

Real Time Assessment of Cognitive State: Research and Implementation Challenges

Michael C. Trumbo¹(✉), Mikaela L. Armenta¹, Michael J. Haass¹, Karin M. Butler¹,
Aaron P. Jones², and Charles S.H. Robinson²

¹ Sandia National Laboratories, Albuquerque, USA
{mctrumb, mlarmen, mjhaass, kbutle}@sandia.gov

² University of New Mexico, Albuquerque, USA
{aaronjones, charob}@unm.edu

Abstract. Inferring the cognitive state of an individual in real time during task performance allows for implementation of corrective measures prior to the occurrence of an error. Current technology allows for real time cognitive state assessment based on objective physiological data though techniques such as neuroimaging and eye tracking. Although early results indicate effective construction of classifiers that distinguish between cognitive states in real time is a possibility in some settings, implementation of these classifiers into real world settings poses a number of challenges. Cognitive states of interest must be sufficiently distinct to allow for continuous discrimination in the operational environment using technology that is currently available as well as practical to implement.

Keywords: Cognitive state · Real time · Eye tracking · Attention

1 Introduction

The link between the mind and the brain makes it possible to use objective physiological signals to infer the mental state of an individual. Decoding of mental states, however, is currently limited with regard to both precision and efficiency. In a typical scenario, data is collected over many trials that occur under varying conditions designed to elicit distinct cognitive states, then the aggregate data is subjected to statistical analyses in order to determine if discrimination between cognitive states is possible based on physiological data associated with each task condition (see [1] for review). This approach increases statistical sensitivity by virtue of including a large number of samples in the analysis, but it does not allow for cognitive state inference on a trial-by-trial basis, thereby limiting efficiency.

Construction of a classifier capable of real time cognitive state inference is desirable in a number of contexts as it allows for implementation of adjustments that prevent errors or enhance performance while a task is ongoing. For instance, assessment of cognitive state in real time may prove valuable in determining if a driver is distracted or engaged in the driving task [2], in providing feedback with regard to emotional state and cognitive load of operators [3], in classifying an individual as alert or fatigued during visual search [4], and in detecting deception allowing for adjustment of questioning [5]. Researchers

have used a number of technologies to accomplish real time cognitive state assessment, often by using neuroimaging data such as electroencephalogram (EEG; [6]), functional magnetic resonance imaging (fMRI; [7]), or functional near-infrared spectroscopy (fNIRS; [8]), as well as physiological data such as galvanic skin response (GSR) and heart rate [9]. Such technologies, however, have several limitations. For instance, MRI equipment is expensive, immobile, noisy, necessitates a shielded room, and requires participants to lie perfectly still during scanning, making it infeasible to implement in most job contexts. EEG, fNIRS, GSR, and heart rate monitoring all require an individual to wear equipment that limits mobility and may become uncomfortable over time, making these methods ill-suited for data collection over long durations of time, such as a regular workday of 8 h [10].

Eye tracking technology uses video recordings from cameras which may be unobtrusively placed in nearly any environment to quantify eye phenomena such as pupil size, eye fixations, eye movements, and blinks [11]. The temporal resolution of these recordings is on the order of milliseconds, allowing for discrimination between cognitive states in real time using these eye metrics [4]. While the notion that eyes are the window to the soul may be debatable, eye metrics are thought to reflect brain activity, thereby acting as a window to cognitive processes [12–14]. Brain state information may be less susceptible to conscious or unconscious countermeasures than indirect metrics such as blood pressure, respiration, and GSR [15], thereby making incorporation of brain state information particularly valuable as a relatively clean source of data reflecting the mental state of an individual.

Broadly speaking, the human visual system may be divided into two modes of attentional processing: ambient and focal. During ambient processing attention is allocated broadly across the visual field. This type of processing is subserved by a dorsal neural pathway in the brain dedicated to determining where in space objects exist and how to interact with them. During focal processing attention is allocated to a region of the visual field defined by an object. Focal attention is associated with a ventral neural pathway concerned with object identification [16]. These distinct brain processes manifest in the amplitude and duration of the saccadic eye movements that occur between fixations (a relative stillness in eye position lasting from tens of milliseconds to several seconds and largely considered indicative of attention to the given position), may be categorized as related to ambient or focal processing based on the amplitude and duration of saccades, or the quick motions of the eye from one fixation to another, prior and subsequent to the fixation [11, 17]. Thus, for an individual engaged in a task with a visual component it is possible to monitor the balance between ambient and focal attention in real time. Eye tracking is therefore efficient, as its temporal resolution is conducive to real time monitoring, as well as relatively precise, in that it enables classification depth beyond a coarse inference of whether or not an individual is allocating attentional resources to a task by further allowing determination of type of attentional resources allocated.

