

Through a Scanner Quickly: Elicitation of P3 in Transportation Security Officers Following Rapid Image Presentation and Categorization

Michael C. Trumbo^(✉), Laura E. Matzen, Austin Silva,
Michael J. Haass, Kristin Divis, and Ann Speed

Sandia National Laboratories, Albuquerque, USA
{mctrumb, lematze, aussilv, mjhaass, kmdivis,
aespeed}@sandia.gov

Abstract. Numerous domains, ranging from medical diagnostics to intelligence analysis, involve visual search tasks in which people must find and identify specific items within large sets of imagery. These tasks rely heavily on human judgment, making fully automated systems infeasible in many cases. Researchers have investigated methods for combining human judgment with computational processing to increase the speed at which humans can triage large image sets. One such method is rapid serial visual presentation (RSVP), in which images are presented in rapid succession to a human viewer. While viewing the images and looking for targets of interest, the participant's brain activity is recorded using electroencephalography (EEG). The EEG signals can be time-locked to the presentation of each image, producing event-related potentials (ERPs) that provide information about the brain's response to those stimuli. The participants' judgments about whether or not each set of images contained a target and the ERPs elicited by target and non-target images are used to identify subsets of images that merit close expert scrutiny [1]. Although the RSVP/EEG paradigm holds promise for helping professional visual searchers to triage imagery rapidly, it may be limited by the nature of the target items. Targets that do not vary a great deal in appearance are likely to elicit useable ERPs, but more variable targets may not. In the present study, we sought to extend the RSVP/EEG paradigm to the domain of aviation security screening, and in doing so to explore the limitations of the technique for different types of targets. Professional Transportation Security Officers (TSOs) viewed bag X-rays that were presented using an RSVP paradigm. The TSOs viewed bursts of images containing 50 segments of bag X-rays that were presented for 100 ms each. Following each burst of images, the TSOs indicated whether or not they thought there was a threat item in any of the images in that set. EEG was recorded during each burst of images and ERPs were calculated by time-locking the EEG signal to the presentation of images containing threats and matched images that were identical except for the presence of the threat item. Half of the threat items had a prototypical appearance and half did not. We found that the bag images containing threat items with a prototypical appearance reliably elicited a P300 ERP component, while those without a prototypical appearance did not. These findings have implications for the application of the RSVP/EEG technique to real-world visual search domains.

Keywords: Rapid serial visual presentation · Visual search · EEG · P300

1 Introduction

A wide variety of domains, ranging from medical diagnostics to intelligence analysis, involve searching through large sets of imagery to find and identify specific items. These domains rely on people's ability to discriminate between relevant and irrelevant images as accurately and efficiently as possible. While computer vision systems have been employed in image classification [2], purely computerized systems may lack the sensitivity, specificity and ability to generalize possessed by humans [1, 3], making fully automated systems infeasible in these complex domains. Since visual search and inspection tasks rely primarily on human judgment, researchers have sought other methods for increasing the speed at which humans can triage large sets of imagery. In one such method, termed rapid serial visual presentation (RSVP), images are presented serially in a fixed location, typically at a rate of 3–20 items per second [4]. Intraub [5] demonstrated that participants can accurately identify targets within a rapid stream of images. The RSVP technique has subsequently been employed to study phenomena ranging from language processing [6], to emotion [7], to attention [8].

Recently, researchers have investigated combining the RSVP technique with brain-computer interface (BCI) technology. In this approach, participants typically view image chips in which a larger image is segmented into many small parts. The chips are presented rapidly, in short bursts, and the participants judge whether or not there was a target present in any of the images in that group. Meanwhile, participants' brain activity is recorded using electroencephalography (EEG), a neuroimaging technique that provides temporal resolution on the order of milliseconds [9]. EEG signals can be time-locked to the presentation of stimuli, producing event-related potentials (ERPs) that provide information about the brain's response to those stimuli. The participants' judgments about whether or not each set of images contained a target and the ERPs elicited by target and non-target images are used to identify subsets of images that merit close expert scrutiny [1]. This approach can allow imagery analysts to hone in on the relevant information very rapidly. The ERP signals can also be combined with machine learning techniques to develop classifiers, which can then be used to process additional data and identify blocks of images that are likely to contain a target based on the degree of similarity to trained data [1, 3, 10–12].

