Adaptive Training for Visual Search

Kelly S. Hale, Angela Carpenter, Matthew Johnston, Jing-Jing Costello, Jesse Flint
Design Interactive, Inc.
Oviedo, FL
kelly@, angela@, matthew@, jingjing@, jesse.flint@designinteractive.net

Stephen M. Fiore
University of Central Florida
Orlando, FL
sfiore@ist.ucf.edu

ABSTRACT

Effective training is a vital foundation for transportation security officers required to learn strategies for identifying anomalies within X-ray images that may indicate a potential threat. Past research has shown that adaptive training is a powerful tool to increase detection performance, however, adaptive training strategies in this domain have typically utilized exposure training techniques exclusively. This paper outlines the science behind adaptive training for anomaly detection, including (1) real-time advanced performance measures associated with visual search tasks and (2) training strategies to target identified root cause(s) of error. Specific strategies discussed in this paper include exposure training and discrimination training to optimize training within the baggage screening domain. A proposed adaptive training framework and resulting system is presented.

Empirical results from a preliminary investigation into the benefits of adaptive training are presented. Thirty novice participants completed a mixed between and within design, where independent variables were training strategy (Traditional or Adaptive) and test session (Session 1, Session 2, Session 3), and dependent variables were sensitivity (d'), response criterion (c), hit rate, false alarm rate, miss rate, response time, and gaze data. In addition, eye tracking data from 4 experts was collected to evaluate differences in scan patterns and visual search strategies between novices and experts. Results showed repeated training in either group improved performance in terms of a decrease in the number of threat items missed and response time. Traditional training resulted in greater sensitivity and fewer false alarms in early training sessions. Gaze data showed that overall dwell time is positively related to the clutter density for the expert group. Analyses are ongoing to examine additional search strategy data (e.g., saccade distance, direction, changes in visual search direction, etc.) to further quantify distinct patterns in eye scan behavior to define novice versus expert performance. Future research will include further investigation into Exposure and Discrimination training to quantify benefits of each training strategy, which can better inform when and how to adapt training over time to target individualized deficiencies/inefficiencies and increase training effectiveness and efficiency. Additionally, future research should consider a longer training period, as current results did not show performance stabilization, indicating that learning may still be occurring.

ABOUT THE AUTHORS

Kelly Hale is Sr. Vice President of Technical Operations at Design Interactive, Inc., and has over 12 years experience in human systems integration research and development. Her R&D efforts are focused in augmented cognition, adaptive, personalized systems, multimodal interaction, training sciences, and virtual environments. Through these efforts, Kelly and her team have developed advanced neurophysiological measurement techniques and have advanced real-time mitigation strategy framework and induction techniques to optimize training, situation awareness, and operational performance through optimization of user cognitive and physical state. She received her BSc in Kinesiology/Ergonomics Option from the University of Waterloo in Ontario, Canada, and her Masters and PhD in Industrial Engineering, with a focus on Human Factors Engineering, from the University of Central Florida.

Angela M. (Baskin) Carpenter is a Research Associate II at Design Interactive, Inc. Her work has focused on using multimodal design science to optimize operator situational awareness and workload in C4ISR environments, development of neuro-physiological metrics to assess signal detection, and assessment of astronaut cognitive state via a hand-held game. She received B.A.s in Psychology and Spanish/Latin American Studies from Flagler College in 2003, and her Master of Science in Human Factors and Systems with a dual human factors and systems engineering track,from Embry-Riddle Aeronautical University in 2005.

Matthew Johnston is Director of Emerging Markets and Technology at Design Interactive Inc His R&D focus includes cognitive readiness assessment tools, next generation human systems, adaptive training systems and he is the company lead for commercial product evaluation and usability testing. He has led numerous efforts involved in furthering the field of human factors engineering including the development of a tactile communication system for dismounted soldiers and an online, game based cognitive assessment and diagnostic tool (CogGauge). Matthew came to DI with a significant background in consumer product evaluation and usability at Ford Motor Company and Nortel Networks. He received a BSc from the University of Waterloo in Kinesiology/Ergonomics Option and an MSc from Loughborough University in England in Ergonomics.

