1st low workload period (M = 1.98 s, SD = 0.34 s) compared to the 3rd–5th low workload periods (3rd period: M = 2.99 s, SD = 0.77 s, $p < 0.0001$, $d_{\mathrm{rm}} = 1.627$; 4th period: M = 2.74 s, SD = 1.30 s, $p < 0.001$, $d_{rm} = 0.800$; 5th period: M = 3.06 s, SD = 0.70 s, $p < 0.0001$, $d_{\mathrm{rm}} = 1.93$). In addition the 2nd low workload period (M = 2.23 s, SD = 0.76 s) was significantly faster than the 3rd low workload period ($p < 0.001$, $d_{\mathrm{rm}} = 1.000$) and 5th low workload period ($p = 0.0001$, $d_{\mathrm{rm}} = 1.127$). For the sudden transitions, the 2nd low workload period (M = 1.9 s, SD = 0.30 s) was significantly faster than the 3rd low workload period (M = 2.51 s, SD = 0.39 s, $p = 0.022$, $d_{\mathrm{rm}} = 1.724$). Finally, there was one significant difference between the gradual and sudden transitions; the 5th low workload period of gradual transitions was significantly slower than the 5th low workload period of sudden transitions (M = 2.28 s, SD = 0.43 s, $p < 0.0001$, $d_{\mathrm{rm}} = 1.260$). Figure 4 shows response time over time based on workload and transition rate.

4.3.2 Accuracy. There was no main effect of transition rate ($F(1,18) = 2.923$ $p = 0.105$, $\eta_G^2 = 0.017$). However, there was a significant main effect of workload level ($F(1, 18)=360.452$, $p < 0.001$, $\eta_G^2 = 0.521$) and period ($F(4, 72) = 5.501$, $p < 0.001$, $\eta_G^2 = 0.043$. As expected, post-hoc tests showed that accuracy for the low workload (M = 90%, SD = 5.7%) was significantly higher than high workload (M = 66%, SD = 6.5%). The significant effect of period indicated there was a workload history effect.

Again, here we focus on the three-way interaction effect as it completes our investigation of the first research question, i.e., is there a workload history effect present in the performance results and does this effect differ for different transition rates and workload levels? There was a significant three-way interaction between transition rate, workload level, and period ($F(4,72) = 3.274$, $p = 0.016$, $\eta_G^2 = 0.027$). For gradual transitions, accuracy for the 1st low workload period (M = 99%, SD = 4.6%) was significantly higher than all subsequent low workload period (2nd period: M = 86%, SD = 12.7%, $p = 0.008$, $d_{rm} = 1.466$; 3rd period: M = 85%, SD = 16.1%, $p = 0.006$, $d_{rm} = 1.208$; 4th period: M = 88%, SD = 25.5%; $p = 0.077$, $d_{\mathrm{rm}} = 0.638$; 5th period: M = 81%, SD = 23.1%; $p < 0.0001$, $d_{\mathrm{rm}} = 1.152$). With sudden transitions, the 4th high workload period (M = 75%, SD = 7.2%) had significantly