Thorpe and colleagues [13] demonstrated the feasibility of pairing EEG with rapid image presentation by asking participants to classify nature scenes presented for 20 ms under a go/no-go paradigm. They found a frontal negativity specific to no-go trials that developed approximately 150 ms following stimulus onset. In the domain of intelligence analysis, Mathan and colleagues [3] used an EEG/RSVP approach with analysts examining satellite imagery. They showed that neurophysiologically driven image classification with rapid image presentation exhibits roughly a five-fold reduction in time required to identify targets relative to conventional image analysis, while retaining a high degree of accuracy. This technique has also been demonstrated using experts searching for masses in mammogram images [12].

The EEG signals used in these imagery triage applications are typically event-related potentials (ERPs). ERPs are obtained when an EEG signal is time-locked to a relevant stimulus [14]. In research settings, ERPs are often averaged across many trials in order to wash out noise (from sources such as eye blinks and facial muscle activity) that can overwhelm the ERP signals. However, the method of averaging across repeated trials is impractical for triage implementations where efficiency is of critical value. In such domains, promise lies in single-trial ERP detection which incorporates spatial information across EEG sensors [15, 16]. Such spatiotemporal EEG activity has revealed distinct patterns for target-present and target-absent images following stimulus presentation that could be exploited for purposes of constructing a single-trial ERP classifier [17].

One of the most useful ERPs for single-trial applications is the P300, or P3. The P3 refers to a positive deflection in voltage that occurs in the latency range of 250–500 ms, typically evoked using an oddball task in which an infrequent “oddball” target (e.g., an image containing a threat) is displayed within a series of frequent distractor stimuli (e.g., innocuous images), and the participant is asked to discriminate between target and non-target stimuli [18, 19]. P3 amplitudes are significantly larger in response to infrequent target items, though in order to evoke a P3 the task must force attention and categorization of stimuli [19]. There are thought to be two subcomponents of the P3, referred to as P3a and P3b, which have distinct neural generators that present as particular scalp topography in the EEG signal. The P3a subcomponent is thought to reflect stimuli-driven frontal attention mechanisms and is therefore maximal over frontal and central electrode locations, while the P3b subcomponent is associated with temporal and parietal lobe activity reflective of memory processing [18]. Therefore, one potential mechanism for the P3 wave as a whole is stimulus detection that engages memory processes [18].

Although the RSVP/EEG paradigm holds promise for helping professional visual searchers to triage imagery rapidly, it may be limited by the nature of the target items. Targets that do not vary a great deal in appearance are likely to elicit ERPs that can be classified by brain-computer interfaces, but more variable targets may not. In the present study, we sought to extend the RSVP/EEG paradigm to the domain of aviation security screening, and in doing so to explore the limitations of the technique for different types of targets. Airport screeners typically inspect X-ray images of baggage in search of threats, such as guns or explosive devices, and other prohibited items, such as flammable materials. As in the other domains in which the RSVP/EEG technique has been applied, the screeners must contend with large sets of imagery and time pressure while making high-consequence decisions. However, unlike domains such as mammography and satellite imagery analysis, the targets that are of interest to an aviation security screener can vary quite drastically in appearance and are sometimes deliberately concealed. In this study, we presented professional Transportation Security Officers (TSOs) with rapid successions of image chips taken from false color baggage X-rays in order to determine if various types of threat items could elicit P3 ERPs. We hypothesized that targets that have a prototypical appearance would elicit a useable P3 signal, but concealed targets or targets that do not have a prototypical appearance would not.

2 Method

2.1 Participants

Twelve individuals (3 female; mean age 32.7, range 21–63), currently working as TSOs with duties that include baggage screening, participated in this experiment and were paid for their time. All participants provided written informed consent and were right-handed, had no early exposure to languages other than English, had no history of neurological disease or defect, and possessed normal or corrected-to-normal vision and hearing.