Jingjing Wang Costello is a Research Associate at Design Interactive, Inc. Since she joined DI, her research has focused on adaptive training, training system design, and human computer interaction. She has research experience in situation awareness and synthetic speech comprehension of native vs. non-native speakers. She holds a Ph.D. from the University of Central Florida in Applied Experimental Psychology and Human Factors, a MS from Syracuse University in Telecommunications and Network Management, and a BS from Syracuse University with a focus on Information Science and Technologies.

Jesse Flint is a Research Assistant for Design Interactive and has a background in Cognitive Psychology with a focus on auditory sensation and perception. His interests include the application of signal detection models, loudness perception, implicit learning, memory, attention, visual search, sound localization, and perception of complex natural sounds. He earned a bachelor's degree in Psychology at Stony Brook University and currently holds a Master's degree in Cognitive Psychology from Binghamton University. He is also a doctoral candidate in Cognitive Psychology at Binghamton University.

Stephen M. Fiore, Ph.D., is faculty with the University of Central Florida's Cognitive Sciences Program in the Department of Philosophy and Director of the Cognitive Sciences Laboratory at UCF's Institute for Simulation and Training. He earned his Ph.D. degree in Cognitive Psychology from the University of Pittsburgh, Learning Research and Development Center. He maintains a multidisciplinary research interest that incorporates aspects of the cognitive, social, and computational sciences in the investigation of learning and performance in individuals and teams. He is co-Editor of recent volumes on Macrocognition in Teams (2008), Distributed Learning (2007), Team Cognition (2004), and he has co-authored over 100 scholarly publications in the area of learning, memory, and problem solving at the individual and the group level.

Adaptive Training for Visual Search

Kelly Hale, Angela Carpenter, Matthew Johnston, Jing-Jing Costello, Jesse Flint
Design Interactive, Inc.
Oviedo, FL
kelly@, angela@, Matthew@, jinging@, jesse.flint@designinteractive.net

Stephen M. Fiore
University of Central Florida
Orlando, FL
sfiore@ist.ucf.edu

INTRODUCTION

Baggage screening is a repetitive visual search task that often has a very low probability of encountering a threat, but extremely high consequences if a serious threat is missed. Due to the importance of screening accuracy, screeners are required to complete extensive training both before going on the job and while employed. Software-based training systems are often limited to observable behavioral metrics (e.g., detections, false alarms), and are thus limited in their ability to identify root cause(s) of visual search performance errors (i.e., scan vs. recognition error).

To address these limitations, a prototype training system was developed that incorporates real time analysis of performance through a diagnostic module that uses eye tracking to provide insight into trainee knowledge and skill. The prototype system results in training that adapts training strategy (exposure or discrimination training), training content (image attributes), and training difficulty level (distribution of specific image attributes across the session) based on individual needs, with the goal of enhancing visual search learning. As outlined by Sireteanu and Rettenbach (1994), this does not simply mean improvement in perceiving any particular feature or combination of features, but in improving higher order pattern recognition and search strategy, which leads to increased accuracy and throughput.

BACKGROUND

Perceptual learning is not just about taking in visual cues, but involves meaningful integration of what is perceived visually (Hoffman & Fiore, 2007). While this process is initially slow and inefficient, studies have shown that perceptual ability can be enhanced through training (Seitz & Dinse, 2007). This enhanced skill may be the result of both bottom-up and top-down processes such as improved signal discrimination and top-down biasing signals (Baluch & Laurent, 2010). Further, research involving radiologists has shown experts employ a global search phase early in visual search, allowing them to better interpret the image as a whole and locate potential anomalies for further focus and identification (Kundel, Nodine, Conant & Weinstein, 2007).

Training Methods

Training such skill has typically been accomplished through exposure training, where trainees are shown a variety of images and asked to identify whether an object of interest (i.e., threat) is present. This 'mass exposure' technique can lead to improved performance and automaticity. For example, one week of exposure training to find a specific color coded stimuli resulted in increased accuracy and reaction time, as well as significant changes in brain activity, reflecting increased visual cortex processing with decreased attention (Greenlee, Frank, Reavis, & Tse, 2011).

