Forensic Science International 251 (2015) 202–208

Contents lists available at ScienceDirect

Forensic Science International

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

The impact of fatigue on latent print examinations as revealed by behavioral and eye gaze testing

Thomas Busey [a,*], Henry J. Swofford [b], John Vanderkolk [c], Brandi Emerick [d]

[a] Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA
[b] U.S. Army Criminal Investigation Laboratory, Defense Forensic Science Center, Forest Park, GA, USA
[c] Indiana State Police Laboratory, Fort Wayne, IN, USA
[d] Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA

ARTICLE INFO

Article history:
Received 17 October 2014
Received in revised form 23 March 2015
Accepted 27 March 2015
Available online 7 April 2015

Keywords:
Fatigue
Errors
Fingerprints
Latent print examinations
Eye tracking

ABSTRACT

Eye tracking and behavioral methods were used to assess the effects of fatigue on performance in latent print examiners. Eye gaze was measured both before and after a fatiguing exercise involving fine-grained examination decisions. The eye tracking tasks used similar images, often laterally reversed versions of previously viewed prints, which holds image detail constant while minimizing prior recognition. These methods, as well as a within-subject design with fine grained analyses of the eye gaze data, allow fairly strong conclusions despite a relatively small subject population. Consistent with the effects of fatigue on practitioners in other fields such as radiology, behavioral performance declined with fatigue, and the eye gaze statistics suggested a smaller working memory capacity. Participants also terminated the search/examination process sooner when fatigued. However, fatigue did not produce changes in inter-examiner consistency as measured by the Earth Mover Metric. Implications for practice are discussed.

Latent print examiners often find themselves in situations where fatigue may affect their performance. The Expert Working Group on Human Factors [5] identified fatigue along with stress, complacency, overconfidence, and task overload as examples of adverse mental states of an examiner. In addition to these mental state changes, physical changes can occur such as eyestrain and back pain from looking through a loupe at latent print impressions. The Working Group report specifically identified several factors that could contribute to fatigue, including being on call for crime scene investigation, working through backlogs, and long hours as practices that could reduce performance. However, the literature contains relatively little direct evidence of the effects of fatigue on latent print examination.

To measure the effects of fatigue in tasks that resemble actual casework, we tested examiners both before and after a visual matching task that was designed to induce fatigue. We recorded both behavioral accuracy as well as eye gaze, which will allow within-subject comparisons in both measures to assess the effects of fatigue and allow for a mechanistic explanation of the changes in behavioral performance.

While relatively little work has been done with latent print examiners and fatigue using eye gaze methodology, work with radiologists has suggested that fatigue can produce reduced accommodation abilities due to eye strain after extended radiological viewing [7]. The authors suggest that this difficulty in focusing may make it more difficult for radiologists to detect abnormalities and may increase viewing time.

Van Orden et al. [13] used a nonlinear regression model to explore the relation between fatigue and eye blink duration, eye fixation duration, and pupil diameter. They replicated prior work suggesting that eyelid 'perclose' (a slight droop in the eyelid) was related to fatigue, and also identified a much stronger relation between fatigue and fixation dwell time. Their model suggests that not only do eye movements reflect elements of fatigue in a visual processing task, but that they could be used as an online diagnostic tool to measure fatigue.

More recent work has argued for the integration of fatigue measurement tools into workstation design and a daily routine [10]. This would allow a metric that relates the amount of fatigue, as estimated through external measures such as eye movements,

* Corresponding author at: Department of Psychological and Brain Sciences, Program in Cognitive Science, Indiana University, 1101 E. 10th St, Bloomington, IN 47405, USA. Tel.: +1 812 8554261.
E-mail address: busey@indiana.edu (T. Busey).

http://dx.doi.org/10.1016/j.forsciint.2015.03.028
0379-0738/

to productivity. Such a metric would quantify the costs of extended viewing sessions, to suggest when a period of diminishing returns might be reached.

Fatigue can also influence the decision making process, which has been termed *decision fatigue*. Danziger et al. [4] studied parole-based decisions and found that parole board members were much less likely to grant parole toward the end of a set of parole hearings. However, the parole granting rate rises abruptly after a food break, when presumably the parole board members are feeling refreshed.

Working memory is subject to limitations due to fatigue. Helton and Russell [6] addressed the effects of fatigue on verbal and spatial working memory tasks during a successive target detection task. Fatigue induced decrements in performance, especially when paired with a concurrent working memory task. This suggests that high cognitive resource demands brought on by a challenging task will use, and ultimately exhaust, working memory resources. This reduces sensitivity during target detection, as well as increasing reaction time.

