J Forensic Sci, November 2006, Vol. 51, No. 6
doi:10.1111/j.1556-4029.2006.00269.x
Available online at: www.blackwell-synergy.com

Adrian G. Dyer,[1] Ph.D.; Bryan Found,[1,2] Ph.D.; and Doug Rogers,[1] Ph.D.

Visual Attention and Expertise for Forensic Signature Analysis

ABSTRACT: Eye tracking was used to measure visual attention of nine forensic document examiners (FDEs) and 12 control subjects on a blind signature comparison trial. Subjects evaluated 32 questioned signatures (16 genuine, eight disguised, and eight forged) which were compared, on screen, with four known signatures of the specimen provider while their eye movements, response times, and opinions were recorded. FDEs' opinions were significantly more accurate than controls, providing further evidence of FDE expertise. Both control and FDE subjects looked at signature features in a very similar way and the difference in the accuracy of their opinions can be accounted for by different cognitive processing of the visual information that they extract from the images. In a separate experiment the FDEs re-examined a reordered set of the same 32 questioned signatures. In this phase each signature was presented for only 100 msec to test if eye movements are relevant in forming opinions; performance significantly dropped, but not to chance levels indicating that the examination process comprises a combination of both global and local feature extraction strategies.

KEYWORDS: forensic science, document examination, eye movement, cognition, decision making

One of the major tasks of forensic document examiners (FDEs) is to compare handwriting features associated with a known writer to those associated with a sample of questioned handwriting to express an opinion as to whether or not they were written by the same writer. The general process by which FDEs carry out this task has been described (1–7). Although FDEs' evidence regarding handwriting has been accepted in many countries for over a century, since the late 1980s the discipline has attracted both academic and legal criticism for not having provided empirical evidence regarding its claims to expertise (8–10). In response there has been a number of detailed validation studies all of which show that the opinions of FDEs are significantly more accurate than those of lay persons (11–16). Only three of these studies have dealt specifically with the examination of questioned signatures (13–15).

Signature examinations are a subset of general handwriting examinations and exhibit a range of interesting properties. Signatures are movement artifacts that are produced in order to validate the document on which it appears as having been exposed to the signer. The signature itself may or may not be based on a copy book system. Invariably there are features within the formation that are introduced by the writer in order to personalize the form. It is not surprising, given the purpose of the signature, that they become disputed. FDEs are required to extract features from the questioned signature and compare those features with the specimen signature population to determine whether or not there is support for the proposition that either the signature is genuine (the product of normal signing behavior by the writer of the specimens) or the product of a nongenuine writing process. Examples of a nongenuine writing process are forged (copies of a genuine signature) or disguise (where the genuine writer purposefully alters the features of their signature in order to deny it at a later time). Research regarding the type and frequency of handwriting features that are used as predictors of each of genuineness, forgery, and disguise have been reported (17–21).

Although studies examining the validity of subjects' opinions on questioned signatures have been published, none to date have attempted to directly study the strategy that FDEs and lay persons use to form their opinions but rather have been studies of the opinion scores themselves. Both FDE and lay subjects have therefore remained a "black-box" where visual images were provided to the box and answers were returned. In order to objectively investigate the process by which opinions are formed it is necessary to monitor not only the opinions of the subjects participating in the task but elements of the behavior of the instrument processing the information. There are two clear components of the human examination system both of which are attractive to study in this respect; the cognitive system (that system making decisions about the significance of the features extracted) and the visual system (that system used to search for features). The experiments described here focus on the visual component of the signature examination and comparison task in order to determine what additional characteristics of FDEs expertise may be elucidated.

To view visual information in detail it is necessary to direct one's gaze so that the image falls on the high acuity region of the retina, the fovea (22–24). To direct the fovea to visual stimuli of interest, the human eye must make ballistic eye movements, called saccades that require a shift of attention (25). Thus measuring eye movements and subject gaze can be a particularly useful method for quantifying visual attention (23,24,26). One way to record subject gaze is to use a video-based eye tracking system that makes use of the different reflective properties of the eye to infrared radiation (23,27). A subject is first calibrated to a grid stimulus of known spatial dimensions, and then when test stimuli are viewed it is possible to accurately quantify the different regions of the test stimulus to which the subject pays attention (23). The use of this technique directly enables the measurement of subject attention to the different components of a stimulus (23,24,27).

[1]Handwriting Analysis and Research Laboratory, School of Human Biosciences, La Trobe University, Victoria 3086, Australia.
[2]Document Examination Team, Victoria Police Forensic Services Department, Macleod, Victoria 3085, Australia.

Received 21 Jan. 2006; and in revised form 6 June 2006; accepted 1 July 2006; published 8 Nov. 2006.

Previous studies have sought to evaluate if eye movements can reveal how experts are able to gain an advantage in the use of visual information used to solve tasks (28–31). However, currently it is not known what features of signatures FDEs pay attention to when making decisions about whether a particular signature is the genuine handwriting of a specimen provider, a forgery, or an attempt of the specimen provider to disguise their own signature. In this study we seek to clarify this situation by examining the eye movements of an expert group (FDE subjects) and a control group while performing a signature trial that simulates the type of task that FDEs perform. We thus seek to reveal how expertise enables FDEs to solve the visual/cognitive task of making decisions about signatures.

Materials and Methods

In this study subjects were required to make decisions about questioned signatures which were displayed on a TFT computer monitor. While viewing stimuli subject eye movements were recorded to quantify attention to the stimuli.