An alternative training method is discrimination training, which involves pairs of targets with or without salient differences presented in two separate side-by-side bag images - the task is to determine whether threat items within each image are the same or different (Fiore et al., 2006). The degree of similarity of these items is varied, thus providing a range of difficulty, with some requiring a careful visual interrogation to determine whether threats are the same or different. Under this training paradigm, it is theorized learning results from development of stimulus-specific knowledge through repeated exposure, as well as development of strategic skills through making comparisons (Doane et al., 1999). Previous studies in the context of baggage screening have found that discrimination training that incorporates holistic threat training (Schuster et al., 2010) and training under higher levels of bag complexity (Sellers et al., 2010) result in improved performance, both in terms of discrimination performance (d′) and response time. Further, in addition to enhanced performance, individuals reported less workload during tests when training contained difficult discriminations in the presence of complex stimuli (Fiore et al., 2006).

An additional training challenge within this community is preparing for the low prevalence of threats. Only approximately 2% of bags screened contain actual threats (Wolfe, Horowitz, and Kenner, 2005). In such low prevalence visual search tasks, miss rates increase dramatically (Wolfe, Horowitz, and Kenner, 2005). Attempts to correct the increase in miss rates have resulted in changes to response criterion (c) rather than sensitivity (d′) (Wolfe and Van Wert, 2010), indicating that participants answer "yes" (there is a threat) more often. This leads to a decrease in the miss rates, but only at the cost of increasing false alarm rates. Wolfe et al., (2007) demonstrated that providing feedback during 60 trial bursts of higher prevalence also resulted in a lasting shift in c during extended sessions of low prevalence. Other studies, however, have found that training environment complexity using exposure methods can influence target detection sensitivity. For example, Fiore, Scielzo and Jentsch (2004) showed that increasing the complexity of the training environment, by adding clutter, improves sensitivity to target detection, but does so dependent upon spatial abilities and test item difficulty. Thus, the current study evaluates both sensitivity and response criterion to evaluate this potential increase in false alarms that may occur in conjunction with decreased misses.

Neurophysiological Measures of Performance

Eye tracking has been used in numerous studies to better understand visual search and pattern perception, particularly to distinguish differences between novices and experts. Capturing fixation patterns provides a *process* level measure of visual search, and can be used to differentiate subtle differences in how one approaches the task that is not otherwise observable. For example, previous studies have shown that experts have larger visual spans as represented by a greater number of fixations between objects of interest (e.g., spaces between pieces on a chess board) and greater number of fixations on objects that are relevant to decision outcome compared to novices (Charness, Reingold, Pomplun & Stampe, 2001; Reingold, Charness, Pomplun & Stampe, 2001). Further, Baluch and Laurent (2010) found a decrease in intersaccadic interval (i.e., time between two successive saccades), but no change in saccade count, which was interpreted as an improvement in discrimination and selection focus in terms of 'quality' as opposed to quantity of items scanned.

Such eye movement data can also be used to identify root cause(s) of search errors, for example whether a trainee failed to fixate on the object of interest (e.g., threat) or whether they fixated, but failed to recognize the object as a threat (Carroll, Fuchs, Hale, Dargue & Buck, 2010). Thus, a more detailed understanding of performance breakdowns can be realized by further decomposing traditional signal detection theory outcomes (Hit, Miss, False Alarm, Correct Rejection). For example, if an image is classified as a Miss, did participants look at the missed threat or not? If no, then training may focus on scan patterns and overall search strategy to locate potential anomalies. If yes, then training may utilize discrimination training to focus trainees on specific details of anomalies/threats.

REAL-TIME, ADAPTIVE TRAINING SYSTEM

A prototype real-time adaptive training system for baggage screening, ScreenADAPT, was developed to incorporate the benefits of both exposure and discrimination training with detailed root cause analyses based on eye scan and behavioral data to optimize learning for individual trainees. This system integrates test and training sessions to provide an individualized training paradigm that increases image and threat difficulty while focusing on specific underlying root causes of inefficient/deficient performance, such as inefficient search, inability to locate anomalies, and inability to correctly identify anomalies. Within targeted training strategies, feedback is provided immediately for each image so that stimulus and response associations may be formed due to their simultaneous occurrence in time (based on Guthrie, 1935) to accelerate the acquisition of target skills without detrimental effect on learning or retention (Corbett & Anderson, 1991). Subsequent test sessions provide feedback in summary form at the end of the session in an after action review (AAR) to provide the trainee with a summary of individual strengths and areas for improvement, and to avoid the potential "mindlessness" of continual immediate feedback (Anderson, 1970), or its use as a crutch (Druckman & Bjork, 1991). The AAR includes performance details like percent correct, most prevalent error types and suggestions for improving, average response time, and examples of performance and eye scanning errors.