Together these results suggest that fatigue can affect reaction times, search accuracy and attention, as well as low-level eye gaze metrics such as fixation durations and working memory. In addition, decision fatigue may affect the conclusions drawn by participants while fatigued. Given that all of these perceptual and cognitive capacities are required to conduct latent print examinations, the goal of the current study is to provide an initial comparison between fresh and fatigued performance during tasks that resemble latent print examinations. Should we find effects of fatigue with a relatively modest sample, the results may indicate the extent of the problem, as well as suggest the manner in which fatigue affects performance and guide more extensive studies using larger samples.

1. Method

1.1. Overview

Testing was done with a relatively small number of participants (five) but tested in a within-subjects design that improves experimental power. Participants were tested in two 20-min eye tracking sessions, separated by a demanding hour-long perceptual matching task that used fingerprint fragments embedded in visual noise. The two eye tracking sessions were limited to 20 min, which is the comfort limit of our eye tracking headgear. We designed the set of comparisons such that there were more comparisons than an examiner could be expected to complete in this 20 min period.

The first eye tracking session was conducted in the morning while participants reported still being relatively alert. Participants were then sent to lunch, after which they undertook an hour-long matching task that was designed to induce fatigue. Immediately after this session they participated in a section eye gaze experiment to examine the effects of fatigue on eye gaze and behavioral responding.

1.2. Participants

Participants were recruited from a large forensic science laboratory. Participants ranged in age from 26 to 31, and ranged in latent print experience between 3 and 8 years. All participants were aware that they would be tested twice, and that the two sessions would be separated by lunch and an hour-long perceptual matching task just prior to the second recording session. Participants were otherwise naïve to the goals of the study. Of the five participants, three wore glasses or contacts, and all three were in a corrected state when participating in the study. The other two had uncorrected vision.

1.3. Equipment and stimuli

There were two testing booths, each of which contained a 15″ Macbook pro laptop. This allowed for two participants to be tested at a time. Images were displayed on the built-in laptop screen using custom Java software that included gamma correction. The stimuli were chosen from an extensive database of latent prints collected from participants at Indiana University that were typically of tenprint quality. The images were degraded by visual noise that was sampled from actual latent prints and then extended using a texture synthesis algorithm [9]. The visual noise was then combined with a fingerprint image using a multiplicative model that simulates the manner in which inks combine, as illustrated in Fig. 1. To combine the friction ridge detail and the noise in a realistic manner, both the noise and the fingerprint ridge detail were modeled as neutral density filters, each of which blocks

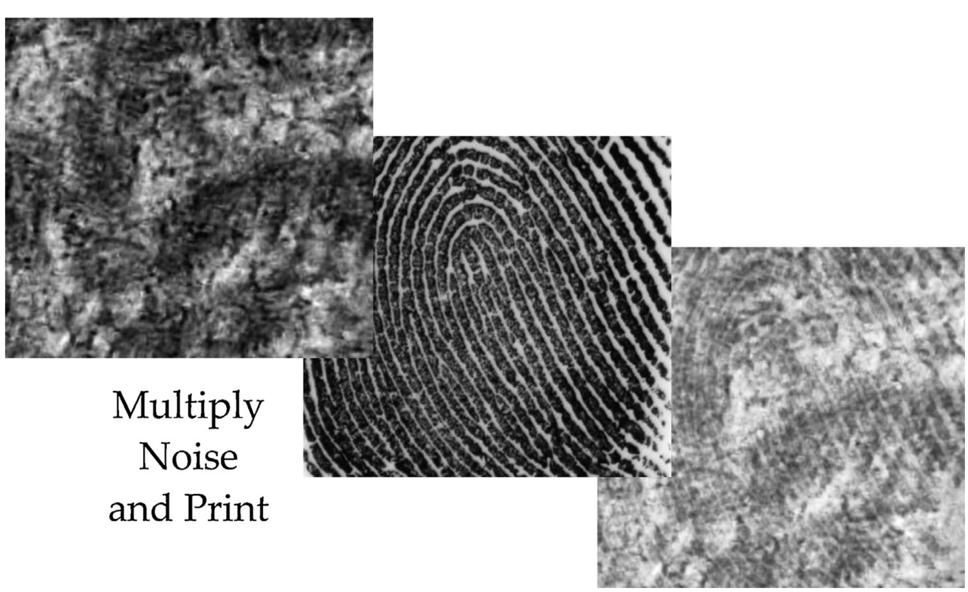

Fig. 1. Example image illustrating how multiplicative noise added to a ten-print quality image can create a realistic latent print while still allowing access to the ground truth ridge detail.

a certain proportion of light. The multiplicative model essentially computes the proportion of light that survives both filters to reach the paper and be reflected back.