Subjects

Two experimental groups were used in the study. The control group consisted of 12 subjects (mean age 22.0 ± 3.5 SEM years) who were La Trobe University students that volunteered to participate in the study as part of a third year student research subject. The test group consisted of nine FDEs (mean age 36.4 ± 1.8 SEM; years experience $= 11.7 \pm 2.3$ SEM). Eight of the test group were qualified by their organization to present expert evidence regarding signatures, while one FDE subject was a trainee with 1 year's experience. Subjects were informed about the nature of the study and signed informed consents before participating. All subjects had better than 6/12 uncorrected visual acuity and no history of neurological disorders. The study did not consider differences between male and female subjects.

Stimuli

A single person signing their name on white A4 paper provided genuine signatures. Four of these signatures were used as specimen signatures for the comparison task, and 16 other signatures of the specimen provider were used as part of the questioned signature group. Furthermore, this person provided eight signatures that were intentionally disguised. The complete set of questioned stimuli for the signature trial comprised of 32 handwritten signatures; 16 of these were the natural signatures of a specimen provider, eight were signatures that had intentionally been disguised and eight signatures were forgeries that were provided by eight different individuals that had been given the opportunity to freehand copy the signature of the specimen provider (16). This provided a series of 32 questioned signatures that represented the typical signature types that a FDE might encounter in casework. The signatures were scanned into a computer using Adobe Photoshop (version 7.0) software and saved as 8-bit grayscale 1024×768 pixel jpeg files (10 maximum quality) to enable their display on a Tobii 1750 eye tracking system.

Data Recording

Eye movements were recorded with a Tobii 1750 binocular eye tracking system (Tobii Technology, Stockholm, Sweden). Calibration for each subject to a 16-point reference grid was carried out providing for a resolution of subject gaze to better than 0.5° of visual angle (Fig. 1). Calibrations were also confirmed before each test phase of the experiments. Stimuli were displayed using an integrated Tobii 1750 1280×1024 pixel TFT monitor. Subjects viewed the screen from a distance of 57 cm so that the visual angle of the screen was $33^\circ \times 27^\circ$ ($W \times H$), and that the width of a typical questioned signature subtended a visual angle of approximately 28°.

Eye fixations were determined using criteria of eye position remaining within a 50 pixel area for a time of greater than 100 msec (23). Data collection, fixation measurements, and analysis of areas of interest (AOI) data were determined with Tobii Clearview 2.1.2 software. To analyze subject attention to different features within the signatures a range of AOI were defined. Subsequent quantitative analysis was then conducted by exporting data from the Clearview software to custom written software. Statistical analysis was conducted with a SPSS computer package (SPSS for Windows version 11.5.0, Chicago, IL (32)).

Experiment Design Rationale

The study sought to understand how visual attention, measured with eye movements, might be used to further characterize the nature of FDEs expertise regarding the authenticity of signatures. Eight key questions regarding subjects performance and the relationship of their opinions to their visual attention strategies were addressed: (1) Are FDE subjects better than control subjects on the signature comparison task? (2) On what type(s) of signatures (genuine, disguised, or forged) do FDE subjects perform better on than control subjects? (3) Are eye movements relevant to understanding how subjects evaluate signatures? (4) Do FDE and control subjects spend different amounts of time making decisions about signatures? (5) Do FDE and control subjects refer to specimen signatures differently depending upon the independent conditions of genuine, disguised, and forged signatures? (6) Does more time viewing signatures lead to improved accuracy? (7) What regions of signatures do subjects pay the most attention to? (8) Do FDE and control subjects look at different features in the last five fixations before a decision is made?

These eight key questions were tackled in three separate experimental phases:

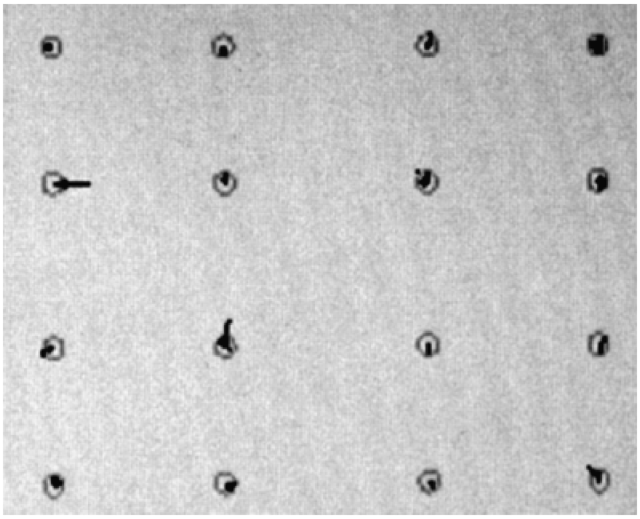

FIG. 1—*An example of a 16-point subject calibration to the Tobii eye tracking system. The circles represent 0.5° of visual angle and plotted within the circles are the eye movements of a typical subject during a calibration.*

Phase 1: Learning—The control subjects, as a group, were given a 30 min presentation on what a genuine, disguised, and forged signature was and the role of a FDE in discriminating between these different classes of stimuli. The control subjects were given an opportunity within the talk to sign their own name a number of times so as to understand the degree of signature variability and to simulate another person's signature. The control subjects were made aware of the potential consequences of a FDE making errors in a real life scenario.

For both the FDE and control groups each subject was given the opportunity to familiarize themselves with the specimen signatures. Initially subjects viewed the four specimen signatures sequentially for 10 sec each while displayed on the Tobii 1750 monitor. Each subject was then given a further 5 min to examine high resolution ink-jet prints of the four specimen signatures presented simultaneously on a piece of A4 paper. Subjects were informed that during the eye movement experiment (Phase 2) that the four specimen signatures would also be displayed at the bottom of the screen while the questioned signature would be displayed at the top of the screen (Fig. 2). This was done to give each subject an opportunity to both learn the specimen signatures, and to refer to the signatures during the experiment in a way that is consistent with the normal process of examining and comparing specimen with questioned signatures (7).