Ten difficulty levels, which vary based on image difficulty (e.g., threat orientation, location, type; presence of distractors, clutter), threat-to-non threat ratio and stimulus presentation time, are designed to step a user through the training process and increase his/her individual expertise level over time. The images used in training sessions are tailor-made on-the-fly to focus on attributes that are most challenging for the user while maintaining representation of all image attributes within the set. Generating images on-the-fly rather than pulling images from a pre-set library serves to reduce the possibility of presenting the user with the same image more than once. With continued training, a user may repeat a training level indefinitely (i.e., highest level), requiring a large number of differing images, and self-generating images ensures a unique, challenging image set each training session.

METHOD

Participants

Thirty novice participants (15M; 15F) ranging in age from 18 to 46 completed this study. All met minimum recruitment requirements for the Transportation Security Administration (TSA). In addition, 4 Transportation Security Officers (TSOs) completed two sessions of Exposure Training to provide a representative sample of expert eye scan data for preliminary comparison evaluation.

Apparatus

The experimental setup was designed to mimic, though not replicate exactly, a TSO's workstation, and consisted of a computer controlling an LCD display and eye tracking hardware. Images were presented full-screen on a 17" LCD display set at 1280x1024 resolution placed approximately 60 cm away from participants at eye level. Positioned directly below the display was an easyGaze® eye tracker – a stand-alone, non-intrustive unit that utilizes Near-Infrared (NIR) Light-Emitting Diodes to generate even lighting and reflection patterns in the eyes of the user. The system collected a variety of time-stamped quantitative gaze data simultaneously from both eyes at a frequency of 50Hz and a spatial resolution of 0.25 degrees. Participants responded using a mouse and keyboard. It was expected that data collected on this simulated TSO workstation is representative of novice operational behavior.

Task Stimuli

A set of 60 X-ray representative passenger bag images (no threats) combined with a set of 28 threat (gun or knife) and 105 distractor items (some intentionally similar to threats in order to balance difficulty of non-threat images, e.g. hair dryer to correspond to gun) generated from publically available 3D model imagery (manipulated to mimic X-ray view) were utilized in this study. Each image was generated using an in-house image generator software to insert distractors and/or threats into X-ray images of representative passenger carry-on luggage at different positions and different angles within the images. Within the Traditional Training condition, a pre-defined set of images were used across all participants. During testing sessions, a ratio of 1:1 threat-to-clear bags was used, and a ratio of 2:1 threat-to-clear bags was used for the training. Within the Adaptive Training condition, testing sessions also utilized a 1:1 threat-to-clear bag ratio. But Adaptive Training incorporated two distinct training methods, each which used a different threat-to-clear bag ratio. Exposure training used 2:1 ratio (same ratio as Traditional Training). Discrimination Training used 100% threat bags (i.e., two images presented simultaneously contained a threat – participant was to determine whether they were the same or different). Training within the adaptive condition (both exposure and discrimination) used results from the previous testing session to select images attributes for a given training session's focus.

Procedure

When participants arrived, they were escorted to the testing room and provided an informed consent document to review and sign, as well as a demographic questionnaire. Participants were seated in front of the display, and the eye tracker was then calibrated. Participants were given written instructions that outlined what constituted a threat for this experiment, as well as instructions on how to operate the testbed, followed by a practice session. The practice session consisted of 10 trial images (5 threat and 5 non-threat images). Once participants completed the practice session correctly, they completed a baseline test that included 80 images. They were then randomly assigned to one of two training groups. The Traditional Training group received fixed content exposure training of 50 images in three successive sessions. The Adaptive Training group received customized training content of 50 images based on (a) training strategy implemented and (b) training content across three successive sessions. After each training session, a post-test was completed that consisted of 80 images. After the experiment was completed, participants were compensated for their time.