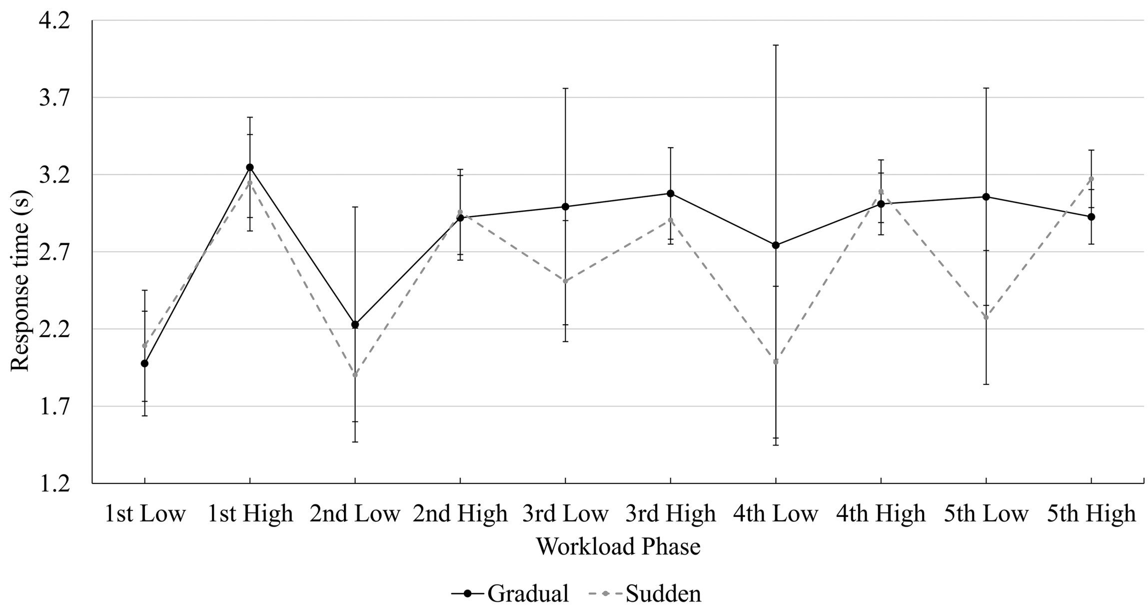

Fig. 4. Mean primary task response time for each workload period for the gradual and sudden transition scenarios.

Fig. 5. Mean primary task performance accuracy for each workload period for the gradual and sudden transition scenarios.

higher accuracy than 5th high workload period (M = 62%, SD = 9.6%; $p = 0.009$, $d_{\text{rm}} = 1.472$). Figure 5 shows accuracy rates over time based on workload and transition rate.

4.3.3 Coefficient K. There was no main effect of transition rate ($F(1,18) = 0.053$, $p = 0.820$, $\eta_G^2 < 0.0001$). There was a main effect of workload level ($F(1,18) = 40.314$, $p < 0.0001$, $\eta_G^2 = 0.185$) and period ($F(4,72) = 6.596$, $p = 0.0001$, $\eta_G^2 = 0.074$). The main effect of workload level addresses our second research question (i.e., is the coefficient K sensitive to the differences between low and high workload periods?). Post-hoc tests indicate

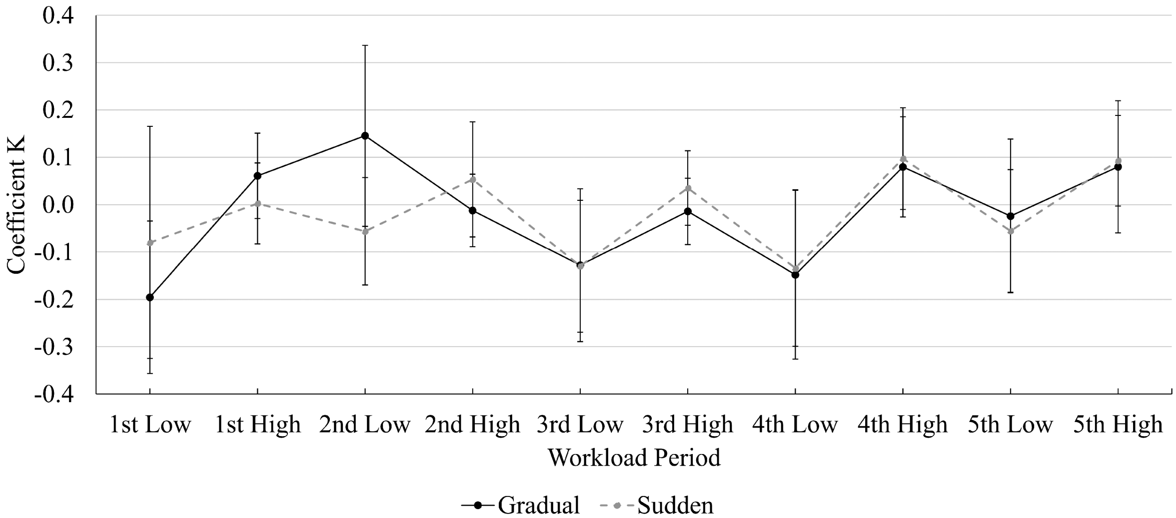

Fig. 6. Mean coefficient K values for each workload period for the gradual and sudden transition scenarios.