2.2 Stimuli

False color X-ray images, created using the same types of scanners that are used in airport security checkpoints, were supplied by the Transportation Security Administration (TSA). These images were created by scanning actual pieces of luggage and were representative of the types of bags that are typically seen by TSOs at the airports. Each image presented a single piece of luggage (e.g., a briefcase, a duffel bag). For every piece of luggage there were two images, one showing a top view and one showing a side view. Some of the bags contained a prohibited item (threat bags), some contained no prohibited items (clear bags), and some threat bags were imaged again with the threat item removed (cleared threat bags). The threat bags contained one of two types of weapons, one of which is generally easier to detect than the other. We will refer to the two types of weapons as Threat A (easier to detect) and Threat B (more difficult to detect). The cleared threat bags were identical to the threat bags in all respects other than the absence of the threat item. The difficulty of the bags was rated by the TSA as easy, medium or hard, based on the amount of clutter in the bags and the types of concealment used for the threat items. Only bags rated as easy by TSA were used for this study.

Each of the false color X-ray images was decomposed into image chips and grouped into blocks of 50, with all images in a given group consisting of either 400×400 pixel chips (generated from images depicting the top view of luggage) or 400×250 pixel chips (generated from images depicting the side view of luggage). Within each block of 50 image chips, there were 49 distractor images taken from clear bags and one target item. The target image chip either contained a threat or the equivalent section of a cleared threat bag. For target images that contained a threat, the entirety of the prohibited item was presented in the image. Within each block of 50 images, all of the images were of the same type and were taken from the same quadrant of a bag. In other words, if the target image showed the top left corner of the top view of a suitcase, all of the distractors within that same block also showed the top left corner of the top view of a suitcase. If the target image was the bottom right corner of the side view of a backpack, all of the distractors showed the same quadrant and same view of other images of backpacks, and so forth.

A total of 10 blocks of images were used for training and 100 blocks of images were used in the main experiment. Of the 100 trials in the experiment, the target image

chip was a threat in 60 trials and a cleared threat in 40 trials. Given the finite number of images provided by TSA and the high number of image chips that were required to generate all of the trials, some of the distractor images were used more than once in different trials. Among the 5,500 image chips that were used (5,000 for the task trials and 500 for the 10 trials in the training block), there were 1,653 distractor image chips that appeared more than once. No target images were repeated, and the order of distractor repetition was balanced across participants.

2.3 Procedure for EEG Recording

The EEG was recorded from 128 silver/silver-chloride electrodes embedded in an elastic cap (ANT WaveGuard, “Duke” layout) using a high-impedance amplifier with active shielding. The electrodes were referenced on-line to the average of all electrodes. Following the experiment, the electrodes were re-referenced off-line to the average of the left and right mastoids. All of the electrodes were tested prior to recording in order to ensure that their impedance was below 50 KOhms. The EEG was digitized with a sampling rate of 256 Hz.

ERPs were computed at each electrode for each experimental condition by averaging the EEG data from 100 ms before the onset of an image chip until 920 ms after onset. Trials containing blinks, eye movement or muscle activity were excluded from the averages. The mean amplitude of the ERPs within time windows of interest was calculated using data digitally filtered off-line using a bandpass filter of 0.2 to 20 Hz.

2.4 Rapid Serial Visual Presentation (RSVP) Task

Participants were seated in a dimly-lit, sound-attenuating booth at a viewing distance of approximately 92 cm from a computer monitor with a refresh rate of 60 Hz. Trial presentation was consistent with similar RSVP studies [3, 10]. Each trial began with a fixation cross that was presented in the center of the screen for 1000 ms. Participants were instructed to keep their eyes on the fixation cross for the duration of its presentation, and to avoid blinking or moving their eyes during the subsequent presentation of images. The stimuli within each trial consisted of a group of 50 images that were presented serially against a white background in rapid succession. Following each set of images, participants were asked to indicate via a button press whether or not they believed a threat to be present in any of the images in that set. The participants were given 5 s to make their response. See Fig. 1 for an illustration of the trial structure.

During an initial training period of 10 trials, images were presented at the rate of 5 images/second (200 ms/image) and participants were given feedback following each trial regarding the accuracy of their response. Following training, presentation rate was set to 10 images/second (100 ms/image) and feedback was no longer provided. 100 trials were presented in this fashion; 60 trials contained a threat and 40 trials did not. Within each trial, image chips presented a consistent view and resolution; 50 trials consisted entirely of 400×400 pixel image chips displaying the front view of a bag, while 50 trials consisted of 400×250 pixel image chips displaying the side view