Often, visual search follows a coarse-to-fine strategy in which initial search is dominated by ambient processing characterized by long saccades and short fixations in order to acquire the gist of a scene, followed by focal search characterized by short saccades and long fixations for purposes of object recognition; essentially exploration followed by inspection [18]. There are, however, exceptions – for instance free viewing of natural

scenes such as land or cityscapes in the absence of a task or goal is characterized by dominant focal attention, both over time and immediately following stimulus onset [17]. Thus, it is important to consider how task demands influence search patterns [19]. Real-time assessment of the balance between ambient and focal attention may be a valuable predictor of performance in a variety of domains, enabling implementation of mitigation measures prior to a critical incident. For instance, drivers have demonstrated selective impairment in response to side tasks depending on if the ancillary task required ambient or focal processing [20], while level of stress [21] and mood [22] bias the amount of time spent in each processing mode. Thus, real-time assessment of visual processing that allows for discrimination between focal and ambient processing may allow performance predication as well as inference of cognitive state.

One method of visualizing the dynamic interplay between ambient and focal attention is to compute the K-Coefficient by standardizing the current fixation duration and the subsequent saccade amplitude and taking the difference between them, with positive ordinates indicating focal processing at the fixation and negative indicating ambient processing [23]. In the current work, we applied this method to two categories of task-oriented viewings of natural scenes. For the first category of images, we gave participants the task of determining either the profession portrayed in a photograph or the season during which the photograph was taken. For the second category, we asked participants either to count the number of vehicles in an image or to count the number of floors in the tallest building in the image. Eye movements were recorded during image search, and fixations were subsequently categorized as ambient or focal using the K-Coefficient algorithm. The results from this analysis are used to underscore obstacles to real time cognitive state assessment.

2 Method

Participants. Forty-one participants (Males = 22; Age: Mean = 26.97, SD = 11.53, Range = 17–65 years), employees, interns, staff, and contractors of Sandia National Laboratories, participated in this study. Three did not provide the requisite demographic information. Technical issues led to loss of data in two instances, reducing usable data to $n = 39$ (Males = 20; Age: Mean = 27.14, SD = 11.69, Range = 17–65 years). Participants were paid their hourly wage for participation in the study.

Stimuli. Photos, 13 in all, were divided into two categories: (1) five with cars and buildings (CB) and (2) eight that we have termed ‘Seasons, Professions’ (SP), photographs of people in various roles ranging from that of an early 20th century factory assembly line to policemen.

Visual Search Task. Participants completed a still image search task in which they were directed to answer a given question based on what they discerned from various photographs. The task was created and presented with EyePresentation, visual stimulus presentation software developed at Sandia National Laboratories. Stimuli were presented at a resolution of 1280×1024 on a Dell 1901FP LCD Monitor (38 cm \times 30 cm). Before viewing the photos, participants were presented with a slide that asked

questions similar to a Yarus task [24]. Questions on instruction slides pertaining to CB images included “How many vehicles can you see?” or “How many stories does the tallest building have?” while questions preceding SP images were “In what season was this photo taken?” and “What job or profession is portrayed in the photo?” While all participants saw all images, questions were counterbalanced by condition (CB and SP). After viewing an instruction slide, participants viewed a fixation cross for 1 s. A fixation cross preceded all instruction, image stimulus and response slides. All participants viewed the photographs in the same order, but two orders of questions were used such that across participants both questions were asked for each image. For instance, in the SP condition both groups saw a photo of an early twentieth century assembly line, however one group ($n = 18$) was asked “What job or profession is portrayed in the photo?” and the other ($n = 21$) responded to, “In what season was this photo taken?” Our other condition (CB) featured city and landscapes in which participants were asked either “How many stories are in the tallest building?” or “How many vehicles are there?” When prepared to answer the question, participants pressed the space bar on a keyboard. The image was replaced by the fixation cross and followed by a response slide. Each phase of the trial was self-paced. Participants pressed the space bar on a keyboard to progress to the next phase when they understood the question, obtained an answer, and had verbalized their answer for the experimenter. There was, however, a ceiling of 45 s for instruction, image stimulus and response slide viewing. Most participants proceeded through the task before reaching this. See Fig. 1 for visualization of the task progression.

