Applied Ergonomics 105 (2022) 103829

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

Scan-based eye tracking measures are predictive of workload transition performance

Shannon P. Devlin [a,b,*], Noelle L. Brown [a], Sabrina Drollinger [c], Ciara Sibley [a], Jawad Alami [b], Sara L. Riggs [b]

[a] U.S. Naval Research Laboratory, Washington, D.C, USA
[b] University of Virginia, Charlottesville, VA, USA
[c] U.S. Naval Aerospace Medical Institute, Pensacola, FL, USA

ARTICLE INFO

Keywords:
Workload transitions
Eye tracking
Growth curve modeling
Supervisory control

ABSTRACT

Given there is no unifying theory or design guidance for workload transitions, this work investigated how visual attention allocation patterns could inform both topics, by understanding if *scan-based* eye tracking metrics could predict workload transition performance trends in a context-relevant domain. The eye movements of sixty Naval flight students were tracked as workload transitioned at a slow, medium, and fast pace in an unmanned aerial vehicle testbed. Four scan-based metrics were significant predictors across the different growth curve models of response time and accuracy. Stationary gaze entropy (a measure of how dispersed visual attention transitions are across tasks) was predictive across all three transition rates. The other three predictive scan-based metrics captured different aspects of visual attention, including its *spread*, *directness*, and *duration*. The findings specify several missing details in both theory and design guidance, which is unprecedented, and serves as a basis of future workload transition research.

1. Introduction

Dynamic and data-rich domains rely on operators to complete various tasks simultaneously. For example, in the multitasking environment of unmanned aerial vehicle (UAV) command and control, the operator is responsible for keeping the mission stable, even when tasks change in frequency, complexity, or priority (Hooey et al., 2017; Sibley et al., 2015; Williams, 2006). When a shift in task demands occurs over a continuous period, it is a *workload transition* (Huey & Wickens,1993; Prytz and Scerbo, 2015). Complex environments may inundate the operator with visual information, making it important to understand how the operator is allocating her visual attention as she manages workload transitions (Abich et al., 2017). However, there is limited research using eye tracking to examine workload transitions (exception: Devlin et al., 2021; Moacdieh et al., 2020) and neither explore whether it can predict performance over time—the focus of this work. The goal here is to determine whether scan patterns are predictive of the performance trends observed over time with workload transitions. Specifically, we examine several eye tracking metrics as predictors in *growth curve models* of performance, i.e., models that predict change over time based on how each individual changes over time (Curran et al., 2010; Hoffman, 2015).

1.1. Review of previous investigations studying workload transitions over time

Performance trends are not consistent when workload transitions. For example, performance after a workload transition has been found to improve (Krulewitz et al., 1975; Matthews, 1986; Matthews and Desmond, 2002; Ungar et al., 2005), deteriorate (Cox-Fuenzalida, 2007; Matthews and Desmond, 2002; Ungar et al., 2005), neither improve nor deteriorate (Helton et al., 2008; Jansen et al., 2016; Kim et al., 2019; Morgan and Hancock, 2011), and/or alternate between improving and deteriorating over time (Devlin et al., 2021; Gluckman et al., 1993; Moroney et al., 1995). For multitasking environments, two explanations are primarily cited:

1. **Resource depletion.** Workload transitions deplete mental resources so performance initially suffers. However, once workload returns to low, the compensatory regeneration component is able to recuperate

* Corresponding author. U.S. Naval Research Laboratory, Washington, DC, USA.
E-mail address: shannon.devlin@nrl.navy.mil (S.P. Devlin).

https://doi.org/10.1016/j.apergo.2022.103829
Received 27 August 2021; Received in revised form 6 June 2022; Accepted 8 June 2022
Available online 2 August 2022

resources, allowing performance to then recover (Gluckman et al., 1993).

2. **Effort regulation.** Workload transition performance is dependent on the individual actively evaluating, recruiting, and deploying the requisite amount of mental resources. Performance is stable as long as the appraisal is correct and workload does not reach levels of overload (Hockey, 1997).

One potential way to further examine the two theoretical explanations is to identify and detail the operators' attentional processes during workload transitions. Previous investigations find participants produce a psychophysiological response to workload transitions (Bowers et al., 2014; Boyer et al., 2015; Cerruti et al., 2010; Kim et al., 2019; McKendrick and Harwood, 2019) meaning there is a specific biological response that corresponds to the response in mental activity (Fairclough, 2009; Matthews et al., 2015). For example, electroencephalograms (EEG) show that the electrical activity in specific, cognitive-related areas of the brain increase as workload increases, suggesting participants actively rely on certain mental resources when managing workload transitions (Bowers et al., 2014; Kim et al., 2019). However, the specific interpretation of EEG measures during workload transitions can have inconsistent and/or convoluted interpretations. Eye tracking is a non-invasive, flexible, and cost-effective technology (Krafka et al., 2016; Sibley et al., 2017) able to capture the "objective and quantitative evidence of the user's visual, overt attentional processes, based on the user's scan patterns'' (Duchowski, 2017, p. 247). Quantifying visual attention allocation patterns relies on *scan-based* eye tracking metrics, i.e., measures capturing the features of visual attention allocation (Poole and Ball, 2006). Examples of these types of metrics include the amount of time visual attention lasts, i.e., fixation duration, or how often the focus of visual attention switches, i.e., gaze transition rate (Goldberg and Kotval, 1999). Scan-based metrics rely on predetermined locations on the display, which are termed areas of interest (AOIs). Eye tracking has been mostly used to inform visual display design (Goldberg and Kotval, 1999; Poole and Ball, 2006), but also it has been used to further the understanding of information processing (Shiferaw et al., 2019a,b), cognitive load (Coral, 2016; Wilson and Russell, 2007), human-automation trust (e.g., Hergeth et al., 2016; Sarter et al., 2007; Victor et al., 2018), situation awareness (Ebeid and Gwizdka, 2018; Ratwani et al., 2010), and individual differences (Jarodzka et al., 2010; Raptis et al., 2017; Shic et al., 2008). However, to date, eye tracking data has seldom been used in the workload transition literature. Of particular interest is understanding if eye tracking can predict performance trends of the operator as workload transitions over time, which seems promising given its past predictive success (Barz et al., 2021; Steichen et al., 2013).

Two publications to date have examined scan-based metrics during workload transitions: Devlin et al. (2021) found studying the trends of a scan-based metric over time helped explain the trends of workload transition performance over time, which has been a topic with diverging, puzzling outcomes (take Jansen et al., 2016 and Morgan and Hancock, 2011 as an example). Moacdieh et al. (2020) compared a set of scan-based metrics between constant and transitioning workload. The scan-based metrics captured one of three aspects of visual attention allocation (Moacdieh and Sarter, 2015): *spread* (where are users looking?), *directness* (how efficiently are users scanning?), and *duration* (how long are users looking at a certain area?). Moacdieh et al. (2020) found spread and directness metrics were different between constant and transitioning workload, with newer-developed metrics being the most enlightening. For example, stationary gaze entropy, which measures how distributed transitions are across the set of AOIs, was lower during workload transitions than during constant workload. Given this was coupled with better multitasking performance, these findings suggested visual attention transitions to task-specific areas of the display should not be equal in multitasking environments. It emphasized the need to include scan-based metrics when studying workload transitions over time. We presently address this need by directly incorporating scan-based metrics into time-based analysis methods, specifically, growth curve modeling. Compared to a repeated measures ANOVA, this more direct and quantitative approach lends to understanding the attentional resources underlying performance outcomes of workload transitions.

1.2. Motivation and research questions

The goal of this study is to understand the extent to which scan-based eye tracking metrics can predict performance trends of three different transition rates (slow, medium, and fast). Specifically we explore a set of scan-based metrics as predictors in *growth curve models* of performance, i.e., models that estimate the change in performance over time based on how performance changes for each individual (Hoffman, 2015). Given workload transition performance can be dependent on the individual (Cox-Fuenzalida et al., 2004, 2006; Devlin and Riggs, 2018; McKendrick and Harwood, 2019; Mracek et al., 2014), it is critical for prediction models to account for the variability amongst individuals. Additionally, this modeling approach also allows for a more specific investigation on the predictive capability of scan-based metrics, as they can be specified to predict *average* performance and/or its *change* over time, allowing for a more informative prediction outcome.

The scan-based metrics used in the current work are from Moacdieh et al. (2020). We selected metrics that discriminated between constant and transitioning workload, while also assuring there was one metric from the three different aspects of visual attention allocation, i.e., *spread*, *directness*, and *duration* (Moacdieh and Sarter, 2015). Table 1 details the scan-based metrics used in the present work.

Previous work has applied growth curve modeling to eye tracking data (e.g., Ayasse and Wingfield, 2020; Barr, 2008; Mirman et al., 2008; Godfroid et al., 2018), but never as a direct predictor in growth curve models of workload transition performance. Given this investigation is the first of its kind, our research questions are as follows:

1. Are visual attention allocation patterns predictive of performance trends over time? and
2. Is this a function of workload transition rate?

We expect multiple scan-based metrics, especially spread and directness metrics, to be predictive of performance trends over time – in other words, multiple metrics will be significant predictors in each growth curve model. If successful, the findings may add the much-needed detail to the theoretical explanations surrounding workload transitions, e.g., *how* mental resources are deployed, and inform display design in complex, multitasking environments. Here, the application is UAV command and control as the Department of Defense is committed to designing displays that reliably assist operators in real-time (United States Department of Defense, 2017, p. 20).

2. Method

2.1. Participants

Sixty student Naval Aviators participated in this study (age: $M = 24.5$ years, $SD = 2.3$ years, 51 males). No participants reported having experience with UAV operations. Each participant completed three trials, i.e., the testbed scenarios detailed in **Workload transition rates**. Participants with more than 20% of their raw gaze samples missing were excluded from the analysis as recommended by Komogortsev et al. (2010). To assure the analysis was based on asymptotic performance, participants who were below the 25th percentile during training, i.e., 64% average accuracy across all tasks, only had their second and third trials, i.e., scenarios, included in the analysis. The final subset of data included 95 trials from 40 participants. Each workload transition rate was represented relatively equally (32, 34, and 29 trials from slow,

Table 1
The scan-based metrics explored as predictors of growth curves of performance (note: these measures were calculated for the entirety of each transition rate).

