J Forensic Sci, September 2008, Vol. 53, No. 5
doi: 10.1111/j.1556-4029.2008.00794.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.*

An Insight into Forensic Document Examiner Expertise for Discriminating Between Forged and Disguised Signatures

ABSTRACT: It has previously been shown that forensic document examiners (FDEs) have expertise in providing opinions about whether questioned signatures are genuine or simulated. This study extends the exploration of FDE expertise by evaluating the performance of eight FDEs and 12 control subjects at identifying signatures as either forgeries or the disguised writing of a specimen provider. Subject eye movements and response times were recorded with a Tobii 1750 eye tracker during the signature evaluations. Using a penalty scoring system, FDEs performed significantly better than control subjects ($t = 2.465$, $p = 0.024$), with one FDE able to correctly call 13 of the 16 test stimuli (and three inconclusive calls). An analysis of eye movement search patterns by the subjects indicated that a very similar search strategy was employed by both groups, suggesting that visual inspection of signatures is mediated by a bottom up search strategy. However, FDEs spent greater than 50% longer to make a decision than the control group. The findings are suggestive that for some stimuli FDEs can discriminate between forgeries and disguises, and that this ability is due to a careful inspection and consideration of multiple features within a signature.

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

The expertise of forensic document examiners (FDEs) in handwriting identification compared to control subjects has been reported (1–6). Three of these studies dealt specifically with signature comparison tasks (3–5). In all cases it was found that FDEs were significantly better than the control subjects at correctly discriminating between genuine and simulated (copies of the genuine signature not written by the genuine writer) questioned signatures. Currently however, no expertise study that we are aware of has explored in detail as to whether FDEs are able to discriminate between forgeries (copies of the genuine signature not written by the genuine writer) and disguised signatures (performed by the genuine writer), and whether FDEs show an expertise effect over control subjects performing the same tasks. It is recognized that document examiners are commonly averse to describing a signature as a forgery and that many examiners will not provide an opinion of authorship on a signature they consider to be a simulation. However, for the purpose of this paper we have used the terms "forgery" or "forged signature" and "disguised" as the author and process of production of the signatures were known to the experimenters and the purpose of the study was to explore examiners' ability to distinguish between these two types of behaviors.

A previous study by the authors used eye movement recordings to understand how experts are able to solve forensic signature tasks better than a control group (6). The rationale for pursuing experimental approaches based on monitoring subjects' eye movements is theoretically straightforward. For a subject to view an image in detail the gaze must be directed in a way that the image falls on the high acuity region of the eye (7–9). Shifts of attention between features are carried out using ballistic eye movements called "saccades" (6–11). The measurement of this process therefore provides a mechanism to quantify visual attention. Video-based eye tracking systems have been developed to quantify subjects' eye movements by using the differential reflective properties of the eye (8). Subsequent to calibration, the eye tracking system allows the experimenter to accurately monitor subjects' attention to different regions of the test stimulus.

Dyer et al. (6) produced evidence using a video-based eye tracking system, that eye movements are relevant to understanding how FDEs view and process information relating to questioned signatures. For example, it was found that if signatures were presented to FDEs for only 100 ms (a time period that prevented their eyes from moving at all), then the ability of FDEs to identify genuine or simulated signatures was significantly better than chance. This in combination with the finding that an unrestricted time period resulted in FDEs performing better again, provides evidence that a dual evaluative process is in action. The evaluation incorporates both global processing of stimuli as well as local detailed feature extraction which is eye-movement dependent. The study of eye movements in FDE subjects (6) provides collaborating evidence that forensic experts develop visual expertise whilst dealing with particular test stimuli. Specifically, Busey and Vanderkolk (12) showed that fingerprint examiners demonstrate an expertise effect over control subjects, and provided evidence of forensic experts using configural processing to match fragments of fingerprints when presented with image noise.

The research reported here develops on the findings of Dyer et al. (6) and is focused on the issue of the validity of opinions relating to the discrimination between forged and disguised signatures by FDEs and a control subject group. The following four key questions (KQ) were considered:

KQ1: Can subjects correctly discriminate between disguised and forged signatures, and do FDEs show an expertise effect compared to the control subjects?

[1]Handwriting Analysis and Research Laboratory, School of Human Biosciences, La Trobe University, Vic. 3086, Australia.
[2]Document Examination Unit, Documents and Digital Evidence Branch, Victoria Police Forensic Services Department, Macleod, Vic. 3085, Australia.
*Current address: Centre for Brain and Behaviour, Department of Physiology, Monash University, Clayton, Vic. 3800, Australia.

Received 14 Jan. 2007; and in revised form 14 June 2007; accepted 7 Dec. 2007.