coefficient K for low workload (M = −0.08, SD = 0.06) was significantly lower than high workload (M = 0.048, SD = 0.04), indicating visual attention was more likely to be more ambient during low workload and more focal during high workload.

To fully address our third research question, i.e., is coefficient K informative of workload history effects, we examined the three-way interaction effect between transition rate, workload level, and workload period. There was a significant three-way interaction between transition rate, workload level, and period ($F(4,72)= 6.568$, $p = 0.0001$, $\eta^2_G = 0.064$). This interaction effect specifically addresses our third research question. For gradual transitions, the 2nd low workload period (M = 0.15, SD = 0.19) had significantly higher coefficient K values than all other low workload periods (1st period: M = −0.19, SD = 0.16, $p < 0.0001$, $d_{\mathrm{rm}} = 1.922$; 3rd period: M = −0.13, SD = 0.16, $p < 0.0001$, $d_{\mathrm{rm}} = 1.547$; 4th period: M = −0.15, SD = 0.17, $p < 0.0001$, $d_{rm} = 1.587$; 5th period: M = −0.02, SD = 0.16, $p = 0.028$, $d_{rm} = 0.955$). In addition, coefficient K for the 2nd low workload period was also significantly higher than the 2nd low workload period of sudden transitions (M = −0.06, SD = 0.11, $p = 0.001$, $d_{\mathrm{rm}} = 1.270$). Also with gradual transitions, the 1st low workload period had a significantly lower coefficient K value than the 5th low workload period ($p = 0.024$, $d_{\mathrm{rm}} = 1.06$). Figure 6 shows coefficient K values over time based on workload and transition rate.

5 DISCUSSION

The overarching goal of this work was to assess whether eye tracking, specifically visual attention type, can be a non-invasive, cognitive, and quantitative means to further understand how people are impacted by workload transitions over time, i.e., the workload history effect. We specifically explored this by addressing three research questions:

(1) **Is there a workload history effect present in the performance results (i.e., a decrease in response accuracy and increase in response time over time) and does this effect differ for different transition rates and workload levels?**
(2) **Is the coefficient *K* sensitive to the differences between low and high workload periods?**
(3) **Is coefficient *K* informative of workload history effects?**

For the first research question, we expected response time and accuracy to degrade and then plateau over time. The results here did reveal a workload history effect for only low workload periods, but it was not the same for gradual and sudden transitions. Similar to previous work, both workload transition rates had faster response times in earlier low workload periods (i.e., 1st and 2nd) than later ones (3rd–5th). However, the results showed that the workload history effect was more pronounced for gradual transitions as response times increased and then plateaued at these slower response times during the latter low workload periods [2, 50]. For sudden transitions, response time increased significantly during the middle of the scenario (i.e., 3rd low workload period), but then recovered to initial speeds later on [36, 57].

Similarly, for accuracy, we also observed workload history effects that manifested over time during the low workload periods, but only for gradual transitions [51]. Consistent with previous work and our expectations, accuracy was the highest at the beginning of the scenario and then decreased over time [2, 8, 30, 50]. On the other hand, with sudden transitions, accuracy remained relatively stable over time for low workload periods. The only performance difference within sudden transitions was the significant drop in accuracy between the last two high workload periods, although it is unclear if this is due to workload history or fatigue [6]. Nevertheless, the performance results support a subset of previous work and our expectations, in that they clearly show workload history effect are a function of workload transition rate, workload level, and experimental setting.