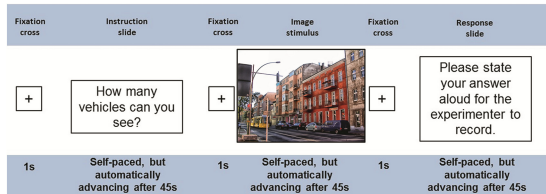


Fig. 1. Timeline of each trial, using vehicle counting as an example. The progression from instructions to image stimulus to response was self-paced, within a 45 s time window for each portion of the trial. Image source: CC0 Public Domain, <https://pixabay.com/en/berlin-germany-urban-buildings-103154/>.

Procedure. After being informed of their rights as research participants and providing informed consent, participants completed a demographic questionnaire providing information on age, sex, and experience with visual imagery work such as photography and other forms of imagery. Participants were then seated in a 148 cm long \times 102 cm wide (58 in. \times 40in.) ‘SE 2000’ soundproof, dark room manufactured by Whisper Room Sound Isolation Enclosures at a viewing distance of 54–92 cm from the computer monitor. The door was propped open to enable verbal communication between the participant and experimenter. The open door allowed ambient light to enter the room and the amount varied between participants. Although we did not take light measurements, typical illumination ranges from 1 lx in a very dim room to a typical value of

90 lx in a work or living space [25]. Prior research indicates that illumination in the range of 0 to 1000 lx does not significantly affect accuracy or precision of eye tracking data [26].

We collected eye tracking data using EyeWorks Suite (v. 3.12.5395.30517) on a DELL Precision T3600 using Windows 7 operating system on an Intel Xeon CPU E5-1603 0 @ 2.80 GHz with 8 GB of RAM. All stimuli were presented at a resolution of 1280×1024 on a DELL 19" LCD monitor. We utilized a 60 Hz FOVIO eye-tracker manufactured by Seeing Machines and verified calibration through a five-point calibration procedure in EyeWorks Record prior to the task. Calibration was considered sufficient if the dot following the eye movement trajectory was sustained (indicating that the eye movement monitor was not losing tracking) and if the calibration dot was accurate (falling on the calibration check targets at the center and corners of the screen when the participant was instructed to look at them, with inaccuracy of up to one centimeter for the upper two corner targets). The eye-tracker was located between 9.5 cm and 8 cm beneath the bottom of the viewing screen. Following calibration, participants completed the task as described above, and in Fig. 1.

After completing this task, the FOVIO was recalibrated before moving on to a Smooth Pursuit task. As this task is not pertinent to this article, it is not discussed further but is described in [27]. Upon completion of this task, the experimental portion of the study was complete and subjects discussed the study with the experimenter before leaving. From consent to debriefing, the study duration spanned roughly 45 min.

3 Results

Fixations were classified as relating to ambient or focal attention using the K-Coefficient [23], which was calculated for each fixation using the difference between the standardized values (z-scores) of the duration of the current fixation (d_i) and subsequent saccade amplitude (a_{i+1}), as such:

$$K_i = \frac{d_i - \mu_d}{\sigma_d} - \frac{a_{i+1} - \mu_a}{\sigma_a}, \quad i \in [1, n - 1] \quad (1)$$

In this equation, μ_d and σ_d represent the mean and standard deviation, respectively, of the fixation duration while μ_a and σ_a represent the mean and standard deviation, respectively, of the saccade amplitude, computed over all n fixations. Positive values therefore represent a relatively long fixation followed by a relatively short saccade, indicative of focal attention for that fixation. Conversely, negative values represent a relatively short fixation followed by a relatively long saccade, indicative of a state of ambient attention for that fixation. Similar to prior research [17], scanpaths with less than four total fixations were excluded from analyses. In addition, the first fixation for a given trial was removed from analysis to compensate for the effect of the fixation cross which biases first fixations toward the center of the screen.

In order to evaluate potential differences between conditions with regard to ambient and focal attention, a RM ANOVA was conducted, with two levels of fixation type (ambient vs. focal) and four levels of condition (cars vs. buildings vs. seasons vs. professions).