To evaluate the utility of using gaze data in addition to behavioral responses to identify root cause of error during visual search in future experiments, 4 TSO participants completed eye tracking calibration, a practice session, and two of the four Exposure Training conditions. This allowed for direct comparison to novice data collected.

Experimental Design

This training paradigm study was a 2x3 mixed between and within study design. The between-group independent variable was Training type (Exposure Training vs. Adaptive Training) and the within-group independent variable was training session. Dependent variables are based on a Signal Detection Theory analysis and an eye tracking analysis. H (hit rate) was defined as the ratio of the number of trials where the threat was correctly identified compared to the total number of trials in which a threat was present. FA (false alarm rate) was defined as the ratio of the number of trials in which a threat was identified by the participant when there was no threat present compared to the total number of trials in which no threat was present. M (miss rate) was defined as the ratio of the number of trials in which the participant failed to identify a threat compared to the total number of trials in which a threat was present. Differences in sensitivity (d′) was calculated as $d' = z(H) - z(FA)$. Criterion (c) is a measure of bias and indicates how willing the participant is to say that there is a threat present. Criterion is defined as $c = -½[z(H) + z(FA)]$.

Hypotheses

- H1: Training with an Adaptive system that provides tailored training will result in significantly higher performance outcomes compared to Exposure Training, specifically:
  - Significant decrease in miss rates across repeat exposure,
  - Significant decrease in criterion (c) across repeat exposure.
- H2: Training with an Adaptive system that provides tailored training will result in significantly shorter time to reach performance criterion (i.e., time to learn to criterion) compared to Traditional Training.
- H3: Novice trainees will show significant differences in eye tracking metrics compared to expert trainees, specifically:
  - Novices will show significantly longer dwell time on AoIs.
  - Time to first fixation on a threat will be significantly shorter for experts compared to novices.

Data Analysis

A 2x3 mixed effects repeated measures ANOVA was used to test for statistical significance with type of training (between subjects) and test session (within subjects) as independent variables and pre-post differences in performance outcomes as dependent variables.

RESULTS

A significant difference was evident in sensitivity change (Δd') from pre-training to post-training for type of training, $F = 4.717$, $p < .05$, $ή^2 = .555$, where Traditional training showed significantly greater average delta across all training sessions (Figure 1). No significant difference was found across test sessions for change in sensitivity (d') from pre-testing.

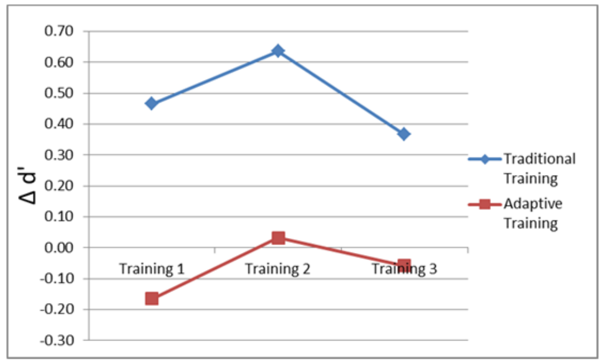

Figure 1. Change in sensitivity (Δ d′) across sessions

A significant difference was evident in miss rate change (ΔM) from pre-training to post-training across test sessions, $F = 9.398$, $p < .01$, $ή^2 = .973$, where the change in miss rate was lower in Session 2 and 3 compared to Session 1 (Figure 2). No significant differences were found for type of training.

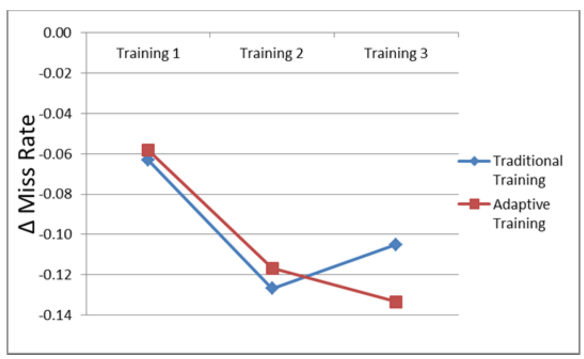

Figure 2. Change in miss rate across sessions

A significant difference was also evident in false alarm rate change (ΔFA) from pre-training to post-training for the main effect of type of training, $F = 9.118$, $p < .05$, $ή^2 = .830$, with Traditional training showing significantly lower increases in false alarm rates from pre-training scores across all sessions (Figure 3). There was no significant difference across test session.