For the second research question, we sought to determine whether coefficient K is sensitive to differences in low and high workload when it transitions. Our findings support our expectations: high workload resulted in more focal visual attention (i.e., positive coefficient K values) whereas low workload results in more ambient visual attention (i.e., negative coefficient K values). This supports the premise that an increase in cognitive load leads to an increase in fixation duration and a decrease in saccade amplitudes, i.e., focal attention. Given the limited work that has employed coefficient K, the findings show that this eye-tracking metric quantifies visual attention type and may be a promising means to assess workload in real-time.

For the third research question, we consider the performance and coefficient K results together, to see whether these two data streams can provide a more complete understanding of the workload history effect. Overall, our expectations for this research question were partially met when considering the result of sudden transitions: there was a decline in performance as participants engaged in more ambient visual attention initially, but performance improved later after coefficient K became aligned with workload level expectations (i.e., positive coefficient K during high workload periods and negative coefficient K during low workload periods). The findings for sudden transitions show that a participant's performance can recover if visual attention allocation strategies are rapidly developed in accordance with each workload level. Oppositely, for gradual transitions, coefficient K peaked during the 2nd low workload period as it was significantly higher than all other low workload periods and the corresponding low workload period of sudden transitions. The significant increase in coefficient K coincides with the first time low workload periods of gradual transitions irreversibly increase in response times and decrease in accuracy rates. This may suggest that workload history effects with gradual transitions may be due to a large, unexpected, increase in focal attention during low workload. Similarly, for gradual transitions, coefficient K was significantly larger in the final low workload period compared to the first. The findings here may indicate that an increase in focal attention during low workload periods of workload transitions may be an indicator of future performance decrements (i.e., workload history effects) and not an improved acquaintance with the environment like previous work suggests [34, 41, 45]. However, previous work has shown workload history effects are more likely to be apparent during low workload periods [2, 6, 48], but now coefficient K provides a quantitative explanation as to why this is the case—i.e., not using the appropriate visual attention type as a function of workload. Furthermore, even though coefficient K values eventually converged to similar patterns over time for both transition rates, performance trends did not. One potential explanation could be adopting ineffective strategies at the onset may have both immediate and delayed effects on performance, even if more effective strategies are eventually adopted as seen with gradual transitions here. It may also show that visual attention strategies adopted for sudden transitions may not be best for gradual transitions and vice versa. A follow-up study should ideally examine workload history over a longer time frame to see if performance recovers when an effective visual attention management strategy is eventually adopted.

Overall, these results suggest that it is not only important to adopt a strategy to account for workload transitions, but that performance may irrevocably suffer (i.e., a workload history effect endures) if an ineffective visual attention strategy is initially adopted. The findings also support the premise that people initially evaluate the environment to develop a workload transition management strategy, but struggle to update this strategy over time [52, 57]. The findings from this study add to the theoretical explanations of how operators manage workload transitions and how this process impacts performance. Previous research that examined workload history with other psychophysiological measures are successful in detecting workload history effects, but the practical interpretation of these measures is unclear [2, 40]. However, the findings here have shown the potential to be used in real-time and inform the design of intelligent technology [22, 59].

6 LIMITATIONS

More work is needed before concretely relating coefficient K to workload history effects. First, a larger sample size should be used to see if these results replicate. Second, it might be best to investigate how different lengths of time at each workload level impact coefficient K results. Here, gradual and sudden transitions had different amounts of time in low workload due to the experimental setup and it is unclear if this influenced the coefficient K results [42]. Two potential ways this difference in duration could impact the coefficient K comparisons made between the two transition rates include (a) coefficient K values in the low workload periods of the gradual transition were more sensitive by major fluctuations in fixation duration and/or saccade amplitude or (b) the longer durations in the low workload periods of sudden transitions gave participants more time to establish a consistent visual attention strategy for that given workload period. Albeit, this is a limitation of the present study, coefficient K still appears to be a promising metric in capturing changes to visual attention and performance outcomes over time when workload changes. We also recommend that future work compare workload history effects between subgroups of performers (e.g., best and worst performers). Given the spread of the data for both the performance and visual attention type measures was considerable, it warrants further explanation at an individual level. This analysis could also contribute to exploring the potential coefficient K has to be a real-time indicator of any individual differences that are affecting performance.