Metric	Definition and calculation
Spread metrics (where are users generally looking?)	
Spatial gaze density	The number of grid cells containing gaze points divided by the total number of cells. A 20 × 20 evenly-divided grid (128 × 72 pixels per cell) was created to cover the entire interface. A higher spatial gaze density would indicate a larger dispersion of attention (Goldberg and Kotval, 1999)
Stationary gaze entropy (SGE)	Stationary gaze entropy indicates how equally distributed a person's attention is, with larger values indicating more evenly spread attention across areas of interest (AOI) and lower values indicating more narrowed attention (Krejtz et al., 2015). It is calculated using the following equation $H_s = -\sum_{i \in AOIs} p_i \log_2(p_i)$ Where p_i represents the proportion of transitions to the *i*th state, i.e. the *i*th AOI (the AOIs are as defined in Fig. 1), and are based on the Markov property, i.e., transitions to a given state only depend on the current state (Shiferaw et al., 2019a,b). In order to accurately compare it across models, this value was normalized by dividing it by $\log_2(\textit{number of AOIs})$.
Directness metrics (how purposeful is attention transitions?)	
Gaze transition rate [grid cells/s]	The rate visual attention shifts between equal grid cells (Goldberg and Kotval, 1999). A higher rate indicates lower efficiency. The same grid cells used for spatial gaze density were used here.
Gaze transition entropy (GTE)	Gaze transition entropy represents the randomness and complexity of a person's eye movements, with higher values indicating more randomness and lower efficiency (Krejtz et al., 2015). It is calculated based on the following formula: $H_t = -\sum_{i \in AOIs} p_i \sum_{j \in AOIs} p_{ij} \log_2\left(p_{ij}\right)$ Where p_i is as described in stationary gaze entropy, and p_{ij} is the probability of transitioning form state *i* to state *j* in one fixation. Assuming the Markov property holds, this was calculated by counting the number of transitions from *i* to *j* and then dividing by the total number of transitions from *i* (Shiferaw et al., 2019a,b). This was done for each pairing of AOIs (which are defined in Fig. 1). In order to accurately compare it across models, this value was normalized by dividing it by $\log_2(\textit{number of AOIs})$.
Duration metrics (how long, in general, does attention last?)	
Average fixation duration [msec]	The amount of time a fixation lasts. A longer duration suggests that the user is extracting more information from the environment (Jacob and Karn, 2003)

medium, and fast transitions, respectively).

2.2. Experimental setup

This study used the same experimental testbed from Moacdieh et al. (2020) and Devlin et al. (2020, 2021), so further details can be found in these works. The testbed was developed using the Unity game development platform (see Fig. 1) and was based on the "Vigilant Spirit Control Station" (VSCS) used by the Air Force to develop interfaces to control multiple UAVs and included tasks typical of an UAV command and control environment, such as target detection and route planning (Feitshans et al., 2008). The testbed was presented on a ViewSonic 24" monitor (2560 × 1440 resolution, 60 Hz refresh rate) and participants sat approximately 65 cm from the monitor-mounted Gazepoint HD eye tracker (fs = 150 Hz, reported accuracy of 0.5–1°), so their point of gaze could be collected.

2.3. UAV command and control testbed and tasks

Participants were responsible for controlling and managing UAV tasks under three different workload transition rates. There were four tasks (one primary task and three secondary tasks) and each occurred across the five panels on the testbed's interface (Fig. 1).

2.3.1. Target detection task (primary task)

Participants were tasked to monitor up to 16 UAV video feeds on the Video Feed panel for a target, i.e., a semi-transparent cube (Fig. 2). A target was only present when a UAV video feed was active, i.e., when that specific UAV feed illuminated. UAV video feeds were active for 10 s, and targets could appear during this time. Participants were instructed to press the "target" button when they spotted a target, else they were instructed to leave the default "no target" button selected. The "no target" button was the default, as pilot testing suggested this better assessed the participant's ability to detect targets versus clicking quickly. Video feeds cycled between active and inactive and if a target was present in an active UAV, but the participant did not select the target button within 10 s, the participant missed the opportunity to detect that specific target. On average, 20% of active UAVs had a target. The number of simultaneously active UAVs determined the workload level (see section **Workload transition rates** for more details). Participants were instructed to prioritize the target detection task.

2.4. Secondary tasks

As the primary task occurred, participants also had to attend to three secondary tasks as quickly and accurately as possible: (1) reroute task, (2) fuel leak task, and (3) chat message task. One secondary task occurred every 20 s, on average, in a pseudorandom order. All emulated the diversified task load of future UAV operators and details of these tasks can be reviewed in (Devlin et al., 2020).

2.5. Workload transition rates

Workload was manipulated by varying the number of active UAVs, i.e., illuminated video feeds in the target detection task. This approach is consistent with previous studies, where workload transitions are manipulated by the number of tasks (e.g., Cox-Fuenzalida, 2007; Hancock et al., 1995; Matthews, 1986; Prytz and Scerbo, 2015). This approach is also consistent with the long-term goal of increasing the load per UAV operator (Arrabito et al., 2010; United States Department of Defense, 2017). Workload level thresholds, i.e., low and high, were based on pilot testing from Devlin et al. (2020) (p. 78–82). To simulate the situations likely to occur in UAV command and control, the workload transition rate was specific to transitions from low to high workload. The workload transition rates were tested via 15-min scenarios, i.e., missions in the testbed:

1. **Slow transitions**. The number of active UAVs increased steadily from low to high workload. The scenario started at low workload for 100 s, and one active UAV was added every 10 s until high workload was reached (13–16 active UAVs). The scenario would remain at high workload for 100 s, before immediately returning to low workload. The dotted black line in Fig. 3 shows this cycle repeated three times for this scenario.
2. **Medium transitions.** The number of active UAVs increased incrementally from low to high workload. The scenario started at low workload for 20 s, and then one to three active UAVs were added every 10 s until high workload was reached (13–16 active UAVs). The scenario would remain at high workload for 2 min, before immediately returning to low workload. The dashed dark gray line in Fig. 3 shows this cycle repeated five times for this scenario.
3. **Fast transitions**. The number of active UAVs in this scenario increased instantaneously from low to high workload. One minute of low workload (3–5 UAVs) was followed by 2 min of high workload (13–16 UAVs). After the 2 min of high workload, there was an immediate return to low workload. The light gray line in Fig. 3 shows this cycle repeated five times for this scenario.

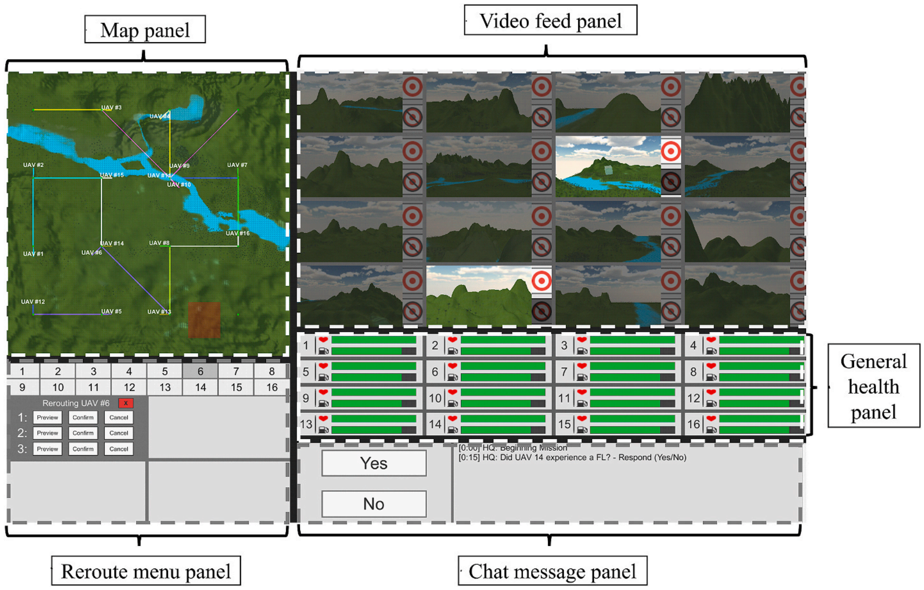


Fig. 1. Interface of the UAV testbed with panels labeled.

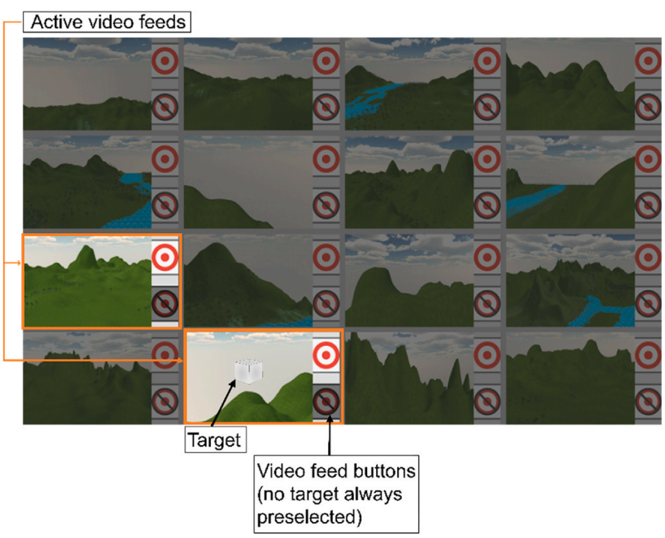


Fig. 2. Video Feed panel showing how any of the 16 video feeds could be active, i.e., video feed is illuminated. Currently, two UAVs are active; all other UAVs are inactive; The UAV in the last row has a target (semi-transparent cube).

2.6. Procedures

This research complied with the APA Code of Ethics and was approved by the Institutional Review Board at the U.S. Naval Research Laboratory. Informed consent was obtained from each participant. Participants then completed a self-paced informational training session where they completed the eye tracker's 9-point calibration procedure and learned about the overall experimental goals. They then completed a 5-min training session where 6–10 UAVs were always active. Participants then completed testbed scenarios in a counterbalanced order and calibrated their point of gaze to the eye tracker before each session.

3. Results

3.1. Preprocessing eye tracking data

Raw gaze data, which consist of the positional (x_i, y_i) and temporal information (t_i), were screened for completeness and accuracy via the data quality metric provided by Gazepoint and trials were removed if they did not reach the quality threshold (as previously mentioned). Velocity profiles were then calculated from the raw gaze points (x_i, y_i, t_i) by differentiating with a six-tap Savitzky-Golay filter of degree 2 (Krejtz et al., 2016). An I-VDT event detection algorithm was then used to determine fixations (Komogortsev and Karpov, 2013). The velocity threshold for fixations was determined by the adaptive algorithm outlined in Nyström and Holmqvist (2010), with it ranging 25.6–60.8°/s across all trials. Then, individual fixations were determined as clusters of raw gaze points that were below the trial's velocity threshold, a maximum distance of 110 pixels from each other, i.e., ~1° visual angle, and occurred within a minimum of 80 ms of each other.

3.2. Model fitting process for conditional growth curve models

The ultimate goal of this work was to identify the scan-based eye tracking metric(s) that predict performance trends over time during different workload transition rates. The five scan-based metrics from Table 1 were explored as time-invariant predictors of the best fitting *unconditional growth curve model*, which is a model that only includes the effect of time and no other predictors. There was an unconditional growth curve model for each performance metric, i.e., primary task response time and accuracy, in each transition rate (slow, medium, and fast transitions). The model fitting process followed the methodology of Barr et al. (2013) and Matuschek et al. (2017), which consisted of fitting models with a backwards selection approach. Parameters were orthogonal polynomial effects of time (in 10 s increments), with the maximal model being a random quintic time model given the results of the local polynomial regression. Parameters were removed if it did not lead to a significant decrease in model fit, which was determined by a likelihood ratio test (LRT) where $\alpha = 0.20$. R was used for all analyses (version 4.0.5; R Core Team, 2021); Response time was modeled as a general linear model and was fit with the lme4 package (Bates et al., 2015) and accuracy was modeled as a generalized linear model and fit with the GLMMadaptive package (Rizopoulos, 2021). Maximum likelihood estimation was used to fit all models.