Pairs of images were chosen from the database to produce fairly challenging trials for a matching task. For trials in which the two images came from the same source, care was taken to choose images that significantly varied in appearance, either due to distortion or the amount of ink on the impressions. Trials that came from different sources were chosen to share very similar appearances. This was typically achieved by left-right reversing the impression from the same finger on the opposite hand of a donor.

Thirty-five image pairs were selected for each of the morning and afternoon sessions. For the morning list, 25 images were mated pairs and 10 were non-mates. For the afternoon list, 23 images were mated pairs and 12 were non-mates. We slightly varied the number of mated and non-mated images as well as varied the order of mated and non-mated trials between the two lists to prevent subjects from using memory or counting strategies to inform their responses. Because participants worked at their own pace, each completed a different number of trials in the 20 min testing period.

To ensure that the item difficulty remained constant, most of the images were repeated in the afternoon session in some form. Fifteen of the trials were identical to the first list, 14 were laterally reversed to reduce the influence of memory from the initial list while still allowing eye gaze comparisons, and 6 included one new image of the pair that converted a matching pair to a non-matching pair or vice-versa. Participants were told that although some images looked familiar, care should be taken not infer an answer based on the first list. If the participants retained some memory for the images in the first list, this could have affected their performance. However, as we will see, performance in the afternoon actually decreased relative to the morning session.

Eye gaze was recorded using custom eyetrackers constructed from NTSC cameras and processed using the ExpertEyes software package [8]. This provides for 30 Hz recording, which over a 20 min recording period provides approximately 36,000 gaze data points per participant per session. We have developed our own algorithm to separate the continuous eye gaze record into fixations and saccades. First, we perform a running median filter over the data, which takes the median of three consecutive points. This serves to reduce the effect of noise in the pupil estimation. Next, we compute the magnitude of velocity at each time point in the data. Finally, we established a velocity threshold to segment the whole continuous stream into several big segments that correspond to dramatic eye location changes. This threshold was set to 7.3°/s, which is somewhat lower than values typically used in the literature, and we chose this in part due to our relatively slow sampling rate (30 Hz) and median filter. However, it is similar to the 20°/s adopted by Sen and Megaw [12]. To avoid spurious brief fixations, we established a minimum duration for a fixation of 67 ms.

1.4. Procedure

During the eye tracking sessions, participants freely viewed up to 35 images while their gaze was tracked using a custom eyetracker developed specifically for this application. To avoid lengthy exposure to the infrared illumination of the eyetracker we terminated the experiment after 20 min of recording by asking the participants to skip the remaining images. However, they were allowed to finish the trial they were working on upon hitting the 20 min mark.

Each participant participated in three separate phases of the experiment. In the morning of the testing day we conducted the first eye tracking session while our participants were still fresh. After lunch, we asked our participants to participate in a demanding behavioral task for 60 min. They were given fragments of images of various sizes and quality, and asked to determine whether a match existed. Fig. 2 illustrates example stimuli from this task, and shows a very faint fingerprint patch on the left side, which must be compared with the image on the right. The purpose of this task was to induce fatigue, and all participants reported that the task was extremely demanding for such a sustained duration. However, we do not discuss the results of this study because we have no basis for comparison of the results, and the overall accuracy is dependent on the amount of added noise. For our purposes it is only important that this task induce some level of fatigue, which all five subjects reported that it did. Immediately