7 CONCLUSION

In summary, this study highlights the nuanced nature surrounding workload history, in that these effects evolve differently over time in dynamic, multitasking domains. The findings here suggest that eye-tracking technology can be a powerful means of explaining these nuances over time. Namely, eye tracking can provide some insight on how visual attention was impacted by workload transitions and subsequently, its impact on performance. This is critical if intelligent technology is going to be implemented to improve performance in dynamic environments like UAV command and control. Specifically, this study addressed the need to consider different transition rates, multitasking, multiple transitions, and eye tracking to understand the nuanced effects of workload history [36]. This work also shows how coefficient K has promise to be a real-time, proactive indicator of visual attention strategies that might later lead to negative workload history effects during low workload. Therefore, coefficient K can be the indicator of when the operator needs assistance **and** how to provide that assistance (e.g., encourage the operator to engage in a certain visual attention type when workload transitions). This potential has not been as promising with other real-time, cognitive-based measures (e.g., EEG) [2, 40]. Overall, this work further supports the value of using process-based psychophysiological measures like eye-tracking to fill research gaps in workload history literature as they provide a more complete picture of how the person manages workload transitions, which can serve as the basis of informing technology design in dynamic environments.

ACKNOWLEDGMENT

The authors thank Aakash Bhagat for his help in developing the testbed used in this study.

REFERENCES

[1] Julian Abich IV, Lauren Reinerman-Jones, and Gerald Matthews. 2017. Impact of three task demand factors on simulated unmanned system intelligence, surveillance, and reconnaissance operations. *Ergonomics* 60, 6 (2017), 791–809.

[2] Margaret A. Bowers, James C. Christensen, and F. Thomas Eggemeier. 2014. The effects of workload transitions in a multitasking environment. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 58. SAGE Publications, Los Angeles, CA, 220–224.

[3] Guy Thomas Buswell. 1935. *How people look at pictures: A study of the psychology and perception in art.* Univ. Chicago Press.

[4] Jacob Cohen. 1988. The effect size index: d. *Stat. Power Anal. Behav. Sci.* 2 (1988), 284–288.

[5] Melissa Patricia Coral. 2016. *Analyzing cognitive workload through eye-related measurements: A meta-analysis*. Ph.D. Dissertation. Wright State University.

[6] Luz-Eugenia Cox-Fuenzalida. 2007. Effect of workload history on task performance. *Human Factors* 49, 2 (2007), 277–291. DOI: https://doi.org/10.1518/001872007X312496

[7] Luz-Eugenia Cox-Fuenzalida, Amanda Angie, Sheri Holloway, and Laura Sohl. 2006. Extraversion and task performance: A fresh look through the workload history lens. *J. Res. Personal.* 40, 4 (2006), 432–439.

[8] Luz-Eugenia Cox-Fuenzalida and Amanda D. Angie. 2005. The effects of workload history on dual task performance. *Curr. Psychol.* 24, 3 (2005), 171–179.

[9] Luz-Eugenia Cox-Fuenzalida, Rhonda Swickert, and James B. Hittner. 2004. Effects of neuroticism and workload history on performance. *Personal. Individ. Dif.* 36, 2 (2004), 447–456.

[10] R. W. Cumming and P. G. Croft. 1973. Human information processing under varying task demand. *Ergonomics* 16, 5 (1973), 581–586.

[11] Edward Cutrell and Zhiwei Guan. 2007. What are you looking for? An eye-tracking study of information usage in web search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* 407–416.