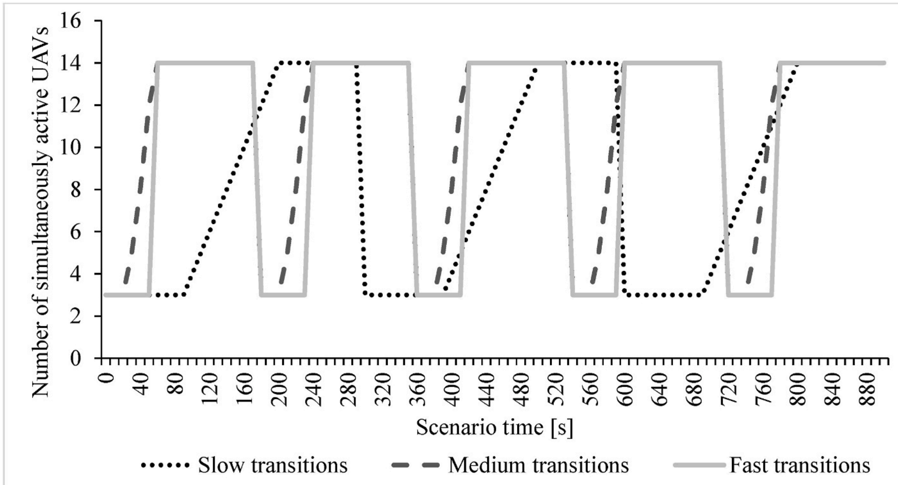


Fig. 3. Number of active UAVs during each 15-min testbed scenario, i.e., transition rate.

After the unconditional growth curve model was established, the scan-based metrics were included in the model as time-invariant predictors. A time-invariant predictor is a measure of the individual that is not expected to change over time or is only reliably measured once throughout an experiment (Hoffman, 2015). The scan-based metrics were modeled as time-invariant predictors because they were only calculated once per participant per each transition rate. The exploration of the predictive ability of scan-based metrics focused on assessing if each scan-based metric was an additive or cross-level time invariant predictor. In this work, an additive time-invariant predictor predicts average performance, i.e., the model's intercept parameter β_{0i}, (see Equation (1)), where a cross-level time-invariant predictor predicts how performance generally trends over time, i.e., the model's linear time slope, β_1 (see Equation (2)).

$$y_{ti} = \beta_{0i} + \beta_1(Time_{ti}) + \beta_2(Time_{ti})^2 + \beta_4(Time_{ti})^4 + \beta_5(Time_{ti})^5 \text{ where }, \tag{1}$$

$$\beta_{0i} = \gamma_{00} + \gamma_{01}(Scan - based\ metric_i) + U_{0i} ,$$

$$\beta_1 = \gamma_{10}$$

$$\beta_2 = \gamma_{20},$$

$$\beta_3 = \gamma_{30} ,$$

$$\beta_4 = \gamma_{40} ,$$

$$\beta_5 = \gamma_{50}$$

$$y_{ti} = \beta_{0i} + \beta_1(Time_{ti}) + \beta_2(Time_{ti})^2 + \beta_4(Time_{ti})^4 + \beta_5(Time_{ti})^5 \text{ where} \tag{2}$$

$$\beta_{0i} = \gamma_{00} + \gamma_{01}(Scan - based\ metric_i) + U_{0i} ,$$

$$\beta_1 = \gamma_{10} + \gamma_{11}(Scan - based\ metric_i)$$

$$\beta_2 = \gamma_{20},$$

$$\beta_3 = \gamma_{30} ,$$

$$\beta_4 = \gamma_{40} ,$$

$$\beta_5 = \gamma_{50}$$

Again, **R** was used for all analyses and a backwards stepwise selection process was used to assess the predictive utility of each scan-based metric in each unconditional growth curve model. This consisted of including all the eye tracking metrics, as both additive and cross-level time invariant predictors, in the established unconditional growth curve model and assessing the change in model fit when each scan-based metric was systematically excluded. The backwards stepwise selection process was then conducted with the *buildmer* package and predictor selection was based on the Akaike's Information Criterion (Matuschek et al., 2017; Voeten, 2019). When appropriate, the significance of each scan-based metric was assessed with a Wald test where degrees of freedom were corrected with the Satterthwaite's method (additive: α = 0.05 and cross-level: α = 0.10; Mathieu et al., 2012; Voeten, 2019). Scan-based metrics that were not significant were excluded, and the model was refitted and compared to the unconditional growth curve model via an LRT (α = 0.10; Gries, 2021). The results from the final conditional growth curve model, i.e., all the parameter estimates and fit statistics, are presented in detail and interpreted.

3.2.1. *For slow transitions, fixation duration predicted average response time and stationary gaze entropy predicted its trend over time*

The unconditional growth curve model for response time during slow transitions had a positive fixed quintic time slope and a random intercept ($\chi^2(1) = 46.188$, $p < 0.001$). The backward stepwise selection process selected fixation duration as an additive time-invariant predictor and stationary gaze entropy as a cross-level time-invariant predictor. These predictors were significant (See Table 2 for details) and lead to a significantly better model fit compared to the unconditional growth curve model ($\chi^2(3) = 15.959$, $p = 0.001$). Therefore, the final conditional growth curve model of response time during slow transitions included average fixation duration as an additive effect and stationary

Table 2
The parameter estimates of the final conditional growth curve models of primary task response time during slow transitions (Satterthwaite's degrees of freedom correction was used when assessing the significance of all parameter estimates).

Model parameter	Parameter estimate (standard error)
Intercept (β_{0i})	3.287*** (0.488)
Intercept (γ_{00})	0.972† (0.487)
Stationary gaze entropy (γ_{01})	−0.005* (0.002)
Average fixation duration (γ_{02})	
Linear time slope (β_1)	−3.938* (1.932)
Linear time (γ_{10})	6.778* (3.049)
Linear time × Stationary gaze entropy (γ_{11})	
Quadratic time slope (β_2)	−0.852*** (0.257)
Cubic time slope (β_3)	2.184*** (0.256)
Quartic time slope (β_4)	0.237 (0.258)
Quintic time slope (β_5)	1.775*** (0.261)

Significance codes: '***' indicates $p < 0.001$, '**' indicates $p < 0.01$, '*' indicates $p < 0.05$, '†' indicates $p < 0.10$.

gaze entropy as a cross-level effect.

Specifically, for every one standard deviation increase in fixation duration ($M = 125.2$ ms, $SD = 17.3$ ms), and assuming all other predictors remained equal, response time **improved** by 0.09 s. For stationary gaze entropy ($M = 0.63$, $SD = 0.08$), every one standard deviation **increase** estimated the total decrement in response time to be 0.20 s **longer**. Table 2 shows the results of the final conditional growth curve model and Fig. 4 depicts the final conditional growth curve model of primary task response time, specifically estimating a quintic trend over time with the scan-based metrics predicting average speed and its change over time.

3.2.2. For medium transitions, stationary gaze entropy predicted average response time

The unconditional growth curve model for response time during medium transitions had a positive fixed quintic time slope and a random intercept ($\chi^2(1) = 2.023$, $p = 0.155$). The backwards stepwise selection process selected gaze transition entropy and stationary gaze entropy as additive time-invariant predictors. However, gaze transition entropy was not a significant additive effect ($\gamma_{01} = -0.938$, $SE = 0.613$, $t = -1.531$, $p = 0.135$) so it was removed. This reduced conditional model was still a significant improvement in model fit compared to the unconditional growth curve model ($\chi^2(1) = 7.934$, $p = 0.005$), meaning stationary gaze entropy was a significant additive effect. Specifically, for every one standard deviation **increase** in stationary gaze entropy ($M = 0.60$, $SD = 0.09$), average response time would be 0.11 s **longer**. Table 3 shows the parameter estimates of the final conditional model and Fig. 5 depicts how changes in stationary gaze entropy impacted the final conditional growth curve model of primary task response time during medium transitions.

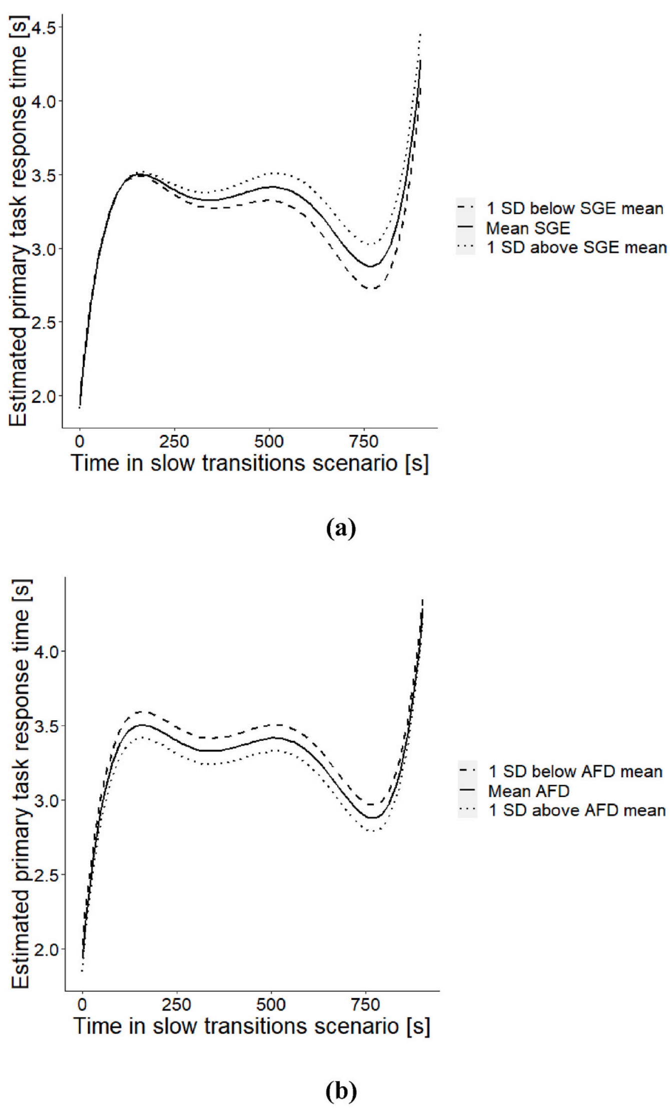


Fig. 4. The final conditional growth curve model of primary task response time during slow transitions, specifically showing the isolated impact of each predictive scan-based metric when it is at its mean and one standard deviation (SD) above and below it: (a) stationary gaze entropy (SGE) and (b) average fixation duration (AFD).

3.2.3. For fast transitions, none of the scan-based metrics predicted response time trends

The unconditional growth curve model for response time during fast transitions had a positive fixed quintic time slope and a random intercept ($\chi^2(1) = 16.01$, $p < 0.001$). The backwards stepwise selection process selected spatial gaze density as an additive and cross-level time-invariant predictor. However, neither was a significant effect in the model ($\gamma_{01} = -0.340$, $SE = 0.483$, $z = -0.705$, $p = 0.487$ and $\gamma_{11} = 5.761$, $SE = 3.574$, $z = 1.612$, $p = 0.107$), so both were removed. Therefore, there was no final conditional growth curve model for response time during fast transitions.