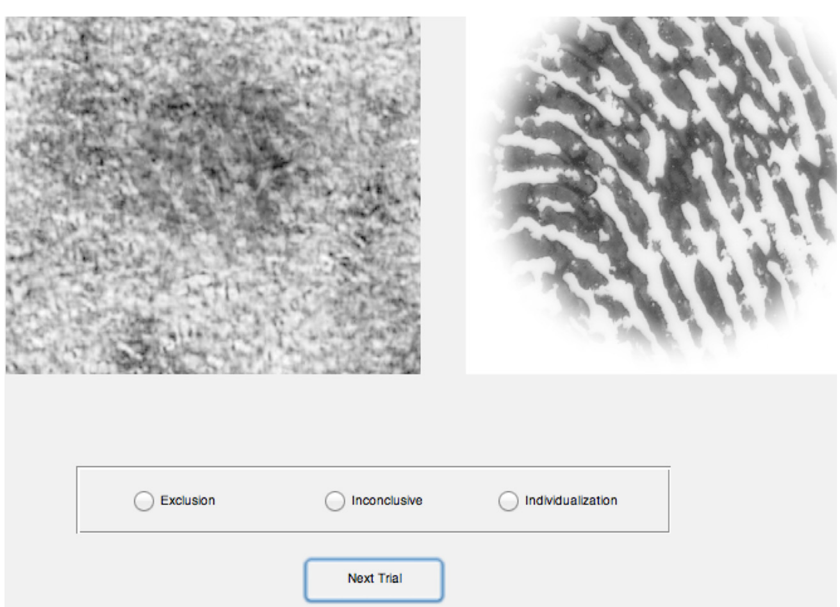

Fig. 2. Example stimuli for the intervening task that was designed to induce fatigue in our participants just prior to the afternoon eye tracking session. The figure on the left contains a fingerprint patch combined with multiplicative noise. Although difficult to see, it is representative of the difficulty of the task.

Table 1
Response frequencies for the two recording sessions. Participants completed more trials during the 20 min. afternoon session than in the 20 min. morning session. Note the higher rate of inconclusive responses for the fatigued session for both mated and non-mated trials.

	Mated pairs				Non-mated pairs				Session totals
	Yes	Inconclusive	No	Mated total	Yes	Inconclusive	No	Non-mated total	
Morning	97	6	0	103	0	0	39	39	142
Afternoon	97	16	0	113	0	5	53	58	171

Table 2
Identification and exclusion rates for morning and afternoon sessions for the 5 participants. Note that both the identification and the exclusion rates drop from morning to afternoon. Responses that are not identification or exclusions are inconclusives.

Participant number	Morning ID rate	Afternoon ID rate	Morning exclusion rate	Afternoon exclusion rate
1	0.90	0.87	1.00	0.92
2	1.00	1.00	1.00	1.00
3	1.00	1.00	1.00	0.92
4	0.90	0.70	1.00	0.92
5	0.92	0.74	1.00	0.83
Mean	0.94	0.86	1.00	0.92

following this task the participants completed a second eye tracking study in which they viewed similar images as seen in the first list.

2. Results

We first discuss the behavioral results. Table 1 presents the frequency data for the two sessions. In the morning session, participants completed 142 comparisons in the time allotted, making 97 identifications out of a total of 103 mated pairs. They made no erroneous exclusions and called 6 mated prints inconclusive. On the non-mated prints, participants completed 39 exclusions with no inconclusive decisions or erroneous identifications (trials in which non-mated prints are labeled as matches).

In the afternoon session, participants completed more trials but gave many more inconclusive ratings. This lowered their overall accuracy. For mated prints, participants made 97 identifications but labeled 16 pairs as inconclusive (compared with only 6 in the morning session). Likewise, for the non-mated pairs, participants made 53 exclusions but labeled 5 as inconclusive (compared with zero in the morning session). It appears that fatigue in this context induces participants to give up sooner on a print than is otherwise warranted.

Table 2 provides the summary data for each of the five participants, focusing on the identification and exclusion rates. For the identification rate, there was no significant drop in the identification rate due to fatigue ($t(4) = 1.83$, $p = 0.14$). However, the exclusion rate data produced a statistically significant drop ($t(4) = 3.16$, $p = 0.034$). This suggests that by giving up sooner on the prints, the participants are labeling prints as inconclusive while in the morning they could have produced an exclusion rating.

2.1. Eye tracking results

These changes in behavioral accuracy between the first and second sessions were accompanied by changes in the eye tracking data. The eye gaze record stores the location of the eye gaze at 30 frames/s. Fig. 3 through Fig. 6 illustrate the raw gaze patterns of example trials in blue, with fixations shown as red plus signs. The eye tends to pause for up to a third of a second at specific locations (called fixations), which are separated by rapid eye movements called saccades. We used standard fixation finding techniques as described previously in the methods to group the raw eye gaze into fixations and saccades. Below we describe summary statistics that