[12] Shannon Devlin, Nadine Moacdieh, Christopher Wickens, and Sara Lu Riggs. 2020. Transitions between low and high levels of mental workload can improve multitasking performance. *IISE Trans. Occupation. Ergonom. Human Factors* 8:2 (2020), 72–87.

[13] Francesco Di Nocera, Marco Camilli, and Michela Terenzi. 2007. A random glance at the flight deck: Pilots' scanning strategies and the real-time assessment of mental workload. *J. Cogn. Eng. Decis. Mak.* 1, 3 (2007), 271–285.

[14] Andrew Duchowski, Krzysztof Krejtz, Justyna Zurawska, and Donald House. 2019. Using microsaccades to estimate task difficulty during visual search of layered surfaces. *IEEE Trans. Visual. Comput. Graph.* (2019).

[15] Andrew T. Duchowski. 2002. A breadth-first survey of eye-tracking applications. *Behav. Res. Methods Instrum. Comput.* 34, 4 (2002), 455–470.

[16] Andrew T. Duchowski. 2007. Eye-tracking methodology. *Theory Pract.* 328, 614 (2007), 2–3.

[17] Tamsyn Edwards, Lynne Martin, Nancy Bienert, and Joey Mercer. 2017. The relationship between workload and performance in air traffic control: Exploring the influence of levels of automation and variation in task demand. In *Proceedings of the International Symposium on Human Mental Workload: Models and Applications*. Springer, 120–139.

[18] EyeTracking, Inc. 2011. *Hardware: Eye Tracking Systems.* EyeTracking, Inc. Retrieved from http://www.eyetracking.com/Hardware/Eye-Tracker-List.

[19] Majid Fallahi, Majid Motamedzade, Rashid Heidarimoghadam, Ali Reza Soltanian, Maryam Farhadian, and Shinji Miyake. 2016. Analysis of the mental workload of city traffic control operators while monitoring traffic density: A field study. *Int. J. Industr. Ergonom.* 54 (2016), 170–177.

[20] Majid Fallahi, Majid Motamedzade, Rashid Heidarimoghadam, Ali Reza Soltanian, and Shinji Miyake. 2016. Assessment of operators' mental workload using physiological and subjective measures in cement, city traffic, and power plant control centers. *Health Promot. Perspect.* 6, 2 (2016), 96.

[21] Philip S. E. Farrell. 1999. The hysteresis effect. *Human Factors* 41, 2 (1999), 226–240.

[22] Karen M. Feigh, Michael C. Dorneich, and Caroline C. Hayes. 2012. Toward a characterization of adaptive systems: A framework for researchers and system designers. *Human Factors* 54, 6 (2012), 1008–1024.

[23] Gregory Feitshans, Allen Rowe, Jason Davis, Michael Holland, and Lee Berger. 2008. Vigilant spirit control station (VSCS): The face of COUNTER. In *Proceedings of the AIAA Guidance, Navigation and Control Conference and Exhibit.* 6309.

[24] Jonathan P. Gluckman, Joel S. Warm, William N. Dember, and Roger R. Rosa. 1993. Demand transitions and sustained attention. *J. Gen. Psychol.* 120, 3 (1993), 323–337.

[25] Joseph H. Goldberg and Xerxes P. Kotval. 1999. Computer interface evaluation using eye movements: Methods and constructs. *Int. J. Industr. Ergonom.* 24, 6 (1999), 631–645.