Phase 2: Recording of eye movements, response times (RTs) and opinions—This phase allowed evaluation of questions 1, 2, and 4–8. The eye movements and RTs of the control and FDE subjects were recorded while subjects carried out the comparison trial involving the 32 questioned signatures displayed on the TFT screen (Fig. 2).

Before the presentation of a questioned signature, subjects were required to fixate on a cross at the top left-hand side of the screen to control for the starting position of the eye. A subject then pressed the space bar on the keyboard to view a stimulus and using the procedure indicated in the subject instructions below sequentially viewed the 32 stimuli and provided their response.

The subject instructions were "You will view a questioned signature at the top of the screen. You may view a signature for as long as required in order to make an accurate decision about whether it is a genuine signature of the specimen provider or not. When a decision about the signature has been made, press the space bar. A screen will be displayed and you will be asked to verbally respond if the signature was either: (a) genuine, (b) forged or disguised, or (c) if you are unsure you may give an inconclusive answer. If you correctly identify the signature as either genuine or forged/disguised you will score one point, but an incorrect opinion receives a penalty of minus one point. An inconclusive response receives a score of zero points." Subjects were not required to distinguish between forgery and disguise, that is, they were not required to offer an opinion on the authorship of the forged/disguised signature.

The penalty scoring system, with an option to pass (that is provide an inconclusive opinion) was used for three reasons: (i) in real world scenarios FDE subjects do give inconclusive decisions regarding signature comparison tasks; (ii) the consequences of making errors in decisions about disputed documents may have serious consequences; and (iii) a penalty scoring system is more likely than "forced choice" to reveal genuine differences in subject performance (33).

Subjects gave a verbal response to the experimenter so as to avoid diversion of attention from the screen to select options on a keyboard.

To analyze subject attention to different features within the signatures a range of AOI were defined.

To analyze key question 5 a single AOI covering the entire questioned signature, and a second AOI covering the specimen signatures was used. These two AOI boxes were of equivalent size and were used to compute the percentage of referral fixations that subjects made to the specimen signatures for each of the questioned signature types.

To analyze key questions 7 and 8 there were 15 different AOI defined for each questioned signature (see Fig. 3*A*). While the spatial relationships of the 32 questioned signatures was a variable it was possible to position the 15 AOI so that for each of the 32 stimuli the same features were selected.

Phase 3: Tachistoscopic presentation of questioned signatures—This task was carried out by the FDE subjects following phase 2 and investigated question 3. The stimuli were the same 32 questioned signatures used in phase 2, but in this phase no specimen signatures were made available as there was insufficient

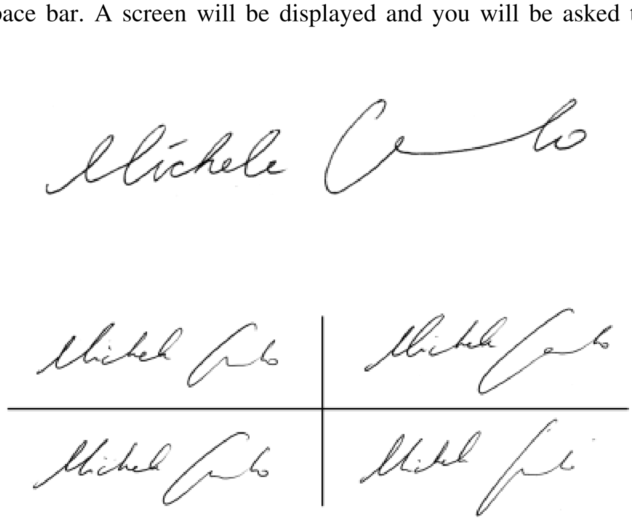

FIG. 2—*An example of the stimulus screen used in the trial where the signature at the top of the figure is a questioned signature (in this example a forgery), and the four signatures at bottom of the figure are known (specimen) signatures.*

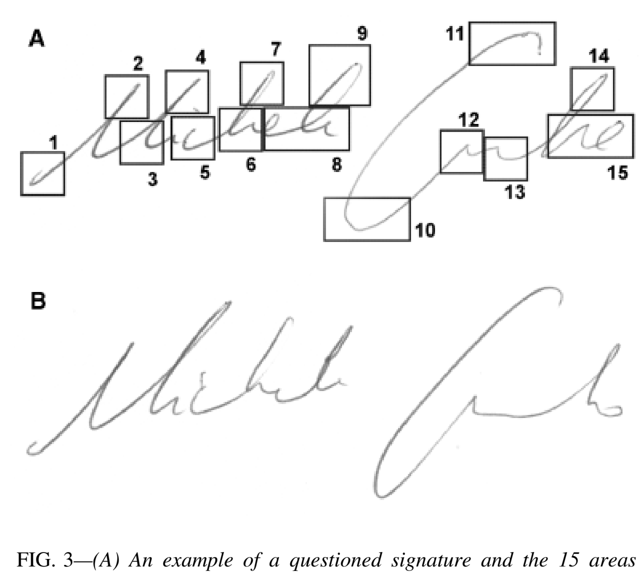

FIG. 3—*(A) An example of a questioned signature and the 15 areas of interest (AOI) boxes that were used to quantify fixations made to the different regions of the signatures. Shown here is a forgery (stimulus signature 20) with evidence of disturbed line fluency (a predictor feature of the forgery process) in AOI 11. (B) Stimulus 17 which is a genuine signature of the specimen provider that also had some evidence of disturbed line fluency.*

time for subjects to view these. The order of presentation of the questioned signatures was different from that of Phase 2. This experiment determined whether eye movements are required for accurately classifying signatures by presenting stimuli for only 100 msec, which is shorter than the time it typically takes to program and execute saccadic eye movements (tachistoscopic presentation (23)).