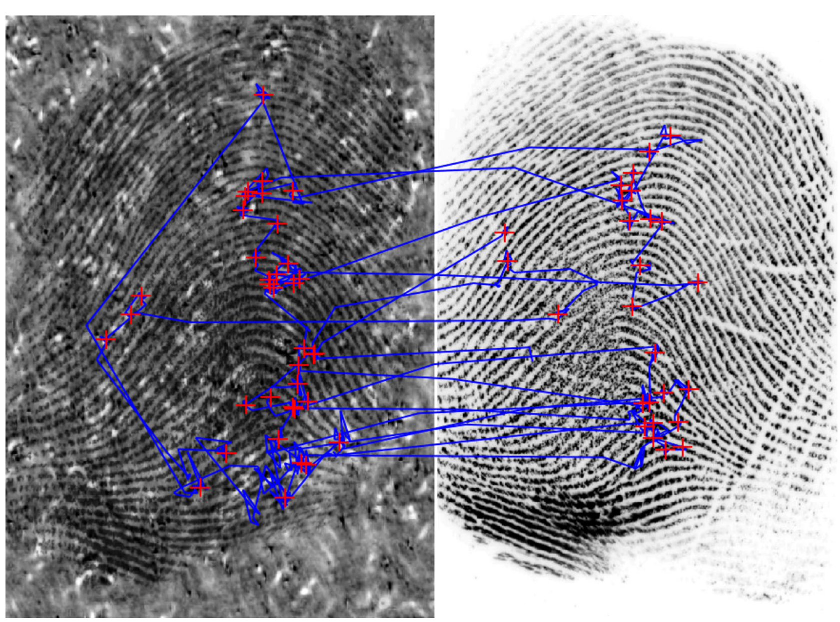

Fig. 3. Eye gaze record from a morning session. Blue lines are the gaze trace, and red crosses are the fixations. (For interpretation of the color information in this figure legend, the reader is referred to the web version of the article.)

illustrate the effects of fatigue, and then refer to the figures to demonstrate particular points.

The second session produced significant changes the statistics of the eye gaze data when examiners view the prints. Note that for all statistics reported below, we compared the second session to the first using within-subject tests and set alpha = 0.05 in two-tailed tests.

In the second session, the examiners tended to make more fixations on the latent print than in the first session: 60.2% of the fixations were on the latent in the afternoon, while 57.4% were on the latent in the morning ($t(4) = 3.09$; $p = 0.036$). In the second session, participants tended to make more saccades between the two prints: 30.0% of the saccades were across the two prints in the afternoon, while 25.1% of the saccades were across the two prints in the morning ($t(4) = 3.45$; $p = 0.026$). In addition, the saccades tended to be longer in the second session (5.9° vs. 7.0°; $t(4) = -3.807$; $p = 0.02$). All three statistics are consistent with the examiner being able to hold fewer visual features in memory when making comparisons while fatigued, and having to go back and forth more often between the two prints. However, there was no difference in terms of the number of saccades made per second (2.2 saccades/s in the morning vs. 2.0 saccades/s in the afternoon; $t(4) = 1.615$, $p = 0.182$). In addition, there was no difference between the morning vs. afternoon mean fixation durations on either the latent print (356.5 ms in the morning vs. 386.2 ms in the afternoon; $t(4) = -0.589$, $p = 0.588$) or the inked print (344.0 ms in the morning vs. 362.9 ms in the afternoon; $t(4) = -0.455$, $p = 0.673$). Thus fatigue does not appear to affect the duration of fixations.

Fig. 3 illustrates the morning session, and demonstrates that the examiner looks at several different features on the left print prior to looking at the print on the right. The search pattern appears more systematic. Contrast this with Fig. 4, which shows a gaze record from the same participant in the afternoon, examining a similar set of prints that had been left-right reversed (although the examiners did not realize this at the time according to post-experiment interviews). Now the participant makes fewer fixations before crossing over to the other print. The examiner also spent less time with this particular print.

Fig. 5 illustrates an image from a different participant from the morning session, and shows a search pattern centered on the core. However, the data in Fig. 6 is from the afternoon and is the same image that has been laterally reversed. Note how the search is more cursory, and the participant makes fewer crossings between the left and right images. Perhaps surprising is the fact that the participant correctly labeled the images in Fig. 5 as a match in the morning session, but only settled on an inconclusive response for the images in Fig. 6 despite the fact that the image quality is identical in the two pairs of images. This highlights the behavioral result that when tired, examiners tend to consider prints as inconclusive more often.

2.2. Fatigue and inter-examiner consistency

The comparisons in the previous section illustrate that examiners tend to alter their looking behavior when fatigued. They make fewer fixations before crossing the midline, more saccades across the midline, and more fixations on the latent print. These suggest that examiners may have difficulty placing multiple features into memory and are instead matching single features across the two impressions. Do these changes in looking behavior have implications for how consistent examiners are as a group? To address this question, we computed a measure of inter-examiner consistency as a test of whether examiners become more or less varied as a group when fatigued. For example, examiners might seek out different regions when fatigued, or avoid areas with marginal quality that they might otherwise venture into when less tired. This could make them more varied as a group, and reduce the consistency between the examiners in terms of the regions they fixate.