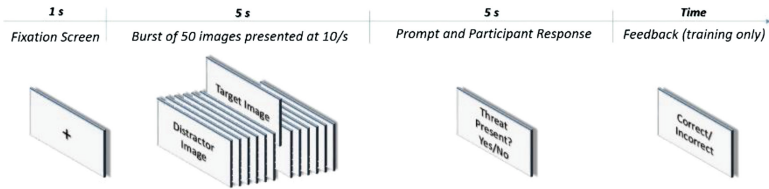


Fig. 1. Time-line of each RSVP trial. Note that response feedback was only provided during the training session.

of a bag. All target image chips were quasi-randomly inserted among the distractor stimuli, with the constraint that target chips were never presented within the first or last 500 ms (5 images) of a trial in order to prevent overlap with ERP signals related to the onset or offset of trials. Participants were given a self-paced break of up to one minute after every 10 trials in order to minimize potential eye strain and fatigue.

3 Results

3.1 Behavioral Results

The participants' average accuracy for each threat condition is shown in Fig. 2.

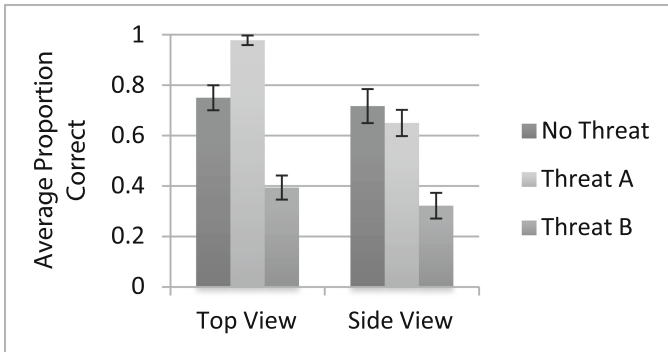


Fig. 2. Average proportion of correct answers for each threat condition

For Threat A, the threat with a stereotypical appearance, the participants responded correctly to an average of 98 % (SD = 3 %) of top-view trials and 65 % (SD = 20 %) of side-view trials. For Threat B, the threat without a stereotypical appearance, the participants responded correctly to an average of 39 % (SD = 18 %) of top-view trials and 32 % (SD = 19 %) of side-view trials. For the trials containing cleared threat bags (i.e. no threat), participants responded correctly to 75 % (SD = 16 %) of the top-view trials and 72 % (SD = 21 %) of the side-view trials. For the participants' average accuracy in each condition, 3 × 2 ANOVA (threat type by bag view) showed a significant main effect of threat type ($F(2, 22) = 26.04, p < 0.01$), a significant main effect of bag view ($F(1, 11) = 51.58, p < 0.01$), and a significant interaction between threat type and bag

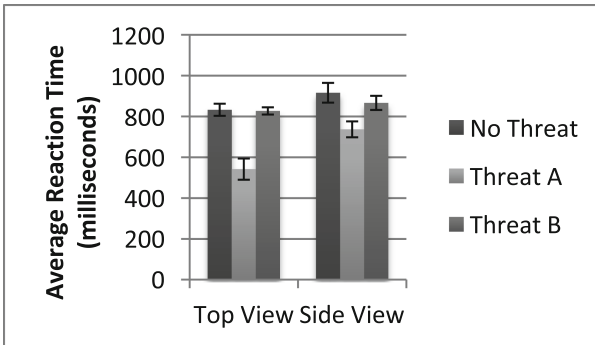


Fig. 3. Average reaction time for responses to each threat condition

view ($F(2,22) = 9.24, p < 0.01$). Pairwise comparisons between the threat conditions using paired t-tests showed that participants were significantly more accurate for Threat A trials than for Threat B trials, in both the top-view ($t(11) = 11.36, p < 0.001$) and side view ($t(11) = 5.92, p < 0.001$) conditions. In addition, participants were significantly more accurate for top-view than for side-view trials for both Threat A ($t(11) = 5.51, p < 0.001$) and Threat B ($t(11) = 2.00, p < 0.05$).

The participants' average reaction times for each threat condition are shown in Fig. 3. For the average reaction times in each condition, a 3×2 ANOVA showed a significant effect to threat type ($F(2,22) = 17.63, p < 0.01$) and a significant effect of bag view ($F(1,11) = 19.66, p < 0.01$). There was not a significant interaction between threat type and bag view ($F(2,22) = 2.25, p = 0.13$). Pairwise comparisons between the threat conditions using paired t-tests showed that participants responded significantly faster to top-view trials than to side view trials, and significantly faster to Threat A trials than to Threat B trials.