A fixation cross was displayed in the center of the computer screen. The fixation cross was a minimum of 6° visual angle from a questioned signature stimulus ensuring that the signature was viewed only with peripheral vision. When a subject was ready to view a stimulus they then pressed the space bar on the keyboard. A stimulus was then displayed for 100 msec and immediately followed by a random noise mask to prevent afterimages assisting the visual recognition process.

Subjects' instructions were as in phase 2, except that subjects were also told that presentation time was only 100 msec and that they should press the space bar so that they viewed the stimulus before a blink.

Results

Question 1 asked if FDE subjects were better than control subjects at the signature trial. An independent samples t-test showed that the score of FDE subjects on the trial (mean = 25.8, SEM = 1.3) was significantly better ($t(19) = 2.273$, $p = 0.035$) than control subjects (mean = 21.2, SEM = 1.4). This finding is consistent with previously reported results (14,15) and thus allows for the distinction of an expert group (FDE subjects) and a control group.

Question 2 asked on what type(s) of signatures do FDE subjects perform better than controls? To evaluate this question the number of errors where subjects incorrectly called the signature and received a penalty score was determined. Both FDE and control subjects made only a negligible number or errors on forged/disguised signatures (a total of two errors from each group, and these errors were made on disguised signatures). However, for genuine signatures the control group made 39 errors while the FDE group made only four errors. Adjusting for group size, the control group made greater than seven times as many errors compared with the FDE group. Both groups gave the same mean number of inconclusive opinions (3.1) for genuine signatures. These observations indicate that FDEs performed better (77.8% correct) than controls (45.3%) on correctly calling genuine signatures with control subjects making more errors (i.e., calling genuine signatures forged/disguised).

Question 3 asked if eye movements are relevant to understanding how subjects evaluate signatures. This question can be addressed by understanding how well do subjects perform at classifying signatures when tachistoscopically presented so that there is insufficient time to permit eye movements (Fig. 4). A paired samples t-test was used for eight FDE subjects (one FDE subject was not available to participate in this phase of the experiment) to compare scores when an open time limit was allowed (mean score = 25.9, SEM = 1.4; mean time taken = 20.9 sec, SEM = 1.5) and when stimuli were presented for only 100 msec (mean score = 16.1, SEM = 1.4). Performance was significantly better with an open time limit ($t(7) = 4.795$, $p = 0.002$). This shows that eye movements are relevant to evaluating signatures, as performance is significantly lower for an experimental condition that prevents eye movements. Interestingly however, the scores for the brief presentation were significantly different from chance. To determine this we empirically calculated that there was

FIG. 4—*An example of a subject's fixation (represented by the gray circle) during a tachistoscopic presentation of a questioned signature. The subject fixated on a cross before the tachistoscopic presentation. The signature then appeared for only 100 msec so that it was not possible for the subject to fixate on any features within the signature (eye remains on the fixation cross).*

less than a 5% chance of scoring eight on the test if answers were randomly assigned; and we then evaluated the frequency with which the subjects scored better than eight on the test (sign test, $p < 0.01$). This suggests that there are two components to signature recognition: (i) global processing of the spatial properties of the entire signature that do not require eye movements; and (ii) local processing of distinct features fixated on during an evaluation which enables improved performance at correctly classifying signatures.

Question 4 evaluated whether FDE and control subjects spend different amounts of time making a decision about the three different types of signatures. Figure 5 shows the mean RT of subjects while making decisions about the three types of stimuli in the signature trial. These data were analyzed considering two groups (control and FDE subjects) and three types of signatures (genuine, disguised, and forged). The dependant variable RT was analyzed with a mixed between–within subjects ANOVA. There was a significant effect on RT depending upon whether subjects were making a decision about a genuine, disguised or forged signature ($F(2, 18) = 14.463$, $p < 0.0005$) showing that subjects spent a significantly longer time on genuine signatures. There was no interaction between groups ($F(2, 18) = 1.214$, $p = 0.320$) showing that both groups spent similar amounts of time making a decision about different types of signatures. There was no significant difference between groups in RT for the different types of signatures ($F(1, 19) = 0.686$, $p = 0.418$). While a plot of RT for disguised signatures does indicate that FDE subjects might spend more time on disguised signatures (Fig. 5), the statistical analysis indicates that this was not significant. This nonsignificant finding is also

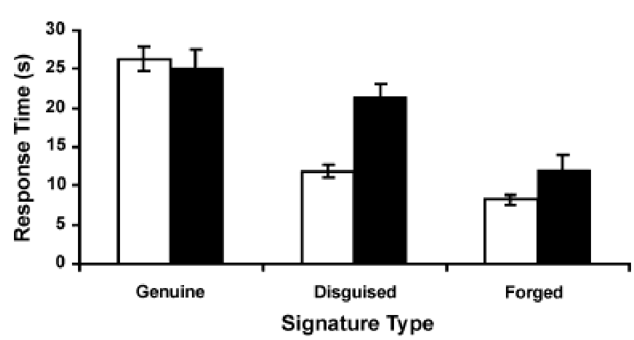

FIG. 5—*The mean response time (+/− SEM) of control subjects (white bars) and FDE subjects (black bars) when making decisions about genuine, disguised and forged signatures.*

collaborated by considering that minimal errors were made on disguised signatures by both the control and FDE groups (see question 2 above). Thus it appears reasonable to conclude that when subject instructions placed an emphasis on accuracy that both control and FDE subjects elect to spend a similar amount of time making decisions about signatures. To test for motivation of the control subjects in relation to the FDE subjects we also compared RT for either responding with an answer or making an inconclusive decision. For this analysis, one subject within each group made no inconclusive decisions and was thus excluded. Both control subjects and FDEs spent significantly more time viewing signatures for which an inconclusive decision was made, indicating that the level of motivation to perform well was high in both groups. For control subjects, mean RT for inconclusive opinions was 37.3 sec compared with 16.4 sec for when a decision was made (paired samples t-test: $t(10) = 3.931$, $p = 0.002$). For FDEs, mean RT for inconclusive opinions was 40.5 sec compared with 19.4 sec for when a decision was made (paired samples t-test: $t(7) = 2.964$, $p = 0.023$).