In previous work [3], we used a measure known as the Earth Mover Metric [11] to quantify the degree to which experts or novices were more consistent as a group. The Earth Mover Metric is a statistical approach to the question of how similar is two sets of fixations. When comparing two observers, each fixation from one observer can be matched with the fixation of another observer, and by doing this with all fixations from both observers we get a measure of how similar the two sets of fixations are, and therefore whether the two observers are viewing similar regions. The difficulty in matching arbitrarily numbered sets of points is that one must compute some form of correspondence between the two sets of points, which can be computationally expensive when

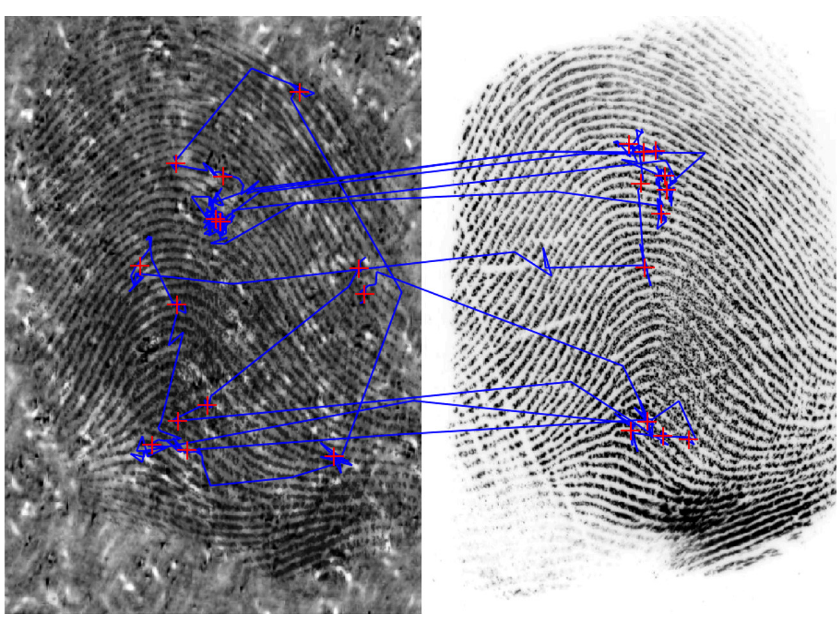

Fig. 4. Gaze record from the afternoon session and shows the same participant as in Fig. 3 with a similar set of prints (the prints were laterally reversed relative to Fig. 3). The search appears more cursory.

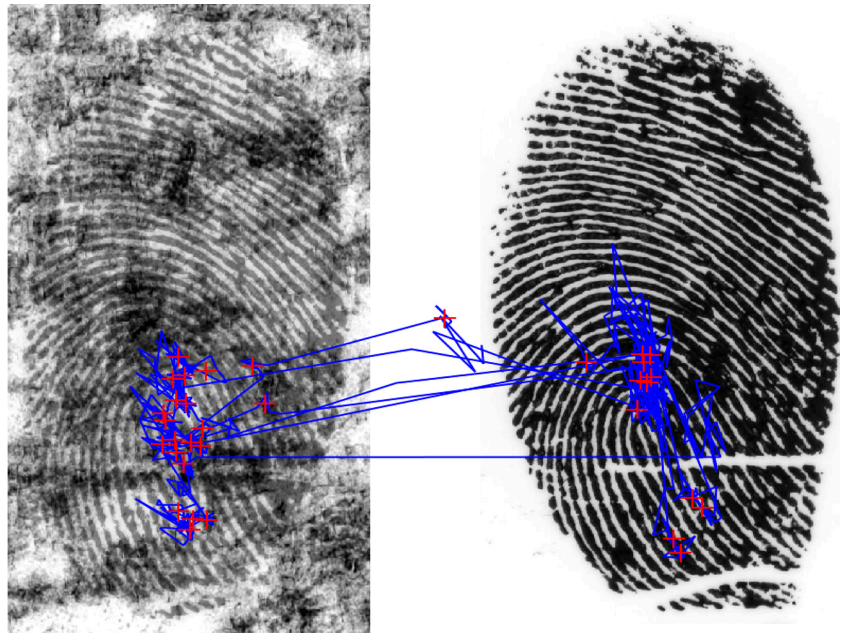

Fig. 5. Search pattern from the morning session illustrating a tight search region around and above the core area, leading to a correct 'identification' decision.

comparing all possible pairs of points across the two observers. In addition, the two sets of eye fixations may be of unequal length. The Earth Mover Metric solves both of these problems by using dynamic programming [1] to find a solution that while not optimal in a global sense, is very likely to be close to optimal.