All pairwise comparisons were planned, and thus no alpha correction was applied. This analysis revealed an attention \times condition interaction, ($F(3,36) = 2.930, p = 0.047$), as well as a main effect of attention ($F(1,38) = 8.946, p = 0.005$) and a main effect of condition ($F(3,36) = 8.325, p < 0.001$). Follow-up pairwise comparisons indicate a greater number of focal vs. ambient fixations (mean difference .617, $p = 0.005$) and that more fixations (collapsed across attention type) occurred in the buildings vs. cars condition (mean difference 4.438, $p < 0.001$), in the buildings vs. professions condition (mean difference 3.651, $p < 0.001$), in the seasons vs. cars condition (mean difference 3.511, $p < 0.001$), and in the seasons vs. professions condition (mean difference 2.727, $p < 0.001$).

Additionally, there was a simple effect of condition within ambient attention ($F(3,36) = 6.774, p < 0.001$) and a simple effect of condition within focal attention ($F(3,36) = 9.460, p < 0.001$). Follow-up pairwise comparisons indicate a greater number of ambient fixations in the buildings vs. cars condition (mean difference 3.628, $p < .001$), in the buildings vs. professions condition (mean difference 2.934, $p = 0.003$), in the seasons vs. cars condition (mean difference 3.325, $p < .001$), and in the seasons vs. professions condition (mean difference 2.630, $p < 0.001$). Additional pairwise comparisons indicate a greater number of focal fixations in the buildings vs. cars condition (mean difference 5.248, $p < 0.001$), in the buildings vs. professions condition (mean difference 4.368, $p < 0.001$), in the seasons vs. cars condition (mean difference 3.697, $p < 0.001$), and in the seasons vs. professions condition (mean difference 2.816, $p = 0.002$).

These pairwise comparisons indicate similar findings for both focal and ambient fixation types (more of both types for buildings vs. cars and vs. professions, and more of both type for seasons vs. cars and vs. professions). In order to determine if these findings may be an artifact of differential search times between conditions (given the nature of the task involved self-terminated search up to a ceiling of 45 s) we conducted a number of pairwise comparisons evaluating the effect of condition on search time, collapsed across attention. These comparisons revealed a greater amount of time (in seconds) was spent searching in the buildings vs. cars condition (mean difference 1.445, $p = 0.004$), in the buildings vs. seasons condition (mean difference 1.541, $p < 0.001$), in the buildings vs. professions condition (mean difference 2.215, $p < 0.001$), and in the seasons vs. professions condition (mean difference .674, $p = 0.021$; see Fig. 2). As these differences closely mapped onto the findings for number of fixations, we re-ran the previous analysis with total search time within the appropriate condition(s) as a covariate. With this control in place, the difference in number of focal or ambient fixations between conditions was no longer significant for any comparisons between conditions (all $ps > .05$).

Due to the overall differences in search times between conditions, and based on previously identified visual processing periods, we divided data into early, middle, and late search periods, as in [28]. The early search period corresponds to the first 1.5 s of scene viewing, the middle to 1.5–3.0 s, and the late 3.0–4.5 s. Early visual search phases are thought to be driven largely by bottom-up visual features (e.g., color, edges, luminance, motion) while during later stages top-down influences (e.g., task goals, motivations, strategies) are thought to dominate [29]. Therefore, we investigated the possibility that bottom-up and top-down search contributions would shift across these search

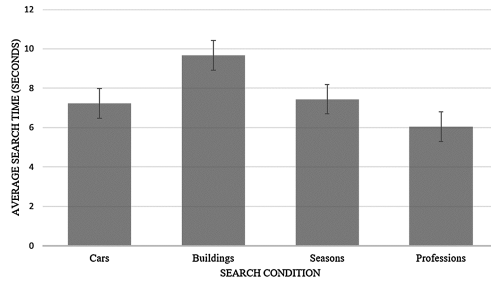


Fig. 2. Average search time in seconds broken down by search condition. Note that overall differences in search time between conditions account for differences between conditions in number of ambient and focal fixations.

periods, reflected in a difference in number of ambient or focal fixations over time. Collapsed by condition, there were no significant differences in number of ambient or focal fixations between time periods (all $ps > 0.05$; see Fig. 3), though the trend of more focal relative to ambient attention over time is consistent with previous literature [28].

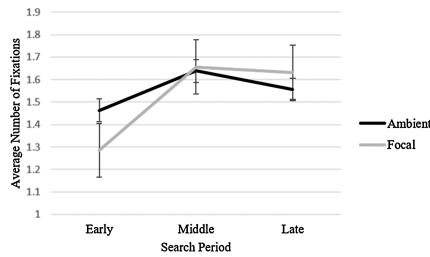


Fig. 3. Average number of fixations by search period and attention type, collapsed across condition.