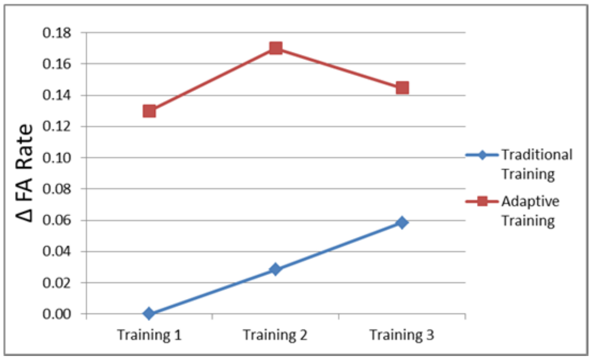

Figure 3. Change in false alarm rates across sessions

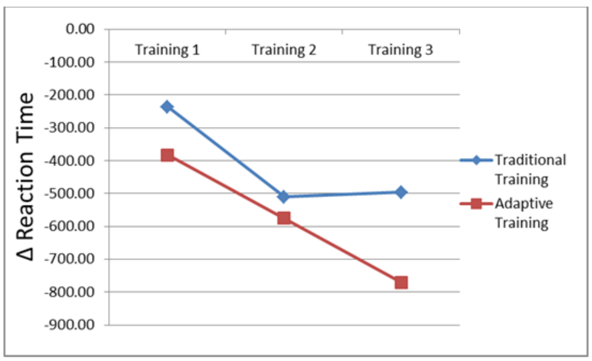

Figure 4. Change in reaction time across sessions

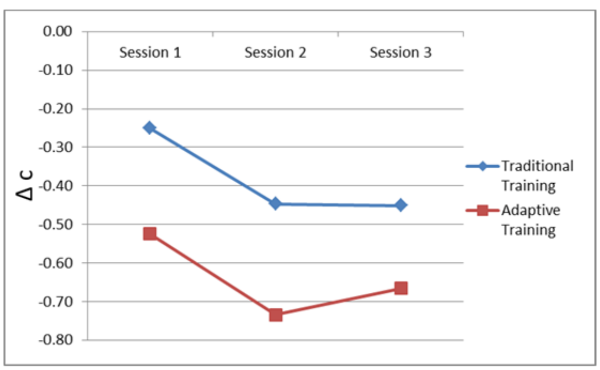

Figure 5. Change in criterion (Δ c) across sessions

A significant difference was found in reaction time change from pre-training to post-training across training sessions, $F = 22.684$, $p < .01$, $\acute{\eta}^2 = .996$., where greater negative deltas were evident in Session 2 and 3 compared to session 1 (Figure 4). No significant differences were found for type of training.

A significant difference was found in criterion change (Δc) from pre-training to post-training across training sessions, $F = 17.360$, $p < .01$, $\acute{\eta}^2 = 1.0$, where both session 2 and 3 showed a significant decrease in delta c compared to session 1 (Figure 5). A significant difference was also found across type of training, $F = 5.117$, $p < .05$, $\acute{\eta}^2 = .589$, where Adaptive Training had significantly higher decreases in delta criterion across training sessions compared to Traditional Training.

Eye Tracking Evaluation

Collecting novice and expert scan data allowed for an initial analysis of the utility of gaze data in further refining root cause error analysis for visual search within a real-time adaptive system as used in the current study. A number of variables, including time to first fixation on the threat, average fixation duration on threats, number of fixations on threats and response time to classification, were tested between the experts and novice groups. A significant difference was found in average fixation duration between experts and novices (t=4.59, df = 105.282, $p < 0.05$). As shown in Figure 6, experts showed a higher average fixation duration on AoIs compared to novices. Time to first fixation on threat, time to classification, number of fixations on threat, and total dwell time on threat did not show any significant differences between novices and experts.