3.2.4. For slow transitions, stationary gaze entropy, average fixation duration, and spatial gaze density predicted average accuracy

The unconditional growth curve model for accuracy during slow transitions was a fixed quintic, random cubic time slope model, excluding all random correlation ($\chi^2(1) = 1.670$, $p = 0.196$). The backwards stepwise selection process selected stationary gaze entropy as an additive time-invariant predictor and average fixation duration, spatial gaze density, and gaze transition rate as cross-level time-invariant predictors. However, the additive time-invariant predictor of gaze transition rate and none of the cross-level time-invariant predictors were significant, so they were removed ($\gamma_{04} = 0.147$, $SE = 0.085$, $z = 1.732$, $p = 0.083$; $\gamma_{11} = 0.616$, $SE = 0.408$, $z = 1.510$, $p = 0.131$; $\gamma_{12} = 9.950$, $SE = 6.092$, $z = 1.633$, $p = 0.102$; $\gamma_{13} = -1.149$, $SE = 0.752$, $z = -1.528$, $p = 0.126$). This reduced conditional model significantly improved model fit compared to the unconditional model ($\chi^2(3) = 14.060$, $p = 0.003$). Therefore, the final conditional growth curve model of accuracy during slow transitions included stationary gaze entropy, average fixation duration, and spatial gaze density as additive predictors.

Specifically, for every one standard deviation **increase** in stationary gaze entropy ($M = 0.63$, $SD = 0.08$), and assuming all other predictors remained equal, average accuracy was predicted to **worsen** by an average of 4.9%. However, for average fixation duration ($M = 125.0$ ms, $SD = 17.1$ ms), every standard deviation **increase** predicted average accuracy to **improve** by an average of 3.7%, assuming all other predictors remained equal. For spatial gaze density, every standard

Table 3
The parameter estimates of the final conditional growth curve models of primary task response time during medium transitions (Satterthwaite's degrees of freedom correction was used when assessing the significance of all parameter estimates).

Model parameter	Parameter estimate (standard error)
Intercept (β_{0i})	3.083*** (0.259)
Intercept (γ_{00})	1.267** (0.425)
Stationary gaze entropy (γ_{01})	
Linear time slope (β_1)	0.217 (0.245)
Quadratic time slope (β_2)	0.688** (0.247)
Cubic time slope (β_3)	0.291 (0.246)
Quartic time slope (β_4)	−0.760** (0.246)
Quintic time slope (β_5)	0.350 (0.243)

Significance codes: '***' indicates $p < 0.001$, '**' indicates $p < 0.01$, '*' indicates $p < 0.05$, '†' indicates $p < 0.10$.

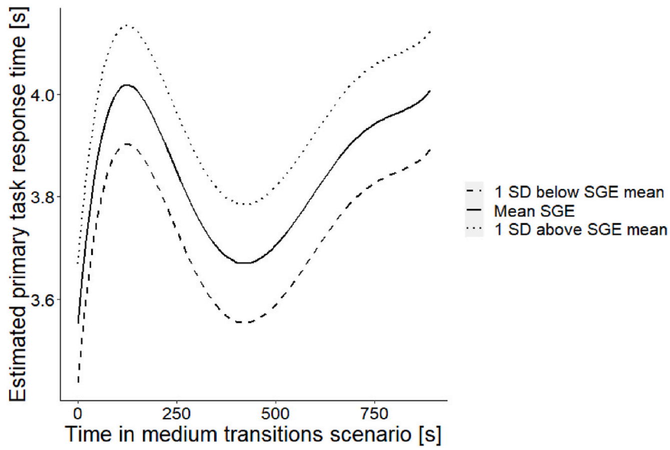


Fig. 5. The final conditional growth curve model of primary task response time during medium transitions, specifically showing the impact of stationary gaze entropy (SGE) when it is at its mean and one standard deviation (SD) above and below it.

deviation **increase** predicted average accuracy to **decline** by an average of 3.8%, assuming all other predictors remained equal. Table 4 shows the parameter estimates of the final conditional model and Fig. 6 depicts how each scan-based metric impacted the final conditional growth curve model of primary task accuracy during slow transitions.

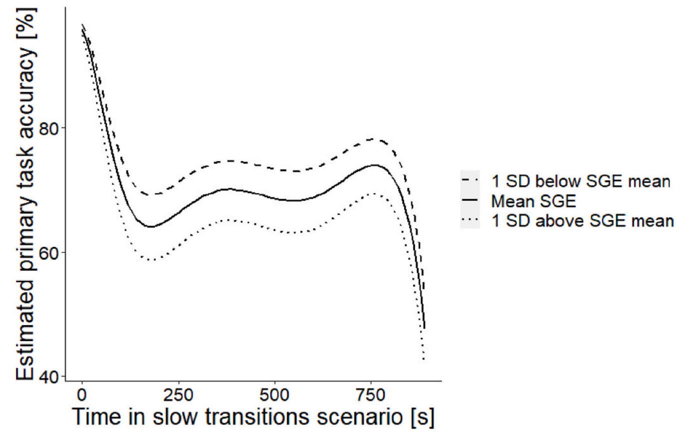
3.2.5. *For medium transitions, stationary gaze entropy predicted average accuracy*

The unconditional growth curve model for accuracy during medium transitions was a model with a fixed quintic and random intercept ($\chi^2(1) = 45.98, p < 0.0001$). The backwards stepwise selection process selected stationary gaze entropy as an additive time-invariant predictor and spatial gaze density and gaze transition rate as cross-level time-invariant predictors. However, none of the predictors other than stationary gaze entropy were significant, so they were removed ($\gamma_{02} = -0.499$, $SE = 0.446$, $z = -1.120$, $p = 0.262$; $\gamma_{03} = 0.083$, $SE = 0.059$, $z = 1.411$, $p = 0.158$; $\gamma_{11} = -5.822$, $SE = 4.194$, $z = -1.388$, $p = 0.165$; $\gamma_{12} = 0.774$, $SE = 0.490$, $z = 1.581$, $p = 0.114$). Including only stationary gaze entropy still lead to a significantly better model fit compared to the unconditional model ($\chi^2(1) = 11.57, p < 0.001$). Therefore, the final conditional growth curve model of accuracy during medium transitions included stationary gaze entropy as an additive effect. Specifically, for every one standard deviation **increase** in stationary gaze entropy ($M = 0.60$, $SD = 0.09$), average accuracy was predicted to **decline** by an average of 5.5% during the medium transitions scenario. Table 5 shows the parameter estimates of the final conditional growth curve model and Fig. 7 depicts how stationary gaze entropy impacted the final conditional growth curve model of primary task accuracy during medium transitions.

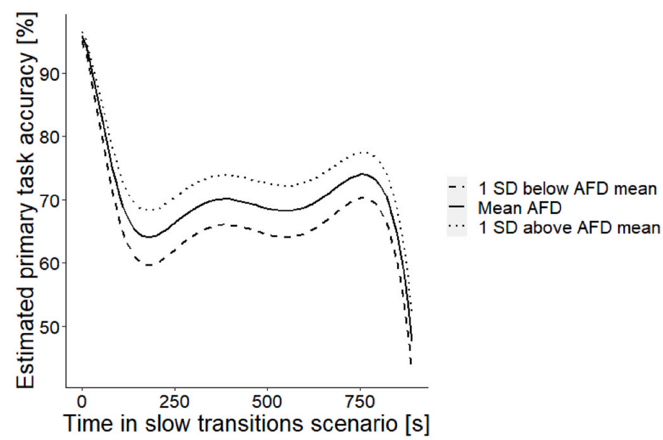
Table 4
The parameter estimates of the final conditional growth curve models of primary task accuracy during slow transitions.

Model parameter	Parameter estimate (standard error)
Intercept (β_{0i})	4.324*** (0.411)
Intercept (γ_{00})	−2.753*** (0.376)
Stationary gaze entropy (γ_{01})	0.187*** (0.037)
Average fixation duration (γ_{02})	−2.618*** (0.491)
Spatial gaze density (γ_{03})	
Linear time slope (β_1)	−1.533*** (0.361)
Quadratic time slope (β_2)	1.483*** (0.378)
Cubic time slope (β_3)	−2.903*** (0.364)
Quartic time slope (β_4)	0.906** (0.352)
Quintic time slope (β_5)	−1.858*** (0.320)

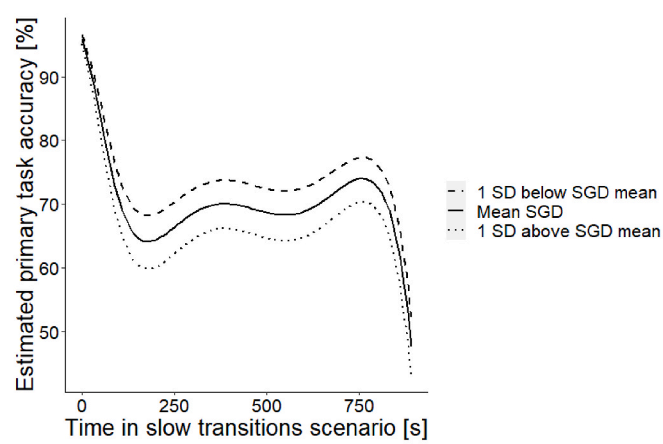
Significance codes: '***' indicates $p < 0.001$, '**' indicates $p < 0.01$, '*' indicates $p < 0.05$, '†' indicates $p < 0.10$.



(a)



(b)



(c)

Fig. 6. The final conditional growth curve model of primary task accuracy during slow transitions, specifically showing the isolated impact of each predictive scan-based metric when it is at its mean and one standard deviation (SD) above and below it: (a) stationary gaze entropy (SGE), (b) average fixation duration (AFD), and (c) spatial gaze density (SGD).

Table 5
The parameter estimates of the final conditional growth curve models of primary task accuracy during medium transitions.

Model parameter	Parameter estimate (standard error)
Intercept (β_{0i})	2.512*** (0.084)
Intercept (γ_{00})	−2.780*** (0.132)
Stationary gaze entropy (γ_{01})	
Linear time slope (β_1)	1.106*** (0.296)
Quadratic time slope (β_2)	0.385* (0.186)
Cubic time slope (β_3)	−0.818** (0.273)
Quartic time slope (β_4)	−0.025 (0.207)
Quintic time slope (β_5)	−1.722*** (0.191)

Significance codes: '***' indicates $p < 0.001$, '**' indicates $p < 0.01$, '*' indicates $p < 0.05$, '†' indicates $p < 0.10$.