Additional planned pairwise comparisons were conducted in order to investigate a potential effect of time within each condition and attention type. These pairwise comparisons revealed a significant difference within ambient attention for the buildings condition, such that the middle search period contained more fixations than the late search period (mean difference .494, $p = 0.013$) and for the professions condition such that the middle search period contained more ambient fixations than the early search period (mean difference .342, $p = 0.031$). Within focal attention, for the buildings condition more fixations occurred within the middle search period than the late search period (mean difference .414, $p = 0.018$), and more within the late search compared to the early search period (mean difference .523, $p = 0.003$). For the seasons condition, more focal fixations occurred within the middle than the early search period (mean difference .560, $p < 0.001$) and more focal fixations occurred within the late vs. early search period (mean difference .517, $p < 0.001$). For the professions condition, more focal fixations occurred within the middle vs. early search period (mean difference .431, $p = 0.002$) and within the late vs. early search period (mean difference .520, $p < .001$).

We also calculated planned pairwise comparisons between conditions within early, middle, and late search periods in order to determine if there was a difference within any search period for the different search conditions. These analyses revealed no significant differences in ambient fixations within the early search period (all p s > 0.05), but a significant difference in focal fixations within the early search period such that there were more focal fixations in the cars condition vs. the professions condition (mean difference .325, $p = 0.041$). Within the middle search period there was a difference in ambient fixations, with more fixations in the seasons condition vs. the cars condition (mean difference .678, $p = 0.008$) and more in the profession condition than the cars condition (mean difference .580, $p = 0.021$). There were no differences for focal fixations in the middle search period (all p s > 0.05). In the late search period, there were differences in number of ambient fixations, with more occurring in the seasons vs. cars condition (mean difference .555, $p = .005$), more in the seasons vs. buildings condition (mean difference .739, $p < 0.001$), more in the professions vs. cars condition (mean difference .434, $p = 0.04$), and more in the professions vs. buildings condition (mean difference .618, $p = 0.01$).

4 Discussion

Eye tracking represents a non-invasive method of collecting data that may be implemented in many research and job contexts without interfering with task performance, and is therefore a technology worth pursuing in the realm of real-time assessment. There are, however, a number of challenges to overcome with regard to both research and practical implementations. The preliminary results of the current study indicate the ability to discriminate between ambient and focal attention states and differentiation by task. We found that when search time is self-terminated there were differences in the average time spent searching between image conditions; notably, participants spent more time searching in the buildings condition than in the other search conditions. Controlling for this difference in average search time rendered differences in overall number of ambient and focal fixations between conditions non-significant. Dividing search time into early, middle, and late search periods, however, revealed several differences between conditions with regard to ambient and focal attention within a given search period. Notably, differences between conditions within ambient were limited to middle and late search periods, while differences within focal were limited to early search. Within conditions, we found evidence for an effect of time within attention type. For the seasons and professions conditions, there tended to be more ambient fixations in the middle search period relative to the early, while for the buildings condition there were more ambient fixations in the middle vs. late search period. For focal attention, there were significantly more fixations in the middle and late search periods vs. the early search period for the buildings, seasons, and professions conditions.

Focal attention is thought to be driven by bottom-up stimulus characteristics to a greater degree than ambient attention when participants are free-viewing natural scenes [17], and is associated with top-down processing later on in visual search when knowledge and expectations guide a detailed inspection of a particular image feature [17, 29].

The current task was goal-directed, standing in contrast to previous research that has characterized attentional states under free viewing conditions [17]. Giving participants a search goal prior to the onset of each image, as in the current research, may impact the time course of focal attention associated with top-down, goal-directed behavior. While this seemed to be the case for the cars condition, in all other conditions there was a greater number focal fixations in middle and late search periods relative to early search. This is consistent with prior research suggesting dominance of focal attention over time [28], and collapsing across conditions suggests an overall pattern of a shift toward focal attention as search time increases in goal-direct visual search (see Fig. 3). It is important to note that in the current study search time was self-terminated by participants, with a ceiling of 45 s. This potentially leaves substantial time following the “late” visual processing period cutoff of 4.5 s. We found a main effect of attention type such that across the entirety of the search period there were significantly more focal fixations than ambient fixations, but given the time span of the search it may be useful to use a finer analysis with regard to time. For instance, it is possible in some search conditions participants cycle through periods of ambient and focal domination as they move between areas of interest. This shifting cycle may be different between search conditions or over time, making it a potential valuable contribution to a classifier. Additionally, this may help determine if visual processing periods of the durations used in this and prior work are appropriate for self-terminated search which may substantially exceed 4.5 s.