3.2 ERP Results

The ERPs were calculated by time-locking the EEG data to the onset of the target chip in each trial and averaging across trials in the same condition. The grand average ERPs were calculated by averaging across all trials in each condition for each participant. One participant's ERPs were excluded due to a large number of trials contaminated by blinks. For each threat type, the ERPs were compared for the threat and cleared threat bags. These stimuli were identical apart from the presence or absence of the threat in the bag. Representative ERPs from each scalp region are shown in Fig. 4 for Threat A and Fig. 6 for Threat B. Scalp maps showing all electrodes are shown in Figs. 5 and 7. For Threat A, the threat with a prototypical appearance, there were two ERP components that differed between the threat and cleared threat bags. The first was a positive peak over the front of the scalp, peaking at approximately 400 ms. The second was a positive peak over the central electrodes with a corresponding negative peak over the frontal electrodes, peaking at approximately 700 ms. For Threat B, the threat that does

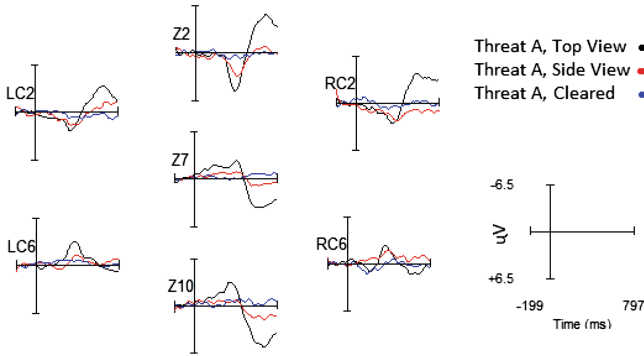


Fig. 4. Representative ERPs from electrodes in each of the seven scalp regions used in the analysis of Threat A.

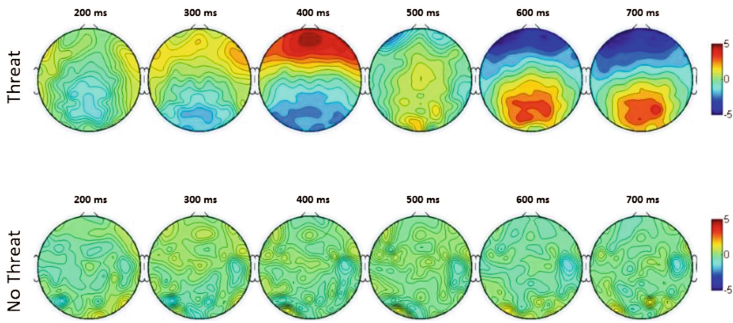


Fig. 5. Grand average ERP scalp maps for the threat and cleared threat trials for Threat A

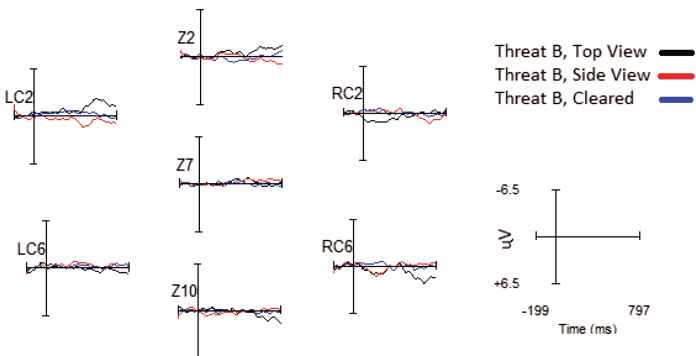


Fig. 6. Representative ERPs from electrodes in each of the seven scalp regions used in the analysis of Threat B.