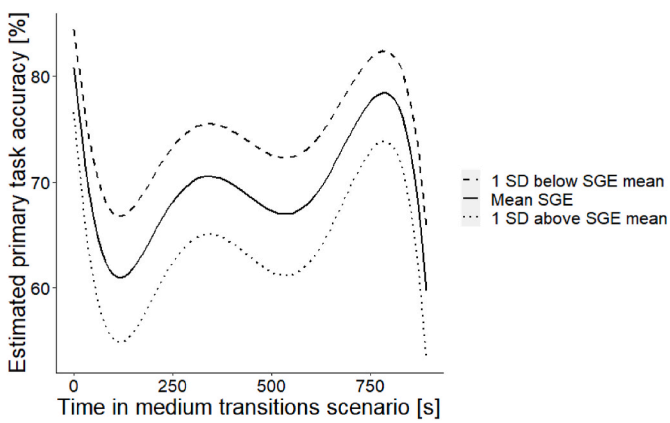


Fig. 7. The final conditional growth curve model of primary task accuracy during medium transitions, specifically showing the impact of stationary gaze entropy (SGE) when it is at its mean and one standard deviation (SD) above and below it.

3.2.6. *For fast transitions, stationary gaze entropy and average fixation duration predicted average accuracy and gaze transition rate predicted its trend over time*

The unconditional growth curve model for accuracy during fast transitions was a random quintic time model ($\chi^2(1) = 2.35, p = 0.1251$). The backwards stepwise selection process selected gaze transition entropy and average fixation duration as additive time-invariant predictors and stationary gaze entropy, spatial gaze density, and gaze transition rate as cross-level time-invariant predictors. However, gaze transition entropy, spatial gaze density, and the cross-level effect of stationary gaze entropy were not significant so they were removed ($\gamma_{01} = 1.151$, $SE = 0.655$, $z = 1.756$, $p = 0.079$; $\gamma_{04} = 0.118$, $SE = 0.063$, $z = 1.872$, $p = 0.061$; $\gamma_{11} = -1.497$, $SE = 3.563$, $z = -0.420$, $p = 0.674$; $\gamma_{12} = -0.491$, $SE = 0.512$, $z = -0.959$, $p = 0.338$). The reduced conditional model was still a significantly better fit than the unconditional model ($\chi^2(4) = 7.91, p = 0.095$). Therefore, the final conditional growth curve model for accuracy during fast transitions included average fixation duration and stationary gaze entropy as an additive predictor and gaze transition rate as cross-level predictor.

Specifically, for every one standard deviation **increase** in stationary gaze entropy ($M = 0.61$, $SD = 0.08$) average accuracy was predicted to **decline** by 4.1% during fast transitions, assuming all other predictors remained equal. For average fixation duration (M = 124.3 ms, SD = 16.5 ms), one standard deviation increase predicted average accuracy to **improve** by 3.6%, assuming all other predictors remained equal. Finally, a one standard deviation increase in gaze transition rate (M = 2.7 grid cells/s, SD = 0.71 grid cells/s) predicted the **decline** in accuracy to be 0.64% **less** on average, assuming all other predictors remained equal. Table 6 shows the parameter estimates of the final conditional model and Fig. 8 depicts how the scan-based metrics impacted the final conditional growth curve model of primary task accuracy during fast transitions.

Table 6
The parameter estimates of the final conditional growth curve models of primary task accuracy during fast transitions.

Model parameter	Parameter estimate (standard error)
Intercept (β_{0i})	2.704*** (0.286)
Intercept (γ_{00})	−2.315*** (0.384)
Stationary gaze entropy (γ_{01})	0.169*** (0.040)
Average fixation duration (γ_{02})	−0.215*** (0.053)
Gaze transition rate (γ_{03})	
Linear time slope (β_1)	−2.978** (1.025)
Linear time (γ_{10})	0.750* (0.367)
Linear time × Gaze transition rate (γ_{11})	
Quadratic time slope (β_2)	0.683** (0.312)
Cubic time slope (β_3)	−1.907*** (0.319)
Quartic time slope (β_4)	2.001*** (0.315)
Quintic time slope (β_5)	−1.161*** (0.290)

Significance codes: '***' indicates $p < 0.001$, '**' indicates $p < 0.01$, '*' indicates $p < 0.05$, '†' indicates $p < 0.10$.

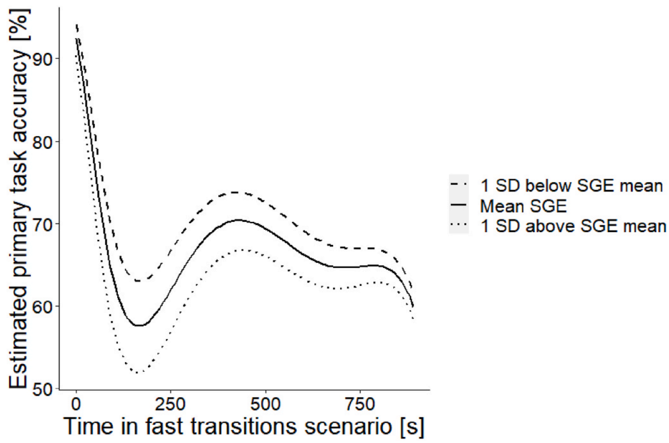
Table 7 summarizes the findings on the ability scan-based metrics have on predicting workload transition performance trends over time.

4. Discussion

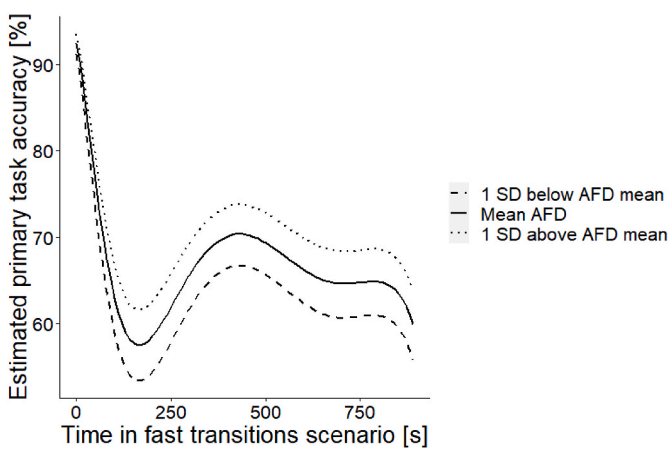
The goal of this research was to assess whether scan-based metrics are predictive of performance trends. Here we found the predictive capability was a function of the performance metric, i.e., response time and accuracy, and workload transition rate. For slow transitions, the duration of visual attention (average fixation duration) and how dispersed its transitions are across the AOIs (stationary gaze entropy) and in general (spatial gaze density) are predictive of both average performance and/or its trends over time. For medium transitions, the dispersion of visual attention transitions across the AOIs (stationary gaze entropy) was predictive of average performance. For fast transitions, the dispersion of visual attention transitions across the AOIs (stationary gaze entropy) and its duration (average fixation duration) is predictive of average accuracy whereas the pace of general visual attention changes (gaze transition rate) is predictive of accuracy's trends over time. Our expectations were mostly met, but with caveats. For instance, across all transition rates, stationary gaze entropy was predictive of average performance and/or its trend over time, which is consistent with other work that studies scan-based metrics in realistic environments (Moacdieh et al., 2020; Shiferaw et al., 2019a,b).

4.1. *Implications of stationary gaze entropy being a significant predictor across all transition rates*

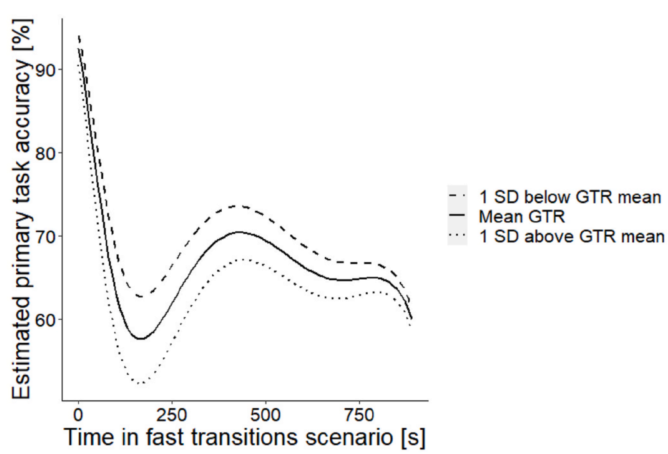
The ability of stationary gaze entropy to predict performance may lie in the inclusion of context-driven AOIs and the Markov property—i.e., *memoryless* transitions. The Markov property is true if, "the probability distribution of future states of the process conditioned on both the past and present states depends only on the present state" (Gudivada et al., 2015). For example, transitions to AOIs on the display is only dependent on the current AOI the participant is looking at. Here the long-term probabilities are the proportion of transitions that go to each AOI (Shiferaw et al., 2019a,b). Stationary gaze entropy may have high predictive capability because it is a measuring spread based on *active* transitions between AOIs (Shic et al., 2008) and the certainty of those transitions. Thus, providing a single quantitative value on the dynamics of visual attention transitions across AOIs, given both micro, i.e., where current visual attention is transitioning to, and macro, i.e., how the proportion of visual attention to each AOI compare to each other over time. Additionally, the current success of assuming transitions between AOIs are a memoryless process—i.e., the Markov property—shows practical potential in deploying effective adaptive assistance based on scan patterns.



(a)



(b)



(c)

Fig. 8. The final conditional growth curve model of primary task accuracy during fast transitions, specifically showing the isolated impact of each predictive scan-based metric when it is at its mean and one standard deviation (SD) above and below it: (a) stationary gaze entropy (SGE), (b) average fixation duration (AFD), and (c) gaze transition rate (GTR).

The eye tracker can intermittently lose contact with the operator's point of gaze, so it is imperative adaptive assistance can provide accurate predictions without requiring the entire scanpath. Vetting this approach is essential as its applicability may change depending on AOI definition or task paradigm.

The overwhelming predictive capability of stationary gaze entropy may not only be due to the Markov property because gaze transition entropy, which also relies on the Markov property, was not effective in predicting performance (similar to Shiferaw et al., 2018). Rather, it may also be due to it being a spread metric, i.e., measuring where someone looks, as gaze transition entropy is a directness metric, i.e., measuring the efficiency of a scanpath. However, between the two spread metrics—i.e., spatial gaze density and stationary gaze entropy—stationary gaze entropy predicted more aspects of performance trends, meaning the way spread metrics are defined matters. Spatial gaze density may have been a somewhat less effective predictor because it reduces the display into a uniform grid, providing no context on how each grid cell relates to the task or how often it is viewed. Conversely, the AOIs of stationary gaze entropy had a direct mapping with a testbed task, which inherently provides more semantic information than a grid cell. Specifically, it informs how participants relied on their visual attention to manage tasks as workload transitioned. Simply knowing how much of the display was viewed, i.e., what spatial gaze density essentially quantifies, could be more dependent on display design than the participant's visual attention allocation patterns (Moacdieh and Sarter, 2015), making it less informative of the individual's workload transition management strategy. In summary, where *and* how frequent visual attention transitioned is a better indicator of performance than where it landed in general.

Finally, stationary gaze entropy is not only useful as a predictor of performance, but also informs display design. Stationary gaze entropy suggests having a balanced number of transitions between the AOIs results in worse performance across all transition rates. Given this study had a primary task, we conjecture that performance improved when most of the transitions were to the primary task's AOI, i.e., Video Feed panel. From a design perspective, attention should be directed to a primary task and minimized elsewhere either via design features and/or adaptive assistance. When reviewing suggested layouts of current UAV command and control tasks, we found that AOIs associated with each task are typically dispersed across the display and organized into several subgroups (Feitshans et al., 2008; Foroughi et al., 2019). It may be strategic to reorganize these displays based on the operator's priorities, especially if multitasking between tasks is not equal.