For the cars and buildings conditions, participants were explicitly directed to locate targets (cars or buildings) and count them. In the seasons and professions condition, we asked participants to look at an image and make a judgement based on the evidence they were able to gather from cues within the image. It may be that this type of judgement requires more ambient processing as one develops a ‘gist’ of the situation. This is supported by a greater number of ambient fixations during the late search period in both the seasons and professions conditions relative to the cars and buildings conditions, suggesting persistence of ambient attention for conditions involving judgement based on image cues. Additionally, image stimuli in the cars and buildings conditions did not include any human beings, while the seasons and professions conditions contained human beings in each image, which are highly salient to attentional and visual processing [30].

Ecological Validity. The current research demonstrates that tasks under the same general domain (e.g., visual search) may elicit different patterns of cognitive states, contingent on such factors as specific stimuli content and task goals. Therefore, when attempting to generalize laboratory research to an operational environment, it is critical that the laboratory circumstances closely mimic those of the environment of interest. Often, real time cognitive state classifiers are built under laboratory conditions of strict control, using tasks specifically designed to elicit drastically different cognitive states in a binary fashion (e.g., alert vs. fatigued; anxious vs. relaxed) thereby maximizing the likelihood of building a successful classifier [1, 4]. Establishing that a classifier is effective under these ideal conditions is a valuable feasibility step. Real-world jobs, however, do not often embody the same method of utilizing extremely different task conditions for purposes of identifying distinct cognitive states based on physiological metrics.

Instead, they may elicit many overlapping cognitive states on a continuum not present in laboratory conditions [1]. In addition, real-world environments may introduce obstacles to data quality not present in a laboratory, such as substantial participant movement and sources of electrical interference that may dictate the types of technologies available for data collection [10]. Therefore, research geared toward application within a particular environment should attempt to mimic the environment of planned deployment as closely as possible.

Classifier Construction. Building a classifier capable of discriminating between cognitive states in real time invites a number of challenging decisions. One such decision is which data to collect and use. For instance, eye tracking offers the ability to collect several data streams in parallel, including eye movement metrics, pupil size, and blinks, all of which have been demonstrated to relate to cognitive state [4]. Previous research using eye tracking data suggests that a classifier constructed implementing all of these data streams significantly outperforms classifiers built using only one data stream [4]. In the current study, ambient fixations resulted in discrimination between certain search conditions within middle and late search periods, while number of focal fixations allowed discrimination during the early search period, suggesting value in tracking both types of fixations. It is therefore important to determine acceptable margins of error weighed against additional costs in terms of computational resources. In a similar vein, it may be true that each data stream allows discriminability between different cognitive states. For instance, blinks may provide critical information when discriminating between an alert vs. a fatigued state, while eye movement metrics may be particularly valuable in discriminating between different types of attention within an alert state, and pupil size may be the best indicator of cognitive load [11]. Thus, multiple data streams may allow for varying levels of cognitive state classification that are particularly well suited to performance prediction at different task stages. It is also worth noting that search time differed between conditions, with average search time in the buildings condition exceeding that of the other conditions. In many operational environments, visual search is self-terminated by the operator (e.g., airport bag screening, radar analysis); if different conditions allow classification of goal state, incorporating search time into a classifier may prove valuable.

An additional consideration is the type of classifier to use. Both linear and nonlinear models are available, and evidence suggests either may be effective [4]. Success, however, may interact with the time interval of data fed into each model. Nonlinear models were found to outperform linear models in two out of three cases when a 1 s interval was used, while linear models performed based when a 10 s interval was used [4]. It is worth noting that model performance varied between individuals such that one model did not exhibit superior performance across all individuals within any given condition. Likewise, an analysis of relative contribution of data streams (i.e., eye movements, blinks, pupil size) revealed that no single metric was a better predictor of cognitive state than the other metrics across all individuals for a given task. Therefore, individual calibration may be necessary when attempting to maximize the accuracy of a classifier.