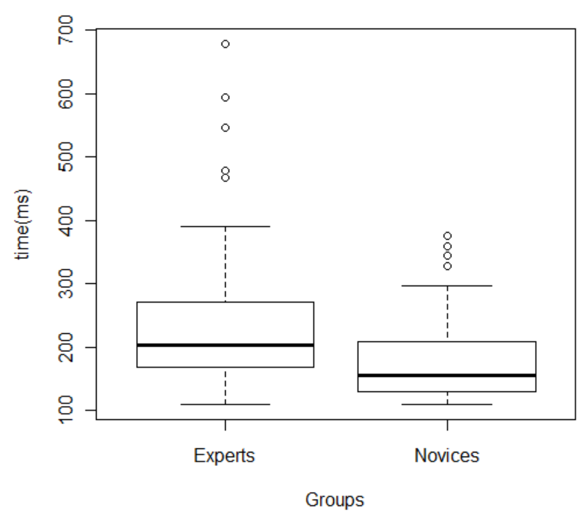

Figure 6. Average fixation duration on threats

Overall dwell time on threats was linearly related to contour density (a clutter measure) for the expert group (p= 0.073, p-value for contour density = 0.0273, meaning this factor is significant, but it is only meaningful when the model itself is significantly fit to the data; Figure 7). No such relationship was found in the novice group.

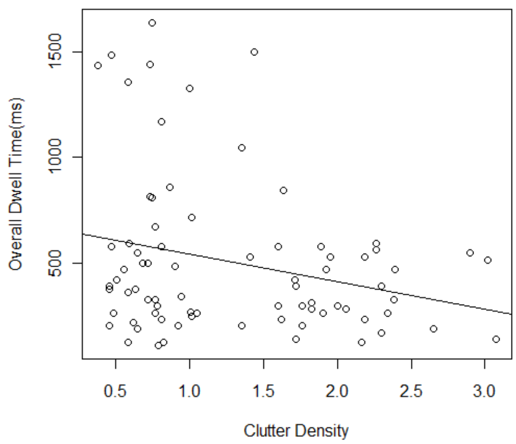

Figure 7. Fitted regression model for overall dwell time and contour density

DISCUSSION

Training in general resulted in a significant impact on accuracy and throughput demonstrated by a reduction in miss rates and response time. The training provided in the current study produced the same lasting change in c as observed in the Wolfe et al. (2007) study. The change (Δ) in c for Adaptive Training was greater than that for Traditional Training, indicating Adaptive Training resulted in fewer misses than Traditional Training. However, results also showed that Traditional Training resulted in greater sensitivity (d′) and lower false alarm rates than Adaptive Training for the three training sessions completed in this study. The trends (though not significant) for d′ and false alarm rates reverse for Adaptive training between session 2 and 3. Thus, with additional training sessions, false alarm rates may continue to decrease and d′ may increase due to training with more, varied threats in successive training sessions. Together, these effects could lead to an overall increase in sensitivity due to Adaptive Training.

Based on findings here and those reported elsewhere (Wolfe, 2007), changes in sensitivity appear to be more difficult to induce than changes in response bias (Δc) during initial training trials. The resulting change in c has the desired effect of decreasing miss rates, but only at the cost of also increasing false alarm rates. Further investigation using a longer training time (i.e., more sessions) may provide further insight into the long term impact of training on false alarm rate, as a low miss rates, at the expense of an increased false alarm rates, is not an ideal solution for aviation security.

Eye Tracking

Behavioral differences between novices and experts have been measured by eye tracking search patterns, percentage of time looking at AoIs, and fixations (Kurland et al., 2005), where experts tend to visually process faster (i.e., shorter fixation duration) and move in shorter jumps from location to location. Results from this study showed experts have *longer* average fixation duration on threats, which is in line with previous findings from intelligent imagery analysis studies (Hale et al., 2008), yet contradicts that from Kurland et al. (2005). Additional data collection is planned to increase the number of expert data points to further investigate eye tracking metrics for real-time adaptive training of visual search tasks and identify additional significant differences in scan and search strategies that can be evaluated in real-time and used to tailor training.

The result that the overall dwell time on the image is negatively related to the contour density for the expert group (Figure 7) may be partially explained by the findings of Lohrenz and Beck (2010) that suggests novices avoid searching in highly cluttered regions of displays. Additional analyses are ongoing to examine other search strategy data (e.g., saccade distance, direction, changes in visual search direction, etc.) to further quantify distinct patterns in eye scan behavior to define novice versus expert performance.