Question 5 determined whether there were differences in how the control and FDE groups referred to the specimen signatures during evaluations of questioned signatures. For each of the 32 stimuli there were two AOI and the dependant variable was taken to be the percentage of fixations made to the questioned signature. Data were arcsine transformed (34) and analyzed with a mixed between–within subjects ANOVA considering the two experimental groups and three types of signatures. There was a significant effect depending upon whether subjects were making a decision about a genuine, disguised or forged signature ($F(2, 18) = 6.351$, $p = 0.008$) indicating that for genuine signatures, subjects made more referrals to the specimen signature box. There was no interaction between groups ($F(2, 18) = 1.750$, $p = 0.202$), showing that both groups paid similar amounts of attention to questioned and specimen signatures. There was no significant difference between groups for the different types of signatures ($F(1, 19) = 0.251$, $p = 0.622$). Thus the only significant finding of question 5 is that both groups refer to the specimen signatures more often when evaluating a questioned signature that is the genuine signature of the specimen provider.

Question 6 evaluated if there was any evidence that with increasing viewing time that subject performance in the signature trial improved. In human cognition the ability to make accurate decisions is often correlated with the amount of time allocated to the task (35), and in this regard we had stressed to subjects in the instructions that the emphasis was on accuracy (35). There was no evidence that increased RT improved performance for either control subjects (Spearman's $\rho = -0.572$, $p = 0.052$) or the FDE subjects (Spearman's $\rho = -0.160$, $p = 0.682$). Indeed, if anything, there appears to be a slight negative relationship between performance and RT, although this does not reach statistical significance for either group. However, this finding must be interpreted in relation to the finding in question 3 where there must be sufficient time to view features on which to make a decision. Thus it appears that after an initial period required in order to view signature features, that additional viewing time does not improve performance at the task. Figure 5 shows the mean RTs for subjects viewing the three types of signatures used in the test, and given the nonsignificant result for the speed accuracy analysis, these times can be taken as a reasonable representation of the time required to evaluate signatures in this type of trial.

Question 7 investigated the regions of the signature to which subjects made the most fixations. Figure 6A shows a typical example of the eye movements of a FDE subject while viewing a questioned signature stimulus, and Fig. 6B shows the amount of attention that control or FDE subjects paid to different regions of the 32 questioned signatures over the entire trial. There is a very similar amount of attention paid to the different features by both the control and FDE groups, which, in collaboration with the RT data from question 4 (Fig. 5), suggests that both groups are obtaining similar types of information about the stimuli. We further investigated why there was not an equal distribution of attention to the 15 AOI (Fig. 6B) by considering the amount of attention to either medial (AOI 3, 5, 6, 8, 12, 13, 15) or peripheral (AOI 1, 2, 4, 7, 9, 10, 11, 14) features by the FDE subjects. Significantly more attention was paid to the medial features of signatures (77.0%) compared with the peripheral features (23.0%; paired samples t-test, $t(8) = 4.882$, $p = 0.001$).

As a control the duration of time that subjects spent fixating in the different AOI was also evaluated compared with the number of fixations; there was a very strong correlation between fixation duration and number of fixations ($R^2 = 0.9523$). This indicates that counting the number of fixations is an accurate