Indeed, individual differences may affect classifier performance in a number of ways. For instance, eye tracking data quality may vary as a function of age, sex,

ethnicity, and disease state [26, 31], cultural variation has been found to influence patterns of visual scene inspection [32], and older adults may exhibit different viewing patterns than younger adults [33]. In addition to differences between individuals, substantial variability may occur within an individual or an individual's environment over time. For instance, adjustments in ambient light intensity and sound, use of caffeine or nicotine, and eye makeup all influence the quality of eye tracking [34]. Controlling for these factors can be a particular challenge if eye tracking is to be implemented under necessarily changing environmental circumstances, such as during driving. A number of studies have indicated that classifiers trained and tested using physiological data obtained over a single session are capable of discriminating between cognitive states [4, 17, 35–37], but classifier accuracy may deteriorate over time due to these individual and environmental changes [38]. Training a new classifier is time consuming, so one compromise is to adjust the classifier at regular intervals using a small amount of newly collected data. For a task involving complex multitasking, an additional 2.5–7 min of data per level of task difficulty was found to significantly enhance classifier accuracy for the remainder of the day, though this improvement was attenuated when extending classifier use for an additional day without incorporating additional training data [38]. Therefore, it may be useful to consider applying additional training data on at least a daily basis in order to account for individual and environmental changes over time.

5 Conclusion

Several technologies offer the possibility of assessing cognitive state in real time using objective physiological data, allowing for online adjustment of task demands to avoid costly errors. There are a number of conditions that must be met for successful implementation of a real time cognitive assessment system in a job setting. The task performed must elicit cognitive states which are discriminable using technology that is practical within the job setting, these cognitive states must map onto performance in a meaningful way, and the classifier must account for differences both between and within individuals and environments. Current eye tracking technology makes this a feasible endeavor within certain contexts, yet there are a number of critical considerations to be cognizant of when attempting translation from a laboratory to an operational environment.

Acknowledgements. We wish to acknowledge James D. Morrow of Sandia National Laboratories, Albuquerque New Mexico for creating the software used in our study to display the visual stimuli and record subject responses.

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

References

1. Haynes, J., Rees, G.: Decoding mental states from brain activity in humans. *Nat. Rev. Neurosci.* **7**(7), 523–534 (2006)
2. Liang, Y., et al.: Real-time detection of driver cognitive distraction using support vector machines. *IEEE Trans. Intell. Trans. Syst.* **8**(2), 340–350 (2007)
3. Neerinx, M.A., Harbers, M., Lim, D., van der Tas, V.: Automatic feedback on cognitive load and emotional state of traffic controllers. In: Harris, D. (ed.) *EPCE 2014. LNCS*, vol. 8532, pp. 42–49. Springer, Heidelberg (2014)
4. Marshall, S.P.: Identifying cognitive state from eye metrics. *Aviat. Space Environ. Med.* **78**(1), B165–B175 (2007)
5. deCharms, R.: Applications of real-time fMRI. *Nat. Rev. Neurosci.* **9**(9), 720–729 (2008)
6. Parra, L., et al.: Response error correction—a demonstration of improved human-machine performance using real-time EEG monitoring. *IEEE Trans. Neural Syst. Rehabil. Eng.* **11**(2), 173–177 (2003)
7. Cox, D., Savoy, R.: Functional magnetic resonance imaging (fMRI) “brain reading”: detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage* **19**(2), 261–270 (2003)
8. Snow, M.P., Barker, R.A., O’Neill, K.R., Offer, B.W., Edwards, R.E.: Augmented cognition in a prototype uninhabited combat air vehicle operator console. In: *Foundations of Augmented Cognition*. 2nd edn., pp. 279–288 (2006)
9. Ikehara, C.S., Crosby, M.E.: Assessing cognitive load with physiological sensors. In: *Proceedings of the 38th Annual HICSS 2005*, p. 295a, January 2005
10. Palmer, E.D., Kobus, D.A.: The future of augmented cognition systems in education and training. In: Schmorrow, D.D., Reeves, L.M. (eds.) *HCII 2007 and FAC 2007. LNCS (LNAI)*, vol. 4565, pp. 373–379. Springer, Heidelberg (2007)
11. Holmqvist, K., et al.: *Eye Tracking*. OUP Oxford, Oxford (2011)
12. Beatty, J., Lucero-Wagoner, B.: The pupillary system. *Handb. Psychophysiol.* **2**, 142–162 (2000)
13. Carlson-Radvansky, L.: Memory for relational information across eye movements. *Percept. Psychophys.* **61**(5), 919–934 (1999)
14. Stern, J., et al.: The endogenous eyeblink. *Psychophysiology* **21**(1), 22–33 (1984)
15. Farwell, L.A., Smith, S.S.: Using brain MERMER testing to detect knowledge despite efforts to conceal. *J. Forensic Sci.* **46**(1), 135–143 (2001)
16. Oliva, A., Schyns, P.G.: Diagnostic colors mediate scene recognition. *Cogn. Psychol.* **41**(2), 176–210 (2000)
17. Follet, B., et al.: New insights into ambient and focal visual fixations using an automatic classification algorithm. *i-Perception* **2**(6), 592–610 (2011)
18. Unema, P., et al.: Time course of information processing during scene perception: the relationship between saccade amplitude and fixation duration. *Vis. Cogn.* **12**(3), 473–494 (2005)
19. Scinto, L., et al.: Cognitive strategies for visual search. *Acta Psychol.* **62**(3), 263–292 (1986)
20. Lenneman, J.K., Lenneman, J., Cassavaugh, N., Backs, R.: Differential effects of focal and ambient visual processing demands on driving performance. In: *Proceedings of the Fifth International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, vol. 5, pp. 306–312. University of Iowa Public Policy Center (2009)
21. Leibowitz, H.W., Shupert, C.L., Post, R.B.: The two modes of visual processing: implications for spatial orientation. In: *PVHD, NASA Conference Publication*, vol. 2306, pp. 41–44, March 1983