4.2. Implications on theory: effort regulation may manifest differently depending on sensory modality and depend on the features of the transition and environment

The interpretations from the final conditional models add to the existing theory on workload transitions. To date, there has been limited work to expand upon the two explanations—i.e., resource depletion and effort regulation, even when psychophysiological measures are included (e.g., Bowers et al., 2014; Boyer et al., 2015). Here, the current findings, allow us a unique opportunity to build upon the effort regulation explanation, given it stipulates that workload transition performance is a function of how mental resources are relied upon. Specifically it states:

Workload transition performance is dependent on the individual actively **appraising**, recruiting, and **deploying** the requisite amount of **mental resources**. Performance is stable as long as the appraisal is correct and workload does not reach levels of overload (Hockey, 1997).

We expand upon the bolded terms in the definition above by adding specifications to: (a) *what* in the environment is appraised, and (b) *how* mental resources are deployed, and (c) the *type* of mental resource. The only aspect of this explanation we do not address is resource recruitment, but future work ideally can (see alternatives presented in Conclusion).

Table 7
Summary of the conditional growth curve modeling results. Significant eye tracking metrics are denoted by a ✓

Transition rate	Performance metric	*Spread metric*		*Directness metric*		*Duration metric*	Takeaway
		Spatial gaze density	Stationary gaze entropy	Gaze transition rate	Gaze transition entropy	Fixation duration	
SLOW	*Response time*		✓			✓	**Longer** *fixation duration* predicts **faster** response times overall. **Larger** *stationary gaze entropy* predicts **larger decrements** in response time over time
	Accuracy	✓	✓			✓	**Larger** *stationary gaze entropy* and *spatial gaze density* predicts **lower** accuracy rates overall. **Longer** *fixation duration* predicts **higher** accuracy rates overall
MEDIUM	*Response time*		✓				**Larger** *stationary gaze entropy* predicts **slower** overall response times
	Accuracy		✓				**Larger** *stationary gaze entropy* predicts **lower** overall accuracy rates
FAST	*Response time*						None of the scan-based metrics predict response time trends
	Accuracy		✓	✓		✓	**Larger** *stationary gaze entropy* predicts **lower** accuracy overall. **Longer** *fixation durations* predicts **higher** accuracy overall. **Larger** *gaze transition rate* predicts **smaller decrements** in accuracy over time

First, we propose workload transition performance is not only dependent on appraising the amount of workload, but it is also dependent on appraising the rate in which workload transitions. The different significant predictors across the final conditional models support this notion. Although all final models predicted performance to worsen when visual attention transitions were more evenly distributed across tasks, i.e., when stationary gaze entropy was larger, there is some evidence that different scan patterns are necessary across the three transition rates. For example, for slow transitions, longer periods of visual attention predicted improved average response time, meaning performing quickly during slow transitions required more thoughtful cognitive processing (Holmqvist et al., 2011; Poole and Ball, 2006). For fast transitions, frequent general attention shifts predicted improved average accuracy and a smaller decrement over time, suggesting performing well during fast transitions requires more frequent changes in attention (Yang et al., 2018). Our results show performance improves when scan patterns account for transition rate.

Second, we propose the deployment of resources should consider the way in which visual attention is allocated. Namely, its location, efficiency, and time span. This is evident by one scan-based eye tracking metric from each category—i.e., *spread*, *directness*, and *duration*—being a significant predictor for at least one performance trend over time. We propose the deployment of mental resources needs to consider the questions regarding the visual attention allocation types: where people are looking (*spread*), how efficiently people are looking (*directness*), and how long people are looking (*duration*).

Third, we propose amending the effort regulation explanation to specify that workload transition performance is dependent on the type of mental resource. Our work suggests it is specifically dependent on *visual attentional* resources. This is supported by the fact that at least one scan-based metric was a significant predictor of performance. Considering the entire experiment was largely visual in nature, this finding is not surprising; however, it highlights that the applicability of this theory may hinge on what sensory modality is considered. This supports the premise of the Multiple Resource Model, which posits different sensory modalities draw from separate attentional resources (Wickens, 1980). Future work should examine whether the effort regulation explanation is applicable to other modalities, i.e., auditory, tactile, etc. We believe detailing the effort regulation explanation in this way leads to a better understanding and expectation of how people perform and allocate their visual attention during workload transitions.

4.3. Implication for design: technology design for each transition rate based on scan-based metrics

Workload transitions continue to be a feature of the environment that needs consideration, but currently there is limited design guidance that accounts for them. Beyond prediction, the current results also provide new information on the success of task strategy. Empirically-based design guidance for each workload transition rate is outlined in Table 8. Although it is by no means exhaustive, as different populations, transition rates, and environments need to be sampled, this design guidance is the first of its kind in the workload transition knowledgebase.

5. Conclusion

The present results suggest scan-based metrics are capable of predicting workload transition performance trends. They are also very informative on differences in performance trends across transitions rates, theory development, and design guidance. Future work should continue to explore how novel measures and methods can test and revise theory surrounding workload transitions in order to innovate the current state of the knowledgebase. This work also finds another application of scan patterns potentially being informative of performance in real-time. Eye tracking methods need to be included in future workload transition research, especially if it pertains to workload transition performance over time.

5.1. Future work and limitations

Although the present work substantially adds to the workload transition knowledgebase, it is not without limitations. First, although the selection of eye tracking metrics included in this study was motivated by

Table 8
Findings, design guidance, and design example for each transition rate.

Transition rate	Finding	Design Guidance	Example
ALL	• Stationary gaze entropy predicted performance to be worse when attention transitions across tasks were relatively equal	• Conduct a task analysis for the content and placement of AOIs o Minimize transitions between AOIs and/or centralize the most used AOIs o Offload less prioritized tasks to other sensory channels (e.g., chat messages to auditory, fuel leaks to tactile; Riggs et al., 2017)	• Alerting participants of potential threats via the tactile channel improved primary task performance without hindering secondary task performance in a simulated combat environment (Oskarsson et al., 2012)
SLOW	• Longer fixation duration and smaller spatial gaze density predicted better average performance	• Design elements on the display so that they prompt the operator to take the time to further examine and encode information and make sure the elements are collocated • Make items essential to good performance, (e.g., the target), engaging and informative so the operator spends time fixating (Jacob and Karn, 2003; Poole and Ball, 2006)	• Increasing the detail and salience of information without increasing clutter helped refocus attention without a cost of cognitive load in nuclear power plant control environments (Kovesdi et al., 2018)
MEDIUM	• Stationary gaze entropy is the only metric able to predict response time and accuracy trends, so it should be further explored to be used in real-time	• Build and test stochastic models of transitions between AOIs, i.e., Markov decision processes, to identify the most to least frequent transitions between AOIs to determine the distance between AOIs	• Gaze location in static image viewing was predicted with 56% accuracy (chance was 33%) with hidden Markov modeling (Coutrot et al., 2018)
FAST	• Larger gaze transition rate and larger fixation duration predicted better average performance and its trend over time	• Prompt efficient scanning by decluttering the display and making key tasks salient and informative, as this will also increase the focus of attention (Moacdieh and Sarter, 2015) • Provide redundancy for the most important tasks by including multiple, informative visual representations in the environment (Yang et al., 2018)	• Comprehension rates and visual attention transition rates increased when information was presented across multiple visualizations (O'Keefe et al., 2014)

existing work (Moacdieh et al., 2020), there are other metrics worth exploring (e.g., Devlin et al., 2021). Several scan-based metrics were selected in the backwards stepwise selection process, so their usefulness may change based on research goals or modeling methods [e.g., the applicability of spatial gaze density in the present work versus Moacdieh et al. (2020)]. In addition, eye tracking metrics used to study cognitive load (e.g., pupillometry, blink rate) may be better equipped to address any applicability of the resource depletion explanation and/or the recruitment of mental resources due to these metrics' ability to specifically address amounts of mental resources. Similarly, psychophysiological measures of task engagement and/or fatigue could help further detail the effort regulation explanation (e.g., Bernhardt et al., 2019; Naeeri et al., 2019). More generally, different types of workload transitions should be explored with this current analysis approach, especially when considering the prevalence and impact of unexpected and dramatic changes in workload (Endsley, 2017).

Relatedly, multivariate growth curve modeling could explore if the trend of scan-based metrics relates to workload transition performance. It would require a rather large sample size (>100 participants; Astivia et al., 2019), but it may be particularly informative when building upon the resource depletion explanation and/or the recruitment of resources, as both depend on time and workload. The current eye tracking metrics were included as time-invariant predictors because of (1) how these metrics are calculated and (2) the ultimate goal to understand the predictive capability of eye tracking on performance trends over time.

Future work should see if Markov models could reliably predict visual attention transitions across AOIs during workload transitions. Previous work has shown some success in using Markov models to predict where visual attention will be allocated (Ebeid et al., 2019; Liechty et al., 2003), but it is unclear how transition rate, individual differences, and task features would influence these findings. It is also worth expanding this kind of investigation to account for the impact of the secondary task and the impact it has on performance for different workload transitions rates. Examining this has been particularly helpful in understanding multitasking strategies in complex environments (Jeong et al., 2019), which is essential to future UAV command and control missions (Cummings, 2014).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank the Command Decision Making program, within the Office of Naval Research, for funding support (award number: N0001421WX00798). We would like to thank Audrey Moore from the U. S. Naval Aerospace Medical Institute for assisting in the data collection and Jad Atweh, Mohamad El-Iskandarani, and Samuel Smith from the University of Virginia for assisting in the eye tracking data analysis.

References

Abich, J., Reinerman-Jones, L., Matthews, G., 2017. Impact of three task demand factors on simulated unmanned system intelligence, surveillance, and reconnaissance operations. Ergonomics 60 (6), 791–809. https://doi.org/10.1080/00140139.2016.1216171.

Arrabito, G.R., Ho, G., Lambert, A., Rutley, M., Keillor, J., Chiu, A., AU, H., Hou, M., 2010. Human factors issues for controlling uninhabited aerial vehicles: Preliminary Findings in support of the Canadian forces joint unmanned aerial vehicle surveillance target acquisition system project. www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA543186.

Astivia, O.L.O., Gadermann, A., Guhn, M., 2019. The relationship between statistical power and predictor distribution in multilevel logistic regression: a simulation-based approach. BMC Med. Res. Methodol. 19 (97) https://doi.org/10.1186/s12874-019-0742-8.

Ayasse, N.D., Wingfield, A., 2020. The two sides of linguistic context: eye-tracking as a measure of semantic competition in spoken word recognition among younger and older adults. Front. Hum. Neurosci. 14 https://doi.org/10.3389/fnhum.2020.00132.

Barr, D.J., 2008. Analyzing "visual world" eyetracking data using multilevel logistic regression. J. Mem. Lang. 59 (4), 457–474. https://doi.org/10.1016/j.jml.2007.09.002.