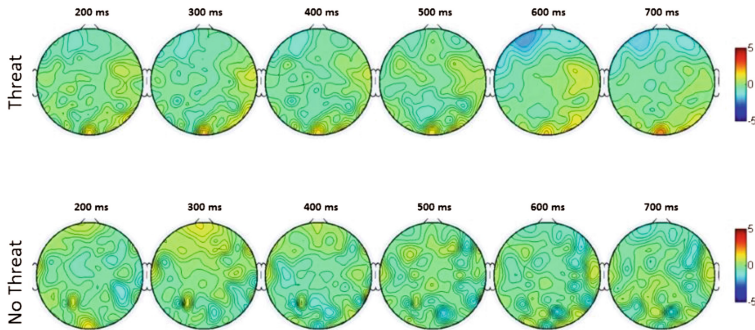


Fig. 7. Grand average ERP scalp maps for the threat and cleared threat trials for Threat B

not have a prototypical appearance, there were no ERP components that differed between the threat and cleared threat bags.

The ERPs were quantified for analysis by computing the mean amplitudes, post baseline correction, of the 300–450 and 600–800 ms intervals in the grand average waveforms. The electrodes were divided into seven scalp regions: left anterior, central anterior, right anterior, central, left posterior, central posterior, and right posterior. Repeated-measures ANOVAs were conducted for each of these time windows in the three central regions with the factors stimulus type (threat or match bag) and electrode site. For top-view Threat A trials, there were significant differences between the threat and match conditions in all three of the central scalp regions in both the 300–450 ms time window (all $F_s > 52.28$, all $p_s < 0.001$) and in the 600–800 ms time window (all $F_s > 24.94$, all $p_s < 0.001$). For side-view Threat A trials in the 300–450 ms time window, there were significant differences between the threat and match conditions in the central anterior and central posterior scalp regions (all $F_s > 18.77$, all $p_s < 0.001$), but not in the central scalp region ($F(1,24) = 2.72$, $p = 0.11$). For side-view Threat A trials in the 600–800 ms time window, there were significant differences between the threat and match conditions in the central anterior and central scalp regions (all $F_s > 8.29$, all $p_s < 0.01$), but not in the central posterior scalp region ($F(1,22) = 0.95$).

4 Discussion

Similar to satellite imagery analysts and radiologists, TSO bag screeners operate in a domain concerned with low-frequency, high-consequence targets buried among innocuous clutter. For satellite imagery analysts, the problem of image search centers on the vast number of continuously updating images in conjunction with an insufficient number of trained analysis [20] such that RSVP/EEG driven search offers the opportunity for otherwise un-reviewed images to be subjected to at least a cursory analysis. TSOs are not confronted with a vast image database in the same way that satellite imagery analysts are, in that every single item of luggage is screened. However, unless an item is flagged for further investigation, it is only viewed once by a single screener,

making this a domain that stands to benefit from a triage technique which would allow for an efficient double-checking scheme.

The aim of the current study was to examine the viability of constructing a neurophysiologically driven classifier within the domain of TSO baggage screening by determining if the basis for constructing such a classifier exists within an RSVP paradigm. A P300 effect was observed for threats with a stereotyped appearance (Threat A), while for trials containing a threat with a highly variable instantiation (Threat B) we did not observe a P300. Additionally, behavioral performance indicated that participants experienced difficulty detecting this more variable class of threats under RSVP conditions. Responses were more accurate when threats were of a stereotyped nature and when presented via top-down (as opposed to side) view. Even when participants correctly indicated the presence of a threat for a given trial, they may not have been basing their response on the critical image in the trial burst, such that a P300 is not timelocked to the chips of interest.

Currently, fully-automated systems are not viable within complex domains due to issues regarding specificity, sensitivity and ability to generalize [1, 3]. Leveraging the human perceptual system may facilitate generalization since brain responses may be specific to detection of attended targets independent of specific target features, thereby obviating the need to train a classifier that is sensitive to each individual target type. In a domain such as luggage screening in which the size, shape, orientation, and nature of targets may vary substantially and change over time, the flexibility of the human brain may continue to prove superior to fully automated methods of image classification. This is evidenced by research demonstrating that variability between images classes (i.e., target vs. distractor) may be low relative to variability within a class (i.e., target-to-target variability; [3]). It is worth noting that participants in the current study were not informed which types of prohibited items could be present in image blocks, but were simply asked to follow their standard operating procedure for identification of threats.