Intuitively, this distance metric treats the fixations as small piles of dirt that must be moved from one location to another, and the Earth Mover solution attempts to move the dirt with as little effort as possible. In this case the fixations are treated as small piles of dirt, normalized by the total number of fixations, and the algorithm moves the dirt from the fixations from one subject to the fixations from another subject for the same image. The result is a single distance metric that is akin to the total work necessary to move the fixations.

For each trial in the experiment for which we had data from at least two participants, we computed the distance between all possible pairs of participants. This was computed separately for morning and afternoon sessions, because the two sessions used different images and the analysis can only be applied to subjects who viewed the same image pair. We computed the mean inter-observer distance for each trial for both morning and afternoon sessions. These distances were then compared using an unpaired test, which reflects the fact that the morning and afternoon trials were not paired (they used different image lists). Note that this comparison relies on the assumption that the morning and afternoon images were not systematically different, which we justify by noting that the initial assignment of images to lists was random, and that some of the images were subsequently left-right

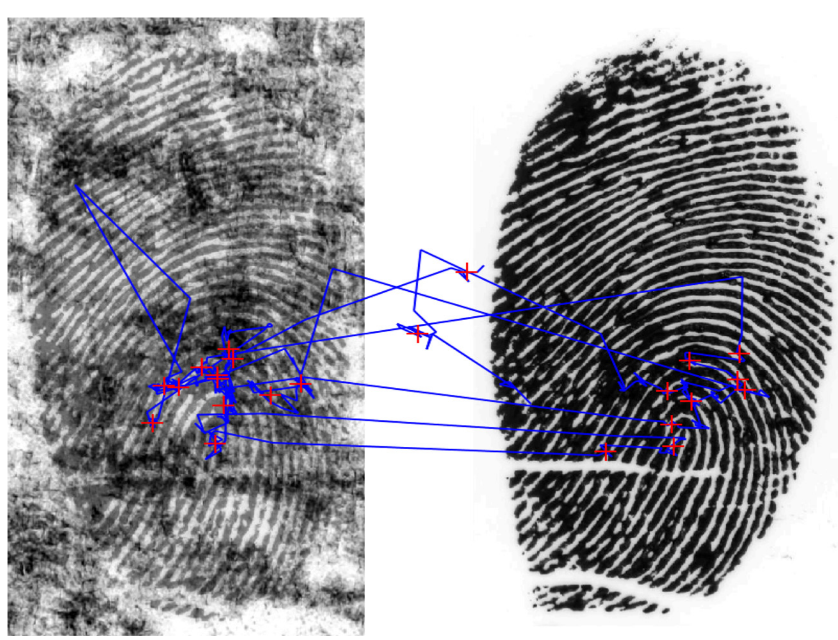

Fig. 6. Data from the same participant as in Fig. 5, looking at left-right reversed images in the afternoon session. The search is more cursory and resulted in an 'inconclusive' decision.

reversed or re-used with different comparison images and therefore had very similar image qualities.

Twenty-two trials in each list had sufficient data to compute the Earth Mover Metric between at least two participants. For each trial we computed the mean of the distances of each examiner to every other examiner for that trial, and compared the two sets of means to see if one set is systematically different from the other using an unpaired two tailed t-test. The mean of the morning session is 2.93° and the mean of the afternoon session is 2.82°, which are not statistically significantly different ($t(42) = 0.423$; $p > 0.05$).

This result demonstrates that there appear to be no strong differences in terms of the consistency of examiners as they become fatigued. Note that this comparison only addresses whether examiners become less or more consistent as a group when fatigued. It remains possible that examiners as a group shifted to a different set of features, yet still remained consistent as a group. However, our current design does not allow us to test this proposition.

3. Discussion

The two testing sessions were designed to estimate both behavioral and eye gaze performance when examiners were fresh and alert in the morning session, and then again when they were fatigued in an afternoon session. All five examiners spontaneously commented on the difficulty of the intervening task that was designed to induce fatigue, and reported experiencing fatigue during the debriefing following the conclusion of the study. Although there may be situations in which more fatigue could be experienced (e.g. 10 h shifts, 3rd shift work, extended field deployments), the current paradigm was designed to approximate the level of fatigue that might be experienced in common workplace settings.

The results of this study demonstrate that fatigue produces noticeable decrements in performance even with relatively few participants and a relatively short testing period. In addition, robust changes in eye gaze behavior are observed, suggesting that participants have difficulty placing multiple features into working memory when tired. In addition, the participants tend to terminate the search process earlier. This may be similar to the results noted by Boksem et al. [2], who showed that fatigue decreased a subject's ability to attend to relevant stimuli. This could make it difficult to place the relevant features into memory for purposes of comparison. The finding that more trials were completed yet more inconclusive conclusions were given is consistent with the concept of decision fatigue, where fatigue produces an inability to make a decision [4].