FUTURE WORK

A follow on study is planned to examine the impact of Exposure versus Discrimination training in isolation to quantify benefits on false alarm rate, miss rate, sensitivity, and response criterion. Based on current findings, it is anticipated that Exposure Training targeted on specific individual deficiencies/ inefficiencies, will reduce misses and increase false alarm rates initially. With repeated training sessions, it is anticipated that false alarm rates will peak and then drop as trainees are exposed to more, varied threats and provided immediate feedback via training sessions. Discrimination Training, while shown to improve hit rates, may also lower false alarm rates through implicit learning, as trainees are focused on specific details of threats through comparison evaluations, and are provided immediate feedback regarding their performance during training. Through quantification of individual benefits of each training paradigm on performance, it is anticipated that an improved Adaptive Training paradigm may be created that optimizes training for individuals, improving both response criterion and sensitivity over repeated training trials.

Further, integration of eye gaze metrics into the real-time adaptive training system is planned to further breakdown the diagnosis of scan errors and tailor training. Future studies will examine the benefit of additional diagnosticity in improving training effectiveness and efficiency.

ACKNOWLEDGEMENTS

This material is based upon work supported in part by the Department of Homeland Security (DHS) under its SBIR program. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views nor the endorsement of DHS.

REFERENCES

Carroll, M., Fuchs, S., Hale, K., Dargue, B. & Buck, B. (2010). Advancing Training Evaulation System: Leveraging neuro-physiological measurement to individualize training. *Proceedings of the Interservice/Industry Training, Simulation & Education Conference (I/ITSEC 2010)*.

Charness, N., Reingold, E.M., Pomplun, M. & Stampe, D.M. (2001). The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memory & Cognition*, 29*0), 1146-1152.

Fiore, S.M., Scielzo, S., Jentsch, F. (2004). Stimulus competition during perceptual learning: training and aptitude considerations in the X-ray security screening process. Content article for special section in Cognitive Technology, 9, 34-39.

Fiore, S.M., Scielzo, S., Jentsch, F. & Howard, M.L. (2006). Understanding performance and cognitive efficiency when training for X-ray security screening. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*, pp.2610-2614.

Greenlee, M.W., Frank, S.M., Reavis, E.A. & Tse, P.U. (2011). From inefficient to pop-out visual search in one week. BIO web of Conferences 1. Available online at: https://www.google.com/search?q=From+inefficient+to+pop-out+visual+search+in+one+week&ie=utf-8&oe=utf-8&aq=t&rls=org.mozilla:en-US:official&client=firefox-a. Viewed May 23, 2011.

Hoffman, R. and Fiore, S.M. (2007). Perceptual (Re)learning: A leverage point for human-centered computing. *IEEE Intelligent Systems*, 22(3), 79-83.

Reingold, E.M., Charness, N., Pomplun, M., & Stampe, D.M. (2001). Visual span in expert chess players: Evidence from eye movements. *Psychological Science*, 12, 48-55.

Schuster, D., Sellers, B., Riviera, Fiore, S.M. & Jentsch, F. (2010). Component versus holistic visual search training for improvised explosive detection. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting,*, pp.1635-1639.

Sellers, B., Rivera, J., Fiore, S., Schuster, D. & Jentsch, F. (2010). Assessing X-ray security screening detection following training with and without threat-item overlap. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting*, pp.1645-1649.

Seitz, A.R., & Dinse, H.R. (2007). A common framework for perceptual learning. *Current Opinion in Neurobiology*, 17, 1-6.

Sireteanu, R. & Rettenbach, R. (1994). Perceptual learning in visual search: fast, enduring, but non-specific. *Vision Research*, 35(14), 2037-2043.

Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005, May 26). Rare items often missed in visual searches. *Nature, 435,* 439–440.

Wolfe, J.M., Horowitz, T.S., Van Wert, M.J., Kenner, N.M., Place, S.S., and Kibbi, N. (2007). Low target prevalence is a stubborn source of errors in visual search tasks. J. Exp. Psychol. Gen. 136, 623–638.

Wolfe, J. M., & Van Wert, M. J. (2010). Varying Target Prevalence Reveals Two Dissociable Decision Criteria in Visual Search. *Current Biology 20*, 121–124, January 26, 2010.