Barr, D.J., Levy, R., Scheepers, C., Tily, H.J., 2013. Random effects structure for confirmatory hypothesis testing: keep it maximal. J. Mem. Lang. 68 (3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001.

Barz, M., Kapp, S., Kuhn, J., Sonntag, D., 2021. Automatic recognition and augmentation of attended objects in real-time using eye tracking and a head-mounted display. In: ACM Symposium on Eye Tracking Research and Applications. https://doi.org/10.1145/3450341.3458766, 1–4.

Bates, D., Mächler, M., Bolker, B.M., Walker, S.C., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Software 67 (1). https://doi.org/10.18637/jss.v067.i01.

Bernhardt, K.A., Poltavski, D., Petros, T., Ferraro, F.R., Jorgenson, T., Carlson, C., Drechsel, P., Iseminger, C., 2019. The effects of dynamic workload and experience on commercially available EEG cognitive state metrics in a high-fidelity air traffic control environment. Appl. Ergon. 77, 83–91. https://doi.org/10.1016/j.apergo.2019.01.008.

Bowers, M.A., Christensen, J.C., Eggemeier, F.T., 2014. The effects of workload transitions in a multitasking environment. In: Proceedings of the Human Factors and Ergonomics Society. SAGE Publishing, 220–224.

Boyer, M., Cummings, M.L., Spence, L.B., Solovey, E.T., 2015. Investigating mental workload changes in a long duration supervisory control task. Interact. Comput. 27 (5), 512–520. https://doi.org/10.1093/iwc/iwv012.

Cerruti, L.M., Estepp, J.R., Miller, W.D., Christensen, J.C., 2010. Transcranial Doppler Assessment of Workload Transition in a Complex Task. In: Proceedings of the Human Factors and Ergonomics Society, vol. 1. Elsevier, pp. 244–248. https://doi.org/10.1518/107118110X12829369200837.

Coral, M.P., 2016. Analyzing Cognitive Workload through Eye-Related Measurements: A Meta-Analysis. Wright State University.

Coutrot, A., Hsiao, J.H., Chan, A.B., 2018. Scanpath modeling and classification with hidden Markov models. Behav. Res. Methods 50, 362–379. https://doi.org/10.3758/s13428-017-0876-8.

Cox-Fuenzalida, L.-E., 2007. Effect of workload history on task performance. Hum. Factors 49 (2), 277–291. https://doi.org/10.1518/001872007X312496.

Cox-Fuenzalida, L.-E., Angie, A., Holloway, S., Sohl, L., 2006. Extraversion and task performance: a fresh look through the workload history lens. J. Res. Pers. 40 (4), 432–439. https://doi.org/10.1016/j.jrp.2005.02.003.

Cox-Fuenzalida, L.-E., Swickert, R., Hittner, J.B., 2004. Effects of neuroticism and workload history on performance. Pers. Indiv. Differ. 36 (2), 447–456.

Cummings, M., 2014. Man versus machine or man + machine? IEEE Intell. Syst. 29 (5), 62–69. www.computer.org/intelligent.

Curran, P.J., Obeidat, K., Losardo, D., 2010. Twelve frequently asked questions about growth curve modeling. J. Cognit. Dev. 11 (2), 121–136. https://doi.org/10.1080/15248371003699969.

Devlin, S.P., Riggs, S.L., 2018. The effect of video game experience and the ability to handle workload and workload transitions. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 736–740. https://doi.org/10.1177/1541931218621167.

Devlin, S.P., Moacdieh, N.M., Wickens, C.D., Riggs, S.L., 2020. Transitions between low and high levels of mental workload can improve multitasking performance. IISE. Trans.Occup. Ergon.Hum. Factors 8 (2), 72–87. https://doi.org/10.1080/24725838.2020.1770898.

Devlin, S.P., Byham, J.K., Riggs, S.L.U., 2021. Does what we see shape history? Examining workload history as a function of performance and ambient/focal visual attention. Trans. Appl. Percept. 8 (2), 1–17. https://doi.org/10.1145/3449066.

Duchowski, A.T., 2017. Eye Tracking Methodology: Theory and Practice, third ed. Springer International Publishing AG.

Ebeid, I.A., Gwizdka, J., 2018. Real-time gaze transition entropy. In: Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications. https://doi.org/10.1145/3204493.3208340, 1–3.

Ebeid, I.A., Bhattacharya, N., Gwizdka, J., Sarkar, A., 2019. Analyzing gaze transition behavior using Bayesian mixed effects Markov models. In: Eye Tracking Research and Applications Symposium (ETRA). https://doi.org/10.1145/3314111.3319839.

Endsley, M.R., 2017. From here to autonomy: lessons learned from human-automation research. Hum. Factors 56 (1), 5–27. https://doi.org/10.1177/0018720816681350.

Fairclough, S.H., 2009. Fundamentals of physiological computing. Interact. Comput. 21 (1–2), 133–145. https://doi.org/10.1016/j.intcom.2008.10.011.

Feitshans, G., Rowe, A., Davis, J., Holland, M., Berger, L., 2008. Vigilant spirit control station (VSCS) the face of counter. In: AIAA Guidance, Navigation and Control Conference and Exhibit. https://doi.org/10.2514/6.2008-6309.

Foroughi, C.K., Sibley, C., Brown, N.L., Rovira, E., Pak, R., Coyne, J.T., 2019. Detecting automation failures in a simulated supervisory control environment. Ergonomics 62 (9), 1150–1161. https://doi.org/10.1080/00140139.2019.1629639.

Gluckman, J.P., Warm, J.S., Dember, W.N., Rosa, R.R., 1993. Demand transitions and sustained attention. J. Gen. Psychol. 120 (3), 323–337. https://doi.org/10.1080/00221309.1993.9711151.

Godfroid, A., Ahn, J., Choi, I., Ballard, L., Cui, Y., Johnston, S., Lee, S., Sarkar, A., Yoon, H.J., 2018. Incidental vocabulary learning in a natural reading context: an eye-tracking study. Bilingualism 21 (3), 563–584. https://doi.org/10.1017/S1366728917000219.

Goldberg, J.H., Kotval, X.P., 1999. Computer interface evaluation using eye movements: methods and constructs. Int. J. Ind. Ergon. 24 (1), 631–645. https://doi.org/10.1016/S0169-8141(98)00068-7.

Gries, S.T., 2021. (Generalized linear) mixed-effects modeling: a learner corpus example. Lang. Learn. 71 (3), 757–798. https://doi.org/10.1111/lang.12448.

Gudivada, V.N., Rao, D., Raghavan, V.V., 2015. Chapter 9: big data driven natural language processing research and applications. In: Govindaraju, V., Raghavan, V.V., Rao, C.R. (Eds.), Handbook of Statistics, vol. 33. Elsevier, pp. 203–238.

Hancock, P.A., Williams, G., Manning, C.M., Miyake, S., 1995. Influence of task demand characteristics on workload and performance. Int. J. Aviat. Psychol. 5 (1), 63–86. https://doi.org/10.1207/s15327108ijap0501_5.

Helton, W.S., Shaw, T., Warm, J.S., Matthews, G., Hancock, P., 2008. Effects of warned and unwarned demand transitions on vigilance performance and stress. Hist. Philos. Logic 21 (April), 173–184. https://doi.org/10.1080/10615800801911305.

Hergeth, S., Lorenz, L., Vilimek, R., Krems, J.F., 2016. Keep your scanners peeled: gaze behavior as a measure of automation trust during highly automated driving. Hum. Factors 58 (3), 509–519. https://doi.org/10.1177/0018720815625744.

Hockey, G.R.J., 1997. Compensatory control in the regulation of human performance under stress and high workload: a cognitive energetical framework. Biol. Psychol. 45, 73–93. https://doi.org/10.1016/S0301-0511(96)05223-4.

Hoffman, L., 2015. Longitudinal Analysis: Modeling Within-Person Fluctuation and Change. Routledge.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., van de Weijer, J., 2011. Eye Tracking: A Comprehensive Guide to Methods, Paradigms, and Measures. Oxford University Press.

Hooey, B.L., Kaber, D.B., Adams, J.A., Fong, T.W., Gore, B.F., 2017. The underpinnings of workload in unmanned vehicle systems. IEEE.Trans.Hum.Mach. Syst. 48 (5), 452–467. https://doi.org/10.1109/TCST.2007.894653.

Huey, B.M., Wickens, C.D. (Eds.), 1993. Workload Transition: Implications for Individual and Team Performance. National Academy Press.

Jacob, R.J.K., Karn, K.S., 2003. Eye tracking in human–computer interaction and usability research: ready to deliver the promises. In: The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research, pp. 573–607.

Jansen, R.J., Sawyer, B.D., van Egmond, R., de Ridder, H., Hancock, P.A., 2016. Hysteresis in mental workload and task performance: the influence of demand transitions and task prioritization. Hum. Factors 58 (8), 1143–1157. https://doi.org/10.1177/0018720816669271.

Jarodzka, H., Scheiter, K., Gerjets, P., van Gog, T., 2010. In the eyes of the beholder: how experts and novices interpret dynamic stimuli. Learn. InStruct. 20 (2), 146–154. https://doi.org/10.1016/j.learninstruc.2009.02.019.

Jeong, H., Kang, Z., Liu, Y., 2019. Driver glance behaviors and scanning patterns: applying static and dynamic glance measures to the analysis of curve driving with secondary tasks. Hum. Factors Ergon. Manuf. 29 (6), 437–446. https://doi.org/10.1002/hfm.20798.

Kim, N., House, R., Yun, M.H., Nam, C.S., 2019. Neural correlates of workload transition in multitasking: an ACT-R model of hysteresis effect. Front. Hum. Neurosci. 12, 535. https://doi.org/10.3389/fnhum.2018.00535.

Komogortsev, O.V., Karpov, A., 2013. Automated classification and scoring of smooth pursuit eye movements in the presence of fixations and saccades. Behav. Res. Methods 45 (1), 203–215. https://doi.org/10.3758/s13428-012-0234-9.

Komogortsev, O.V., Gobert, D.V., Jayarathna, S., Koh, D.H., Gowda, S.M., 2010. Standardization of automated analyses of oculomotor fixation and saccadic behaviors. IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 57 (11), 2635–2645. https://doi.org/10.1109/TBME.2010.2057429.

Kovesdi, C., Le Blanc, K., Spielman, Z., Hill, R., Oxstrand, J., 2018. Quantifying the contribution of individual display features on fixation duration to support human-system interface design in nuclear power plants. In: Joint Proceedings of the AHFE 2018 International Conference on Human Factors in Artificial Intelligence and Social Computing, Software and Systems Engineering, the Human Side of Service Engineering and Human Factors in Energy, 453–464. http://www.springer.com/series/11156.

Krafka, K., Khosla, A., Kellnhofer, P., Kannan, H., Bhandarkar, S., Matusik, W., Torralba, A., 2016. Eye tracking for everyone. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2176–2184. http://gazecapture.csail.mit.edu.

Krejtz, K., Duchowski, A., Szmidt, T., Krejtz, I., González Perilli, F., Pires, A., Vilaro, A., Villalobos, N., 2015. Gaze transition entropy. Trans. Appl. Percept. 13 (1), 1–20. https://doi.org/10.1145/2834121, 4.