[26] Robert A. Goldberg and Manard R. Stewart. 1980. Memory overload or expectancy effect? “Hysteresis” revisited. *Ergonomics* 23, 12 (1980), 1173–1178.
[27] P. A. Hancock, George Williams, and C. M. Manning. 1995. Influence of task demand characteristics on workload and performance. *Int. J. Aviat. Psychol.* 5, 1 (1995), 63–86.
[28] Peter A. Hancock. 1989. A dynamic model of stress and sustained attention. *Human Factors* 31, 5 (1989), 519–537.
[29] Peter A. Hancock and Gerald Matthews. 2019. Workload and performance: Associations, insensitivities, and dissociations. *Human Factors* 61, 3 (2019), 374–392.
[30] William S. Helton, Tyler Shaw, Joel S. Warm, Gerald Matthews, and Peter Hancock. 2008. Effects of warned and unwarned demand transitions on vigilance performance and stress. *Anxiety Stress Coping* 21, 2 (2008), 173–184.
[31] G. R. Hockey. 1997. Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biol. Psychol.* 45, 1–3 (1997), 73–93.
[32] Becky L. Hooey, David B. Kaber, Julie A. Adams, Terrence W. Fong, and Brian F. Gore. 2017. The underpinnings of workload in unmanned vehicle systems. *IEEE Trans. Hum.-Mach. Syst.* 48, 5 (2017), 452–467.
[33] Beverly M. Huey and Christopher D. Wickens. 1993. *Workload Transition: Implications for Individual and Team Performance.* Technical Report. National Research Council, Washington, DC.
[34] David E. Irwin and Gregory J. Zelinsky. 2002. Eye movements and scene perception: Memory for things observed. *Percept. Psychophys.* 64, 6 (2002), 882–895.
[35] Robert J. K. Jacob and Keith S. Karn. 2003. Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In *The Mind’s Eye.* Elsevier, 573–605.
[36] Reinier J. Jansen, Ben D. Sawyer, René van Egmond, Huib de Ridder, and Peter A. Hancock. 2016. Hysteresis in mental workload and task performance: The influence of demand transitions and task prioritization. *Human Factors* 58, 8 (2016), 1143–1157.
[37] Halszka Jarodzka, Katharina Scheiter, Peter Gerjets, and Tamara Van Gog. 2010. In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learn. Instruct.* 20, 2 (2010), 146–154.
[38] Jiepu Jiang, Daqing He, and James Allan. 2014. Searching, browsing, and clicking in a search session: Changes in user behavior by task and over time. In *Proceedings of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval.* 607–616.
[39] Nayoung Kim, Wonjoon Kim, Myung Hwan Yun, and Chang S. Nam. 2018. Behavioral and neural correlates of hysteresis effects during multitasking. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 62. SAGE Publications, Los Angeles, CA, 11–13.
[40] Na Young Kim, Russell House, Myung H. Yun, and Chang S. Nam. 2019. Neural correlates of workload transition in multitasking: An ACT-R model of hysteresis effect. *Front. Hum. Neurosci.* 12 (2019), 535.
[41] Krzysztof Krejtz, Arzu Cöltekin, Andrew T. Duchowski, and Anna Niedzielska. 2017. Using coefficient K to distinguish ambient/focal visual attention during map viewing. *J. Eye Move. Res.* 10, 2 (2017), 1–13.
[42] Krzysztof Krejtz, Andrew Duchowski, Izabela Krejtz, Agnieszka Szarkowska, and Agata Kopacz. 2016. Discerning ambient/focal attention with coefficient K. *ACM Trans. Appl. Percept.* 13, 3 (2016), 1–20.
[43] Krzysztof Krejtz, Andrew T. Duchowski, Izabela Krejtz, Agata Kopacz, and Piotr Chrząstowski-Wachtel. 2016. Gaze transitions when learning with multimedia. *Journal of Eye Movement Research* 9, 1 (2016).
[44] Krzysztof Krejtz, Tomasz Szmidt, Andrew T. Duchowski, and Izabela Krejtz. 2014. Entropy-based statistical analysis of eye movement transitions. In *Proceedings of the Symposium on Eye Tracking Research and Applications.* 159–166.
[45] Krzysztof Krejtz, Katarzyna Wisiecka, Izabela Krejtz, Paweł Holas, Michał Olszanowski, and Andrew T. Duchowski. 2018. Dynamics of emotional facial expression recognition in individuals with social anxiety. In *Proceedings of the ACM Symposium on Eye Tracking Research and Applications.* 1–9.
[46] Judith E. Krulewitz, Joel S. Warm, and Theodore H. Wohl. 1975. Effects of shifts in the rate of repetitive stimulation on sustained attention. *Percept. Psychophys.* 18, 4 (1975), 245–249.
[47] Daniël Lakens. 2013. Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Front. Psychol.* 4 (2013), 863.
[48] Michael L. Matthews. 1986. The influence of visual workload history on visual performance. *Human Factors* 28, 6 (1986), 623–632.
[49] Nadine Marie Moacdieh, Shannon P. Devlin, Hussein Jundi, and Sara Lu Riggs. 2020. Effects of workload and workload transitions on attention allocation in a dual-task environment: Evidence from eye-tracking metrics. *J. Cogn. Eng. Decis. Mak.* (2020), 1555343419892184.
[50] Justin F. Morgan and Peter A. Hancock. 2011. The effect of prior task loading on mental workload: An example of hysteresis in driving. *Human Factors* 53, 1 (2011), 75–86. DOI: https://doi.org/10.1177/0018720810393505
[51] Brian W. Moroney, Joel S. Warm, and William N. Dember. 1995. Effects of demand transitions on vigilance performance and perceived workload. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 39. SAGE Publications, Los Angeles, CA, 1375–1379.
[52] E. A. B. Over, I. T. C. Hooge, B. N. S. Vlaskamp, and C. J. Erkelens. 2007. Coarse-to-fine eye movement strategy in visual search. *Vision Res.* 47, 17 (2007), 2272–2280.