Our results suggest the possibility of implementing a triage technique within the domain of luggage screening. However, there are a number of important limitations to consider. Presentation of cropped or compressed images is typical of this body of research (e.g., [1, 3, 11, 12]), and the current study is no exception, utilizing cropped images (chips) rather than compressing the visually dense broad images in order to retain discriminability of image components. Rapid presentation rate is a necessity for triage techniques to maintain efficiency, but combination with image chips represents a double-edge sword. The pace of image presentation does not afford time for saccadic search of individual stimuli, which improves EEG signal to noise by minimizing artifactual eye movements. However, as targets become distal from the fixation point, detection rate may decrease [11, 17].

In addition, under RSVP conditions, participants have been shown to exhibit difficulty detecting targets that lay in the boundary between chips [3]. In the current study, each target was entirely contained within a particular image chip. This was intentional given the preliminary nature of the study, but in a real world setting, automatic image decomposition is highly unlikely to result in target items entirely falling within the boundary of the generated image chips. It is possible that overlapping image chips, as used in prior RSVP/EEG research [10] would enhance spatial context, thereby

mitigating the issue of boundary items, though such overlap results in an overall decrease in the efficiency of the triage system given that a greater number of images are needed to cover the same amount of image space. Efficiency may be further compromised by the need for frequent breaks to avoid mental fatigue or eye-strain. The current study offered a self-paced break of up to 1 min after every ten trials (500 images); additional work is necessary to determine at what point physical or mental fatigue becomes a factor.

The advantage to EEG provided by limiting eye movement is only valuable if the EEG signal itself is valuable. Recent research suggests that within an RSVP/EEG paradigm, the behavioral performance of participants tracks the detection of evoked response in the EEG signal [11]. In other words, image blocks that contain a target only elicit a distinct EEG signal when the participant is consciously aware that the image block contains a target, such that neuroimaging adds little value beyond overt behavioral information. This stands in contrast to Hope et al. [12] who demonstrated that receiver operating characteristic area under the curve increased from .62–.86 to .75–.94 when moving from a single electrode to multiple electrodes in an RSVP image triage paradigm. Likewise, Healy and Smeaton [21] demonstrated that using a mere 4 channels of EEG increases image classification accuracy by nearly 50 % beyond using only overt behavioral response. The current study did not attempt construction of a classifier or implementation of modeling for automatic detection of P300 s. Instances in which participants responded incorrectly (e.g., threat present to a threat absent trial) are grouped by response and grand averaged such that instances in which a threat is subconsciously detected may have washed out. Behaviorally, there were not enough instances of false negatives to allow for analysis of a potential subconscious P300. Although the current work was unable to evaluate brain responses on a single-trial basis, it does suggest that for certain items a neural response is elicited which may allow for future construction of a classifier capable of automatic peak identification, thereby allowing neurophysiology to identify the presence of threats in a way not captured by behavioral responses. It is also important to note that such a classifier may provide a better basis for localization of target-containing images within a sequence due to the high temporal resolution associated with ERPs relative to the substantial latency inherent in motor responses [3].

It is currently unknown if task familiarity plays a role in image triage performance within this domain. In the current study, we tracked the amount of time each TSO had spent working in the capacity of a baggage screener in order to determine if job experience related to ability to accurately identify targets in a domain-specific RSVP paradigm. However, all participants were naïve to high throughput analysis of images as experienced in this study, and it is possible that training within this paradigm would result in enhanced ability to discriminate between target and non-target blocks of images. Individual differences may also play a role, as previous investigation has demonstrated that a slower rate of presentation may be necessary in order to attain an acceptable level of accuracy for some individuals [17].

Given the equipment expense and time cost of setup and analysis of EEG data, it is important to determine the extent to which EEG provides a benefit above and beyond overt behavioral data, and if task practice and/or identification of individuals adept at high throughput screening may obviate the need for neurophysiological data. While the

current study utilized a 128 channel EEG system, other work has found a small number of electrodes to be sufficient for substantial increases in classification accuracy [12, 21] such that low cost consumer-grade EEG systems may prove a viable option.