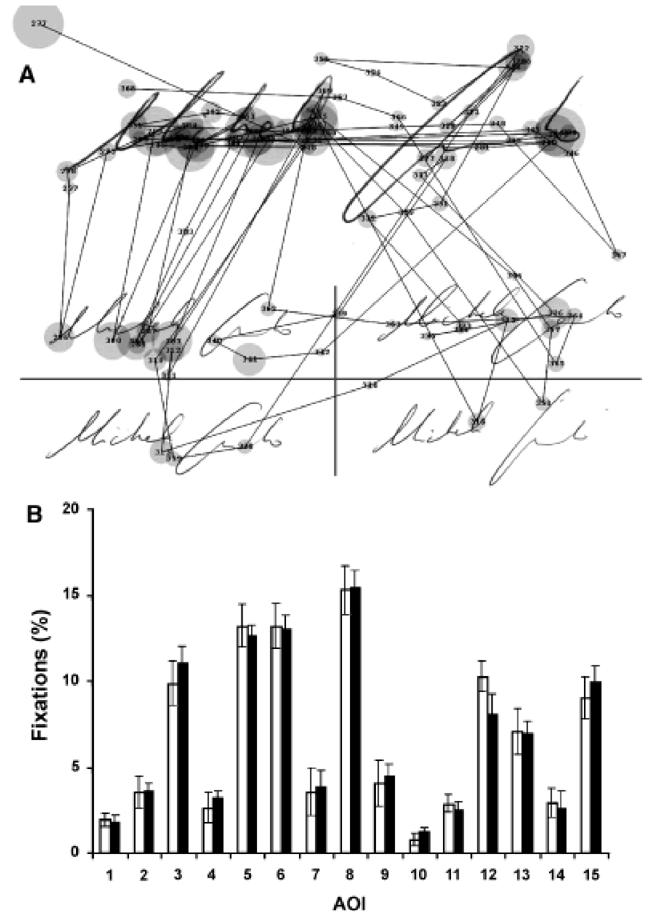

FIG. 6—*(A) Typical eye movement recording of a forensic document examiners (FDE) while making a decision about a questioned signature. In this example the signature was a genuine signature of the specimen provider, and the subject correctly called the signature. The response time of the subject to this stimulus was 25.4 sec which is very similar to the mean response time of FDEs for genuine questioned signatures. Circles show fixations and lines indicate saccadic eye movements. Numbers on the fixations indicate sequence order which begin at fixation 277 at the top left-hand side of the image, where a fixation cross had been displayed before the presentation of the questioned signature. (B) The mean percentage fixations (± SEM) to 15 different areas of interest (AOI; see Fig. 3A) by 12 control subjects (white bars) and nine FDEs (black bars) for all questioned signatures.*

representation of the attention subjects pay to the different regions of the signatures.

Question 8 investigated visual attention to the 15 different AOI during the last five fixations before a decision was made and thus tested the hypothesis that subjects may make a decision triggered by a "salient" feature. The alternative hypothesis was that subjects made a decision using the accumulation of information from multiple points suggesting that attention in the final five fixations should not differ from the average attention paid to the 15 different AOI during the course of the experiment. To test these hypotheses it was possible to consider two particular stimuli that each displayed a line fluency disturbance (a salient feature which may be predictive of a forgery or disguise process) and that might act as a trigger for a decision. These salient features were AOI 11 for a forged signature (Fig. 3*A*) and AOI 9 for a genuine signature (Fig. 3*B*).

All subjects called the forged signature correctly. For the genuine signature, six of the control subjects' opinions were wrong, with one correct and five inconclusive. This signature was called correctly by two FDEs (there was one error and six inconclusive opinions).

The number of fixations that subjects made to the 15 different AOI in their last five fixations before a decision being made for all genuine signatures was used to establish a baseline of image salience describing how much visual attention was paid to these regions. Then the frequencies of fixations during the last five

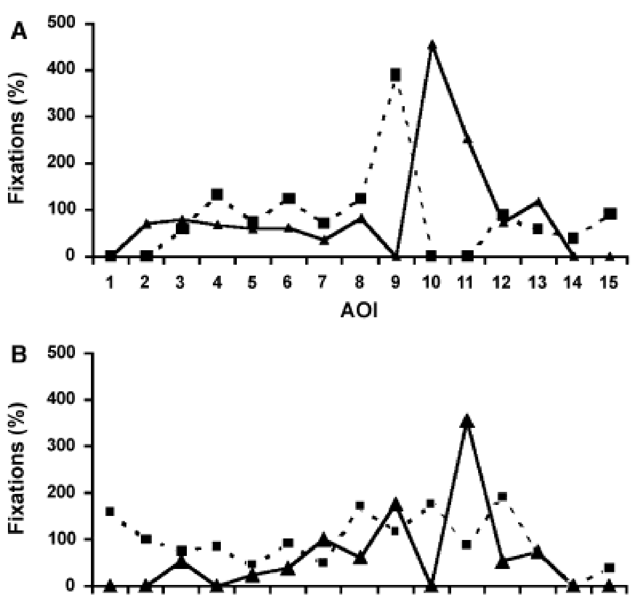

FIG. 7—*Subjects' attention when making a decision about two stimuli that tested the hypothesis that a decision was made following the viewing of a salient feature. (A) Control subjects' fixations (%) as a measure of attention to particular areas of interest (AOI) during the last five fixations before a decision being made. For both a genuine signature (broken line) and a forged signature (solid line) control subjects showed increased attention to a salient feature that could be taken as evidence of the forgery process suggesting that they made a decision based upon the salient feature. (B) Forensic document examiners (FDE) fixations (%). For the forged signature (solid line), FDEs showed increased attention to the salient feature (AOI 11). However, for a genuine signature that contained a predictor of the forgery process the FDEs did not show increased attention to the feature before making a decision (broken line). This is evidence that FDE subjects weight their decision based upon evidence from multiple sources of visual information.*

fixations to the genuine signature or forged signature were computed as a percentage ratio compared with the baseline of image salience (Fig. 7). Figure 7*A* shows the increased attention to the features indicative of forgery or disguise by the control subjects. Control subjects also paid more attention to AOI 10 on the forged signature (see Fig. 3*A*) which also contains similar evidence of line fluency disturbance. Thus decisions by control subjects are mediated by some form of serial search that concludes after viewing a feature that enables classification of the signature. For both the genuine and the forged signature control subjects made a decision based upon a salient feature that could be taken as evidence of the forgery process. For the genuine signature, this approach can account for the control subjects' high error rate for this signature.

Figure 7*B* shows that for the forged signature FDEs do pay increased attention before a decision being made, however, for the genuine signature the FDEs did not make a decision based upon a single feature that may have indicated simulation or disguise behavior. This is evidence that FDE subjects weight their decision based upon evidence from multiple sources of visual information.

Discussion

Consistent with the findings of previous studies (14,15) this investigation shows that FDE subjects perform significantly better on signature tasks than do lay people. This effect was evident in spite of the control subjects being primed to the task (so technically could not be considered "lay persons") and, based on the RT data, were highly motivated to perform well. An evaluation of error scores indicated that the main reason for FDE subjects performing better was a considerably lower number of errors when expressing opinions regarding the questioned signatures that were the genuine signatures of the specimen provider.