[53] Bing Pan, Helene A. Hembrooke, Geri K. Gay, Laura A. Granka, Matthew K. Feusner, and Jill K. Newman. 2004. The determinants of web page viewing behavior: An eye-tracking study. In *Proceedings of the Symposium on Eye Tracking Research and Applications.* 147–154.

[54] Sebastian Pannasch, Jens R. Helmert, Katharina Roth, Ann-Katrin Herbold, and Henrik Walter. 2008. Visual fixation durations and saccade amplitudes: Shifting relationship in a variety of conditions. *Journal of Eye Movement Research* 2, 2 (2008).

[55] Alex Poole and Linden J. Ball. 2006. Eye tracking in HCI and usability research. In *Encyclopedia of Human Computer Interaction.* IGI Global, 211–219.

[56] Fred H. Previc. 1998. The neuropsychology of 3-D space.*Psychol. Bull.* 124, 2 (1998), 123.

[57] Erik G. Prytz and Mark W. Scerbo. 2015. Changes in stress and subjective workload over time following a workload transition. *Theoret. Issues Ergonom. Sci.* 16, 6 (2015), 586–605.

[58] R Core Team. 2019. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/.

[59] Ling Rothrock, Richard Koubek, Frederic Fuchs, Michael Haas, and Gavriel Salvendy. 2002. Review and reappraisal of adaptive interfaces: Toward biologically inspired paradigms. *Theoret. Issues Ergonom. Sci.* 3, 1 (2002), 47–84.

[60] Dario D. Salvucci and Joseph H. Goldberg. 2000. Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the Symposium on Eye Tracking Research and Applications.* 71–78.

[61] Ciara Sibley, Joseph Coyne, and Jeffrey Morrison. 2015. Research considerations for managing future unmanned systems. In *Proceedings of the AAAI Spring Symposium.* 454–459.

[62] Lisa C. Thomas and Christopher D. Wickens. 2004. Eye tracking and individual differences in off-normal event detection when flying with a synthetic vision system display. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 48. Sage Publications, Los Angeles, CA, 223–227.

[63] Jens R. Helmert, Boris M. Velichkovsky, Markus Joos, and Sebastian Pannasch. 2005. Two visual systems and their eye movements: Evidence from static and dynamic scene perception. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, Vol. 27.

[64] James A. Winnefeld and Frank Kendall. 2013. Unmanned systems integrated roadmap FY 2013-2038. *Approved for Open Publication Reference*14-S (2013), 0553.

Received February 2020; revised January 2021; accepted February 2021