References

1. Huang, Y., Erdogmus, D., Mathan, S., Pavel, M.: Large-scale image database triage via EEG evoked responses. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 429–432 (2008)
2. Collins, R.T., Lipoton, A.J., Kanade, T., Fujiyoshi, H., Duggins, D., Tsin, Y., Wixon, L.: A system for video surveillance and monitoring: VSAM final report (CMU-RI-TR-00-12). Carnegie Mellon University, Robotics Institute. (2000). http://www.ri.cmu.edu/pub_files/pub2/collins_robert_2000_1/collins_robert_2000_1.pdf
3. Mathan, S., Whitlow, S., Dorneich, M., Ververs, P., Davis, G.: Neurophysiological estimation of interruptibility: demonstrating feasibility in a field context. In: Schmorow, D. D., Nicholson, D.M., Dexler, J.M., Reeves, L.M. (eds.) Foundations of Augmented Cognition, 4th edn, pp. 51–58. Strategic Analysis, Arlington (2007)
4. Kranczioch, C., Debener, S., Herrmann, C.S., Engel, A.K.: EEG gamma-band activity in rapid serial visual presentation. *Exp. Brain Res.* **169**, 246–254 (2006)
5. Intraub, H.: Rapid conceptual identification of sequentially presented pictures. *J. Exp. Psychol. Hum. Percept. Perform.* **7**(3), 604–610 (1981)
6. Harris, C.L., Morris, A.L.: Illusory words created by repetition blindness: a technique for probing sublexical representations. *Psychon. Bull. Rev.* **8**(1), 118–126 (2001)
7. Vroomen, J., Driver, J., de Gelder, B.: Is cross-modal integration of emotional expressions independent of attentional resources? *Cogn. Affect. Behav. Neurosci.* **1**(4), 382–387 (2001)
8. Chun, M.M., Potter, M.C.: A two-stage model for multiple target detection in rapid serial visual presentation. *J. Exp. Psychol. Hum. Percept. Perform.* **21**(1), 109–127 (1995)
9. Mantini, D., Marzetti, L., Corbetta, M., Romani, G.L., Del Gratta, C.: Multimodal integration of fMRI and EEG data for high spatial and temporal resolution analysis of brain networks. *Brain Topogr.* **23**(2), 150–158 (2010)
10. Bigdely-Shamlo, N., Vankov, A., Ramirez, R.R., Makeig, S.: Brain activity-based image classification from rapid serial visual presentation. *IEEE Trans. Neural Syst. Rehabil. Eng.* **16**(5), 432–441 (2008)
11. Dias, J.C., Parra, L.C.: No EEG evidence for subconscious detection during rapid serial visual presentation. *IEEE Signal Process. Med. Biol. Symp.* 1–4 (2011)
12. Hope, C., Sterr, A., Elangovan, P., Geades, N., Windridge, D., Young, K., Wells, K.: High throughput screening for mammography using a human-computer interface with rapid serial visual presentation (RSVP). In: Proceedings SPIE Medical Imaging 2013: Image Perception, Observer Performance, and Technology Assessment, vol. 8673, (2013). doi:[10.1117/12.2007557](https://doi.org/10.1117/12.2007557)
13. Thorpe, S., Fize, D., Marlot, C.: Speed of processing in the human visual system. *Nature* **381**, 520–522 (1996)
14. Luck, S.J.: An introduction to the event-related potential technique. MIT press, Cambridge (2014)
15. Gerson, A.D., Parra, L.C., Sajda, P.: Cortical origins of response time variability during rapid discrimination of visual objects. *Neuroimage* **28**(2), 342–353 (2005)

16. Parra, L., Alvino, C., Tang, A., Pearlmutter, B., Yeung, N., Osman, A., Sajda, P.: Single-trial detection in EEG and MEG: keeping it linear. *Neurocomputing* **52**, 177–183 (2003)
17. Mathan, S., Whitlow, S., Erdogmus, D., Pavel, M., Ververs, P., Dorneich, M.: Neurophysiologically driven image triage. In: *Proceedings of the 2006 Conference on Human Factors in Computing Systems*, pp. 1085–1090 (2006)
18. Polich, J.: Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* **118**, 2128–2148 (2007)
19. Luck, S.J., Kappenman, E.S. (eds.): *The Oxford handbook of event-related potential components*. Oxford University Press, Oxford (2011)
20. Kenyon, H.S.: Unconventional information operations shorten wars. *Signal Mag. Armed Forces Commun. Electron. Assoc.* **57**, 33–36 (2003). <http://www.afcea.org/content/?q=node/167>
21. Healy, G., Smeaton, A.F.: Optimising the number of channels in EEG-augmented image search. In: *Proceedings of the 25th BCS Conference on Human-Computer Interaction*, pp. 157–162 (2011)