KQ2: Do FDEs perform better than control subjects on forged and/or disguised signatures?

KQ3: Do FDEs and control subjects choose to allocate different amounts of time for making decisions about disguised and forged signatures?

KQ4: Do FDEs or control subjects look at signature simulations differently when making a decision?

Materials and Methods

This study was conducted immediately after the initial experiments reported in Dyer et al. (6). In this current study, subjects were required to make decisions about questioned forged or disguised signatures that were displayed on a TFT computer monitor. While viewing stimuli, subject eye movements and response times (RT) 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 eight FDEs (mean age 37.4 ± 1.8 SEM; years experience = 13.0 ± 2.2 SEM) who volunteered to participate in the study. All examiners were authorized in handwriting and signature examination by NATA accredited Australian government laboratories. The accrediting body, NATA, is recognized by the ASCLAB-LAB group in the U.S.A. Subjects were informed about the nature of the study and signed informed consents prior to participating. Subjects had better than 6/12 uncorrected visual acuity and no history of neurological disorders. The study did not consider differences between M/F subjects.

Stimuli

All participants were shown exactly the same set of test stimuli. The test stimuli used were signatures extracted from a blind validation trial conducted by the university in 2004. None of the subjects had participated in this trial. The signatures were made onto white paper and scanned into a computer using Adobe Photoshop version 7.0 software and saved as 8-bit grayscale 1024 × 768 pixel jpeg files (10 max quality) to enable displaying on a Tobii 1750 eye tracking system. Four genuine signatures (provided by a single specimen provider) were used to allow subjects to familiarize themselves with the normal handwriting of the specimen provider prior to stimulus presentation and were presented on screen with each test signature. The specimen provider also provided eight signatures that were intentionally disguised which were randomly incorporated into the stimulus set along with eight forged signatures provided by eight different subjects that had an opportunity to copy the signature of the specimen provider (Fig. 1).

Each of the forgers was provided with three original normal samples of the signature written by the specimen writer randomly drawn from a pool of 200 samples. Forgers were instructed that they could use any or all of the supplied specimen signatures as models for their forgeries. They were asked to reproduce the signature as accurately as possible. Forgers were also instructed that their forgeries must be unassisted (not tracings). Each forger was provided with a pen and a booklet. The booklet contained pages divided into spaces. These spaces were numbered 1–20. The forgers were asked to inspect the genuine signature and forge it 20 times. From these the eight simulated signatures included in the stimulus set were chosen on the basis that one or more opinions concerning the signature provided by participants in the 2004 signature validation trial were erroneous. That is, that at least one opinion provided by FDEs in the 2004 trial was that the signature was genuine.

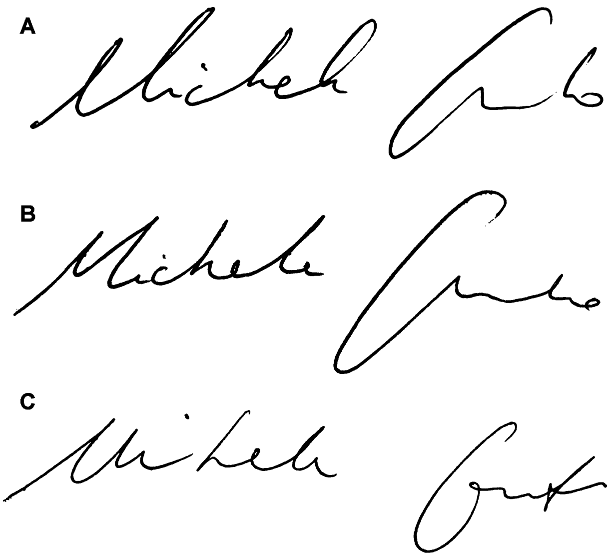

FIG. 1—*Examples of (A) genuine, (B) disguised, and (C) forged signatures.*

None of the examiners reported in the study had prior knowledge of the signature (through examination or association with the individual) nor did they assist in the collection or preparation of the test stimuli.

Data Recording

Eye movements were recorded with a Tobii 1750 binocular eye tracking system (Tobii Technology, Stockholm, Sweden). Subjects viewed stimuli on the TFT screen from a distance of 57 cm so that the width of a typical questioned signature subtended a horizontal visual angle of *c*. 28°. Calibration for each subject to a 16-point reference grid was done providing for a resolution of subject gaze to better than 0.5° of visual angle as described in detail by Dyer et al. (6).

Eye fixations were determined using criteria of eye position remaining within a 50 pixel area for a time of greater than 100 ms (8). Data collection, fixation measurements, RTs, and analysis of areas of interest (AOI) data were determined with Tobii Clearview 2.5.1 software. 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).