3.1. Implications for practice

The current data, although on a relatively small sample of 5 examiners, suggest that even a modest 60 min fatiguing task after lunch can produce demonstrable effects in both behavioral performance and associated eye gaze behavior. Based on these results, we suggest that managers consider organizing daily operations in such a manner to minimize fatigue issues, be cognizant of work schedules, and monitor stress factors that may induce fatigue. One response to these findings might be to discuss a self-monitoring plan with examiners that allows them to switch to less demanding tasks such as report writing when they experience fatigue. Similar recommendations are found in the medical imaging literature [10]. In addition, managers can reduce reward structures that might promote fatiguing behavior (for example, allowing 4/10 work weeks and then allowing overtime on work days might induce fatigue).

The decision fatigue illustrated by the increased number of inconclusive responses in the afternoon eye tracking session suggests that problem comparisons might benefit from a break. Examiners often report the benefits of leaving a difficult decision to the next morning to enable a fresh set of eyes, and the current results support such a strategy in terms of improved perceptual and decision making capacities.

Because the casework of various laboratories may differ, we are reluctant to offer one set of proscriptive remedies. However, we do feel that the lab environment may benefit from an open discussion of those situations that may lead to fatigue, as well as strategies to overcome it. This may lead to individual experimentation that could lead each examiner to explore the amount of time they are capable of sustained attention before a break is necessary.

References

[1] R. Bellman, Dynamic programming and a new formalism in the theory of integral equations, Proc. Natl. Acad. Sci. U.S.A. 41 (1) (1955) 31–34.
[2] M.A.S. Boksem, T.F. Meijman, M.M. Lorist, Effects of mental fatigue on attention: an ERP study, Cogn. Brain Res. 25 (1) (2005) 107–116, http://dx.doi.org/10.1016/J.Cogbrainres.2005.04.011.
[3] T. Busey, C. Yu, D. Wyatte, J.R. Vanderkolk, F.J. Parada, R. Akavipat, Consistency and variability among latent print examiners as revealed by eye tracking methodologies, J. Forensic Ident. 61 (1) (2011) 60–91.
[4] S. Danziger, J. Levav, L. Avnaim-Pesso, Extraneous factors in judicial decisions, Proc. Natl. Acad. Sci. U.S.A. 108 (17) (2011) 6889–6892, http://dx.doi.org/10.1073/Pnas.1018033108.
[5] Expert Working Group on Human Factors in Latent Print Analysis, Latent print examination and human factors: improving the practice through a systems approach, The Report of the Expert Working Group on Human Factors in Latent Print Analysis, NIST NIJ, National Institute of Justice, Washington, DC, 2012.
[6] W.S. Helton, P.N. Russell, Working memory load and the vigilance decrement, Exp. Brain Res. 212 (3) (2011) 429–437, http://dx.doi.org/10.1007/S00221-011-2749-1.
[7] E.A. Krupinski, K.S. Berbaum, Measurement of visual strain in radiologists, Acad. Radiol. 16 (8) (2009) 947–950, http://dx.doi.org/10.1016/J.Acra.2009.02.008.
[8] F.J. Parada, D. Wyatte, C. Yu, R. Akavipat, B. Emerick, T. Busey, ExpertEyes: open-source, high-definition eyetracking, Behav. Res. Methods (2014), http://dx.doi.org/10.3758/s13428-014-0465-z.
[9] J. Portilla, E.P. Simoncelli, A parametric texture model based on joint statistics of complex wavelet coefficients, Int. J. Computer Vision 40 (1) (2000) 49–71.
[10] B.I. Reiner, E. Krupinski, The insidious problem of fatigue in medical imaging practice, J. Digital Imag. 25 (1) (2012) 3–6, http://dx.doi.org/10.1007/S10278-011-9436-4.
[11] Y. Rubner, C. Tomasi, L.J. Guibas, The Earth Mover's Distance as a metric for image retrieval, Int. J. Computer Vision 40 (2) (2000) 99–121.
[12] T. Sen, T. Megaw, The effects of task variables and prolonged performance on saccadic eye movement parameters, in: A.G. Gale, F. Johnson (Eds.), Theoretical and Applied Aspects of Eye Movement Research, Elsevier, Amsterdam, 1984, pp. 103–111.
[13] K.F. Van Orden, T.P. Jung, S. Makeig, Combined eye activity measures accurately estimate changes in sustained visual task performance, Biol. Psychol. 52 (3) (2000) 221–240, http://dx.doi.org/10.1016/S0301-0511(99)00043-5.