22. Biele, C., Kopacz, A., Krejtz, K.: Shall we care about the user's feelings? Influence of affect and engagement on visual attention. In: Proceedings of the International Conference on Multimedia, Interaction, Design and Innovation, p. 7. ACM, June 2013
23. Duchowski, A.T. Krejtz, K.: Visualizing dynamic ambient/focal attention with coefficient K. In: Proceedings of the First Workshop on ETVIS, Chicago, IL, October 2015
24. Yarbus, A.: Eye Movements and Vision. pp. 171–211. Plenum Press, New York (1967)
25. Chang, A., et al.: The human circadian system adapts to prior photic history. *J. Physiol.* **589**(5), 1095–1102 (2011)
26. Blignaut, P., Wium, D.: Eye-tracking data quality as affected by ethnicity and experimental design. *Behav. Res.* **46**(1), 67–80 (2013)
27. Haass, M.J., Matzen, L.E., Butler, K.M., Armenta, M.: A new method for categorizing scanpaths from eye tracking data. In: ETRA 2016, April 2016
28. Pannasch, S., Velichkovsky, B.: Distractor effect and saccade amplitudes: further evidence on different modes of processing in free exploration of visual images. *Vis. Cogn.* **17**(6–7), 1109–1131 (2009)
29. Parkhurst, D., et al.: Modeling the role of salience in the allocation of overt visual attention. *Vis. Res.* **42**(1), 107–123 (2002)
30. Judd, T., Ehinger, K., Durand, F., Torralba, A.: Learning to predict where humans look. In: 2009 IEEE 12th International Conference on Computer Vision, pp. 2106–2113. IEEE September 2009
31. Kuechenmeister, C.: Eye tracking in relation to age, sex, and illness. *Arch. Gen. Psychiatry* **34**(5), 578 (1977)
32. Chua, H.F., Boland, J.E., Nisbett, R.E.: Cultural variation in eye movements during scene perception. *Proc. Nat. Acad. Sci. U.S.A.* **102**(35), 12629–12633 (2005)
33. Romano Bergstrom, J.C., Olmsted-Hawala, E.L., Jans, M.E.: Age-related differences in eye tracking and usability performance: website usability for older adults. *Int. J. Hum.-Comput. Inter.* **29**(8), 541–548 (2013)
34. O'Brien, S.: Eye tracking in translation process research: methodological challenges and solutions. *Methodol. Technol. Innov. Transl. Process. Res.* **38**, 251–266 (2009)
35. Berka, C., et al.: Real-Time analysis of EEG indexes of alertness, cognition, and memory acquired with a wireless EEG headset. *Int. J. Hum.-Comput. Inter.* **17**(2), 151–170 (2004)
36. Freeman, F., et al.: Evaluation of an adaptive automation system using three EEG indices with a visual tracking task. *Biol. Psychol.* **50**(1), 61–76 (1999)
37. Gevins, A., et al.: Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Hum. Factors* **40**(1), 79–91 (1998)
38. Christensen, J., et al.: The effects of day-to-day variability of physiological data on operator functional state classification. *NeuroImage* **59**(1), 57–63 (2012)