By evaluating the ability of FDE subjects to correctly call questioned signatures both when permitted to view all stimulus features and when eye movements were restricted by a tachistoscopic presentation it is possible to conclude that to maximize the accuracy it is necessary for FDEs to evaluate a range of different features. However, even when eye movements are restricted it was still possible for the FDEs to call signatures with some degree of accuracy. This indicates that there is a dual evaluative process that incorporates both global processing of stimuli, as well as local detailed feature extraction which is dependent on eye movements. Most attention is paid to "medial" features (see question 7) where there are more concentrated changes in pen movement that provide more features on which writing process behavior might be predicted. It is possible that this strategy also provides an efficient mechanism to globally process the overall spatial elements of a signature. In addition, it is possible that the information extracted from the "peripheral" features primarily relates to line quality which may be assessed more easily and require less attention (that is, less direct eye fixations).

This current study shows that there are many similarities between how control and FDE subjects view questioned signatures. Both control and FDE subjects spent a similar amount of time making decisions when calling the different types of signatures (Fig. 5), and both groups found it necessary to refer to the specimens more frequently when evaluating questioned signatures that were genuine. Neither group showed evidence of a positive relationship between RT and performance in the signature trial. Furthermore, both groups had a similar percentage of mean fixations to the different AOI within the signatures (Fig. 6*B*), indicating that attention to particular features does not explain the higher level of performance in FDE subjects.

Figure 6*B* provides evidence that both control and FDE groups pay a similar amount of attention to the different features within questioned signatures. This is interesting in the context that there are two distinct types of saccadic eye movements (26,36): (i) reflexive saccades where eye movements are stimulus driven by image salience using bottomup cognitive processes; and (ii) volitional saccades where a topdown or knowledge based cognitive mechanism controls where the eye should look (26,36). Control subjects had not received formal training in relation to volitionally directing their gaze in a systematic manner (a topdown mechanism) to particular features of a signature and it is likely that their approach is more stimulus driven (i.e., bottomup). Therefore, as the visual attention of the FDE subjects to the stimuli was very similar to the control group (Fig. 6*B*), we suggest that the eye movements of FDE subjects are also mainly influenced by bottomup processes of controlling visual attention rather than following a systematic, prescribed approach resulting from training. Indeed, this may be a more effective way to evaluate signatures as it is known that visual attention in humans is more efficiently employed using an anarchic, reflexive-type search rather than volitional search strategies (37).

However, Fig. 7 suggests that the control and FDE groups used different cognitive strategies when making decisions about genuine signatures of the specimen provider. In the case of the control subjects a decision was made after viewing a salient feature that potentially enabled the classification of the questioned signature (Fig. 7*A*). These data for control subjects are consistent with a sequential search of features indicating serial cognitive processing of visual information (38). All the FDE subjects viewed the salient feature in AOI 9 for stimulus signature 17 (a genuine signature), but this did not appear to act as a trigger for a decision to be made (Fig. 7*B*). This suggests that FDE subjects are processing information about a salient feature in conjunction with other information that is available within the signature. These data are more consistent with simultaneous or parallel processing of visual information (38). The data indicates that FDE subjects perform better than the control group because when calling genuine signatures the FDE subjects take into account multiple pieces of visual information and are thus less likely to incorrectly classify these signatures. This finding explains why the FDE group scored significantly higher than the control group in the signature trial. However, the FDE group did make decisions about the forged stimulus (stimulus signature 20) on the basis of a salient feature (Fig. 7B). The explanation for this potentially confounding information might be that there was insufficient evidence within the forged signature to suggest that this questioned signature was consistent with the handwriting of the specimen provider, and so rather than using simultaneous or parallel processing to make the decision, it was possible to "correctly" conclude that this questioned signature was a forgery using a serial process. In this regard, no subject from either group made an error on the forged signature and the RT to make a decision was only 31% of the time subjects spent on genuine signatures. This indicates that this signature was an easy stimulus and only required serial processing of the visual information. The test of expertise was that the FDE group were able to realize that for the genuine signature (stimulus 17) there was conflicting information, and the RT for the FDE subjects on this one particular signature was nearly twice as long as for the control subjects.

From the experimental data it is possible to propose a general model to rationalize why it is that FDEs outperform control subjects in spite of similar comparison strategies, similar RTs for different categories of questioned signatures (genuine, disguised, and forged), and similar attention being paid to the different features. It appears that the increased time taken for FDEs and the control group to identify genuine signatures as opposed to forged or disguised may well be explained by a constant strategy of searching the image primarily for the predictors of the forgery or disguise behaviors (that is searching for dissimilarities between the questioned signature and the population of specimen [known] signatures), rather than primarily searching for features that are predictors of the natural writing process. The line fluency and gross spatial information can be searched very quickly (supported by the tachistoscopic presentation in phase 3) and the predictors of forgery or disguised behaviors can be obvious. For the genuine signatures, the task takes longer to reach the decision phase (in comparison to forged and disguised signatures) as the search for the predictors continues to remain fruitless. The strategy moves from line quality and gross spatial features (in the periphery of the signature) to a more careful assessment of more detailed features (within the medial area of the signature) before the decision that, in the absence of the predictor features, the signature is likely to be genuine. The difference between the FDE and control groups which results in the controls expressing more erroneous opinions lies in the cognitive domain. The evidence presented suggests that control subjects utilize a serial cognitive processing model involving a sequential search for salient features in order to classify signatures. FDEs appear to utilize (at least for difficult stimuli) simultaneous or parallel processing where their decisions are not necessarily influenced by single salient features, but rather by the assessment of the "significance" of each salient feature both in relation to its own "weight" in identifying the writing process, and its relationship to the absence or presence of other salient features within the writing trace.

Acknowledgments

We are grateful for the generous contribution of time by the participants of the study. We thank Dr. Suzane Vassallo for discussions about the design of the experiments and Dr. Lalina Muir for comments on the manuscript.