Krejtz, K., Duchowski, A., Krejtz, I., Szarkowska, A., Kopacz, A., 2016. Discerning ambient/focal attention with coefficient *K*. Trans. Appl. Percept. 13 (3), 1–20. https://doi.org/10.1145/2896452, 11.

Krulewitz, J.E., Warm, J.S., Wohl, T.H., 1975. Effects of shifts in the rate of repetitive stimulation on sustained attention. Percept. Psychophys. 18 (4), 245–249. https://doi.org/10.3758/BF03199369.

Liechty, J., Pieters, R., Wedel, M., 2003. Global and local covert visual attention: evidence from a bayesian hidden Markov model. Psychometrika 68 (4), 519–541.

Mathieu, J.E., Aguinis, H., Culpepper, S.A., Chen, G., 2012. Understanding and estimating the power to detect cross-level interaction effects in multilevel modeling. J. Appl. Psychol. 97 (5), 951–966. https://doi.org/10.1037/a0028380.

Matthews, M.L., 1986. The influence of visual workload history on visual performance. Hum. Factors 28 (6), 623–632. https://doi.org/10.1177/001872088602800601.

Matthews, G., Desmond, P.A., 2002. Task-induced fatigue states and simulated driving performance. Q. J. Exp. Psychol. 55A (2), 659–686. https://doi.org/10.1080/0272498014300050.

Matthews, G., Reinerman-Jones, L.E., Barber, D.J., Abich IV, J., 2015. The psychometrics of mental workload: multiple measures are sensitive but divergent. Hum. Factors 57 (1), 125–143. https://doi.org/10.1177/0018720814539505.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., Bates, D., 2017. Balancing Type I error and power in linear mixed models. J. Mem. Lang. 94, 305–315. https://doi.org/10.1016/j.jml.2017.01.001.
McKendrick, R., Harwood, A., 2019. Cognitive workload and workload transitions elicit curvilinear hemodynamics during spatial working memory. Front. Hum. Neurosci. 13, 1–16. https://doi.org/10.3389/fnhum.2019.00405, 405.
Mirman, D., Dixon, J.A., Magnuson, J.S., 2008. Statistical and computational models of the visual world paradigm: growth curves and individual differences. J. Mem. Lang. 59 (4), 475–494. https://doi.org/10.1016/j.jml.2007.11.006.
Moacdieh, N., Sarter, N., 2015. Display clutter: a review of definitions and measurement techniques. Hum. Factors 57 (1), 61–100. https://doi.org/10.1177/0018720814541145.
Moacdieh, N.M., Devlin, S.P., Jundi, H., Riggs, S.L., 2020. Effects of workload and workload transitions on attention allocation in a dual-task environment: evidence from eye tracking metrics. J.Cognit. Eng.Decis. Making 14 (2), 132–151. https://doi.org/10.1177/1555343419892184.
Morgan, J.F., Hancock, P.A., 2011. The effect of prior task loading on mental workload: an example of hysteresis in driving. Hum. Factors 53 (1), 75–86. https://doi.org/10.1177/0018720810393505.
Moroney, B.W., Warm, J.S., Dember, W.N., 1995. Effects of demand transitions on vigilance performance and perceived workload. In: Proceedings of the Human Factors and Ergonomics Society. SAGE Publications, Inc, pp. 1375–1379. https://doi.org/10.1177/154193129503902104.
Mracek, D.L., Arsenault, M.L., Day, E.A., Hardy III, J.H., Terry, R.A., 2014. A multilevel approach to relating subjective workload to performance after shifts in task demand. Hum. Factors 56 (8), 1401–1413. https://doi.org/10.1177/0018720814533964.
Naeeri, S., Mandal, S., Kang, Z., 2019. Analyzing pilots' fatigue for prolonged flight missions: multimodal analysis approach using vigilance test and eye tracking. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 111–115. https://doi.org/10.1177/1071181319631092.
Nyström, M., Holmqvist, K., 2010. An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. Behav. Res. Methods 42 (1), 188–204. https://doi.org/10.3758/BRM.42.1.188.
Oskarsson, P.A., Eriksson, L., Carlander, O., 2012. Enhanced perception and performance by multimodal threat cueing in simulated combat vehicle. Hum. Factors 54 (1), 122–137. https://doi.org/10.1177/0018720811424895.
O'Keefe, P.A., Letourneau, S.M., Homer, B.D., Schwartz, R.N., Plass, J.L., 2014. Learning from multiple representations: an examination of fixation patterns in a science simulation. Comput. Hum. Behav. 35, 234–242. https://doi.org/10.1016/j.chb.2014.02.040.
Poole, A., Ball, L.J., 2006. Eye tracking in human-computer interaction and usability research: current status and future prospects. Encycl. Hum. Comput. Interact. 211–219. https://doi.org/10.4018/978-1-59140-562-7.ch034.
Prytz, E.G., Scerbo, M.W., 2015. Changes in stress and subjective workload over time following a workload transition. Theor. Issues Ergon. Sci. 16 (6), 586–605. https://doi.org/10.1080/1463922X.2015.1084397.
R Core Team, 2021. R: the R project for statistical computing. In: R Foundation for Statistical Computing. https://www.r-project.org.
Raptis, G., Katsini, C., Belk, M., Fidas, C., Samaras, G., Avouris, N., 2017. Using eye gaze data and visual activities to infer human cognitive styles: method and feasibility studies. In: Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization, pp. 164–173. https://doi.org/10.1145/3079628.3079690.
Ratwani, R.M., McCurry, J.M., Trafton, J.G., 2010. Single operator, multiple robots: an eye movement based theoretic model of operator situation awareness. In: HRI '10 Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction, 55, pp. 235–242. https://doi.org/10.1109/HRI.2010.5453191.
Riggs, S.L., Wickens, C.D., Sarter, N., Thomas, L.C., Nikolic, M.I., Sebok, A., 2017. Multimodal information presentation in support of NextGen operations. Int. J.Aero. Psychol 27 (1–2), 29–43. https://doi.org/10.1080/10508414.2017.1365608.
Rizopoulos, D., 2021. Package 'GLMMadaptive'. J. Comput. Graph Stat. 4 (1) https://doi.org/10.1080/10618600.1995.10474663.
Sarter, N.B., Mumaw, R.J., Wickens, C.D., 2007. Pilots' monitoring strategies and performance on automated flight decks: an empirical study combining behavioral and eye-tracking data. Hum. Factors 49 (3), 347–357. https://doi.org/10.1518/001872007X196685.
Shic, F., Chawarska, K., Bradshaw, J., Scassellati, B., 2008. Autism, Eye-Tracking, Entropy. 7th IEEE International Conference on Development and Learning, ICDL. https://doi.org/10.1109/DEVLRN.2008.4640808, 73–78.
Shiferaw, B., Downey, L.A., Westlake, J., Stevens, B., Rajaratnam, S.M.W., Berlowitz, D. J., Swann, P., Howard, M.E., 2018. Stationary gaze entropy predicts lane departure events in sleep-deprived drivers. Sci. Rep. 8 (1), 1–10. https://doi.org/10.1038/s41598-018-20588-7, 2220.
Shiferaw, B., Crewther, D., Downey, L.A., 2019a. Gaze entropy measures reveal alcohol-induced visual scanning impairment during ascending and descending phases of intoxication. J. Stud. Alcohol Drugs 80 (2), 236–244. https://doi.org/10.15288/jsad.2019.80.236.
Shiferaw, B., Downey, L., Crewther, D., 2019b. A review of gaze entropy as a measure of visual scanning efficiency. Neurosci. Biobehav. Rev. 96, 353–366. https://doi.org/10.1016/j.neubiorev.2018.12.007.
Sibley, C., Coyne, J., Morrison, J., 2015. Research considerations for managing future unmanned systems. In: AAAI Spring Symposium, pp. 54–59.
Sibley, C.M., Foroughi, C.K., Olson, T., Moclaire, C., Coyne, J.T., 2017. Practical considerations for low-cost eye tracking: an analysis of data loss and presentation of a solution. In: Schmorrow, D., Fidopiastis, C. (Eds.), International Conference on Augmented Cognition, vol. 10284. Springer, Cham, pp. 236–250. https://doi.org/10.1007/978-3-319-58628-1.
Steichen, B., Carenini, G., Conati, C., 2013. User-adaptive information visualization - using eye gaze data to infer visualization tasks and user cognitive abilities. In: IUI '13, pp. 19–22.
Ungar, N.R., Matthews, G., Warm, J.S., Dember, W.N., Thomas, J.K., Finomore, V.S., Shaw, T.H., 2005. Demand transitions and tracking performance efficiency: structural and strategic models. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 1523–1526. https://doi.org/10.1177/154193120504901704.
United States Department of Defense, 2017. Unmanned Systems Integrated Roadmap FY 2017-2042.
Victor, T.W., Tivesten, E., Gustavsson, P., Johansson, J., Sangberg, F., Ljung Aust, M., 2018. Automation expectation mismatch: incorrect prediction despite eyes on threat and hands on wheel. Hum. Factors 60 (8), 1095–1116. https://doi.org/10.1177/0018720818788164.
Voeten, Cesko C., 2019. Using 'buildmer'to automatically find & compare maximal (mixed) models. R package version 1, 1–7.
Wickens, C.D., 1980. The structure of attentional resources. In: Nickerson, R.S. (Ed.), Attention and Performance VIII. Psychology Press, pp. 239–257.
Williams, K.W., 2006. Human Factors Implications of Unmanned Aircraft Accidents: Flight-Control Problems. https://doi.org/10.1016/S1479-3601(05)07008-6.
Wilson, G.F., Russell, C.A., 2007. Performance enhancement in an uninhabited air vehicle task using psychophysiologically determined adaptive aiding. Hum. Factors 49 (6), 1005–1018. https://doi.org/10.1518/001872007X249875.1006.
Yang, F.-Y., Tsai, M.-J., Chiou, G.-L., Wen-Yu Lee, S., Chang, C.-C., Chen, L.-L., 2018. Instructional suggestions supporting science learning in digital environments based on a review of eye tracking studies. Educat. Technol. Soc. 21 (2), 28–45. https://doi.org/10.2307/26388377.

Shannon P. Devlin is a General Engineer at the Naval Research Laboratory. She earned her Ph.D. in Systems Engineering at the University of Virginia in 2021.

Noelle Brown is a research scientist at the Naval Research Laboratory. She earned her Ph. D. in Cognitive Psychology from Louisiana State University in 2011.

Sabrina Drollinger is a Personnel Psychologist with the U. S. Navy. She earned her Master's degree in Industrial/Organizational Psychology from the University of Georgia in 2007.

Ciara Sibley is an Engineering Research Psychologist at the Naval Research Laboratory. She earned her M.A. from George Mason University in Human Factors and Applied Cognition in 2009.

Jawad Alami is a Ph.D. student in Systems Engineering at the University of Virginia. He earned his B.E. in Computer and Communication Engineering from the American University of Beirut in 2018.

Sara L. Riggs is an assistant professor in the Engineering Systems and Environment department at the University of Virginia. She earned her Ph.D. in Industrial and Operations Engineering from University of Michigan in 2014.