References

1. Osborn AS. Questioned documents. 2nd ed. Chicago: Nelson-Hall Co, 1929.
2. Harrison WR. Suspect documents: their scientific examination. New York: Praeger, 1958.
3. Conway JVP. Evidential documents. Illinois: Charles C Thomas, 1959.
4. Hilton O. Scientific examination of questioned documents. New York: Elsevier Science Publishing Co. Inc, 1982.
5. Ellen D. The scientific examination of documents: methods and techniques. West Sussex: Ellis Horwood Limited, 1989.
6. Huber RA, Headrick AM. Handwriting identification: facts and fundamentals. Boca Raton: CRC Press, 1999.
7. Found B, Rogers D, editors. Documentation of forensic handwriting comparison and identification method: a modular approach. J Forensic Doc Exam 1999;12:1–68.
8. Risinger DM, Denbeaux MP, Saks MJ. Exorcism of ignorance as a proxy for rational knowledge: the lessons of handwriting identification "expertise." Uni Pennsylv Law Rev 1989;137:731–92.
9. United States v. Starzecpyzel, 880 F.Supp.1027 (S.D.N.Y. 1995).
10. Risinger DM, Saks MJ. Science and nonscience in the courts: daubert meets handwriting identification expertise. Iowa Law Rev 1996;82:21–74.
11. Kam M, Wetstein J, Conn R. Proficiency of professional document examiners in writer identification. J Forensic Sci 1994;39:5–14.
12. Kam M, Fielding G, Conn R. Writer identification by professional document examiners. J Forensic Sci 1997;42:778–86.

13. Found B, Sita J, Rogers D. The development of a program for characterising forensic handwriting examiners' expertise: signature examination pilot study. J Forensic Doc Exam 1999;12:69–80.
14. Kam M, Gummadidala K, Fielding G, Conn R. Signature authentification by forensic document examiners. J Forensic Sci 2001;46:884–8.
15. Sita J, Found B, Rogers D. Forensic handwriting examiners' expertise for signature comparison. J Forensic Sci 2002;47:1117–24.
16. Found B, Rogers D. The initial profiling trial of a program to characterize forensic handwriting examiners' skill. J Am Soc Quest Doc Exam 2003;6:72–81.
17. Leung SC, Cheng YS, Fung HT, Poon NL, Forgery I. Simulation. J Forensic Sci 1993;38:402–12.
18. van Gemert AWA, van Galen GP. Dynamic features of mimicking another persons writing and signature. In: Simner ML, Leedham CG, Thomassen AJWM, editors. Handwriting and drawing research: basic and applied issues. Amsterdam: IOS Press, 1996:459–71.
19. Herkt A. Signature disguise or signature forgery. J Forensic Sci Soc 1986;26:257–66.
20. Wendt GW. Statistical observations of disguised signatures. J Am Soc Quest Doc Exam 2000;3:19–27.
21. Black D, Found B, Rogers D. The frequency of the occurrence of hand writing performance features used to predict whether questioned signatures are simulated. J Forensic Doc Exam 2003;15:17–28.
22. Yarbus AL. Eye movements and vision. New York: Plenum, 1967.
23. Duchowski AT. Eye tracking methodology: theory and practice. London: Springer-Verlag, 2003.
24. Martinez-Conde S, Macknik SL, Hubel DH. The role of fixational eye movements in visual perception. Nat Neurosci 2004;5:229–40.
25. Kustov AA, Robinson DL. Shared neural control of attentional shifts and eye movements. Nature 1996;384:74–7.
26. Parkhurst D, Law K, Niebur E. Modeling the role of salience in the allocation of overt visual attention. Vision Res 2002;42:107–23.
27. Henderson J, Falk R, Minut S, Dyer F, Mahadevan A. Gaze control for face learning by humans and machines. In: Shipley T, Kellman P, editors. From fragments to objects: segmentation processes in vision. New York: Elsevier, 2001:463–81.
28. Reingold EM, Charness N, Pomplun M, Stampe DM. Visual span in expert chess players. Psychol Sci 2001;12:48–55.
29. Knoblich G, Ohlsson S, Raney GE. An eye movement study of insight problem solving. Mem Cog 2001;29:1000–9.
30. Nodine CF, Mello-Thoms C, Kundel HL, Weinstein SP. Time course of perception and decision making during mammographic interpretation. Am J Roentgenol 2002;179:917–23.
31. Russo FD, Pitzalis S, Spinelli D. Fixation stability and saccadic latency in elite shooters. Vision Res 2003;43:1837–45.
32. Pallant J. SPSS survival manual. Crows Nest NSW, Australia: Allen and Unwin, 2001.
33. Petz B. To penalize or not to penalize false answers in the achievement tests of the alternative type. Revija za Psihologiju 1978;8:49–56.
34. Sokal RR, Rohlf FJ. Biometry. New York: WH Freeman, 1981.
35. Rival C, Oliver I, Ceyte H. Effects of temporal and/or spatial instructions on the speed accuracy tradeoff of pointing movements in children. Neurosci Lett 2003;336:65–9.
36. Leigh RJ, Zee DS. The neurology of eye movements. New York: Oxford University Press Inc, 1999:646.
37. Wolfe JM, Alvarez GA, Horowitz TS. Attention is fast but volition is slow. Nature 2000;406:691.
38. Townsend JT, Wenger M. The serial-parallel dilemma: a case study in linkage of theory and method. Psychonomic Bullet Rev 2004;11:391–418.

Additional information and reprint requests:
Doug Rogers, Ph.D.
Forensic Expertise Profiling Laboratory
School of Human Biosciences
La Trobe University
Victoria 3086
Australia
E-mail: d.rogers@latrobe.edu.au