Learning Phase

Prior to experiments, the control subjects as a group were given a 30-min talk on what normal, disguised, and forged signatures were 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 variability within a signature. The control subjects were made aware of the potential consequences of a FDE making errors in a real life scenario. Our previous study (6) showed that this initial briefing promotes a high level of motivation in the control subjects to perform well at a signature evaluation task.

Subjects then each viewed the four specimen signatures sequentially for 10 sec each while displayed on the Tobii 1750 monitor. Eye movements were recorded during this phase but the data did not reveal any information that was deemed important to report in this manuscript. Each subject was then given a further 5 min to examine the four specimen signatures presented simultaneously on a piece of A4 paper. Subjects were informed that during the following eye movement experiment, the four specimen signatures would be displayed at the bottom of the screen while the questioned signature would be displayed at the top of the screen (Fig. 3*A*). 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 similar to the normal operating environment of FDEs.

Testing Phase

The eye movements, RT, and performance at the signature trial of the FDE and control subjects were evaluated while making decisions about 16 questioned signatures. Subjects were informed that each of the 16 questioned signatures was either a forgery or a disguised signature (i.e., none of the questioned signatures were genuine). The FDEs participating in the study do not provide case work opinions using these terms but agreed to the terminology for the purpose of the experiment.

Prior to the presentation of a questioned signature, a subject was 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 a qwerty keyboard to view a stimulus.

To quantify subject performance in the signature trial, a penalty scoring system was used where subjects were awarded 1 point for a correct decision, but an incorrect decision incurred a 1 point penalty (6,13). Subjects were also allowed to respond inconclusively, for which they scored zero points on that stimulus. This type of scoring system with an opportunity for an inconclusive decision models the real world working scenario of FDEs, and the addition of a penalty scoring system is thought more likely to reveal differences in subject performance (13).

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 so as to make an accurate decision about whether it is disguised or forged. 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: (1) disguised, (2) forged, or (3) if you are unsure you may give an inconclusive answer. If you correctly identify the signature as either disguised or forged you will score 1 point, but an incorrect allocation receives a penalty of −1 point. An inconclusive response receives a score of 0 points."

After a decision had been made about a questioned signature, a screen was presented outlining the three options (disguised, forged, or inconclusive) and the subject verbally gave a response to the experimenter. This procedure was done so that the subjects did not have to divert attention from the screen to select options on a keyboard. Using this procedure, each subject sequentially viewed the 16 questioned stimuli.

During the testing, the experimenter was sitting behind the subject. No visual or auditory cues were provided that could be interpreted as feedback as to the validity of any response provided by the subject.

TABLE 1—*Mean number of calls per subject for either control or forensic document examiner (FDE) subjects while making decisions about whether simulated signatures were either forged or disguised.*

	Control	FDE
Forged-correct	5.5	5.1
Forged-incorrect	1.6	0.5
Forged-inconclusive	0.9	2.4
Disguised-correct	3.6	5.0
Disguised-incorrect	3.5	1.4
Disguised-inconclusive	0.9	1.6

Results

KQ1: Can Subjects Correctly Discriminate Between Disguised and Forged Signatures, and Do FDEs Show an Expertise Effect Compared to the Control Subjects?

Table 1 provides the mean number of opinions (calls) for the control and FDE groups for the forged and disguised signatures. While control and FDE groups made a similar proportion of correct calls for either forged or disguised signatures, difference in overall performance can be attributed to a much higher rate of incorrect calls by control subjects for both forged (3.2 times as many errors) and disguised (2.6 times as many errors). The reduced incorrect scores relative to the control group are likely to result from FDEs' higher number of inconclusive opinions. This finding is consistent with that of Sita et al. (5), where the expertise of FDEs is thought to be best characterized by their conservatism indicating that expertise is likely to lead to clearer decisions about when the correct identification of a signature is too difficult to evaluate from the available information.

The same data can be converted to the penalty scoring system. Using this system, if subjects could not discriminate between stimuli and randomly assigned answers or answered inconclusive, then the resulting score would have a mean of zero (maximum score would be 16). It was found that the control group scored a mean of 4.0 (SEM = 1.8) and thus showed a weak ability to differentiate between stimuli, whilst the FDEs scored a mean of 8.5 (SEM = 1.2) and performed significantly better on this task than the control group (independent samples t-test, $t = 2.465$, df = 18, $p = 0.024$). Thus, there is evidence that subjects can discriminate between forged and disguised signatures, and that the expertise effect of trained FDEs reported in previous studies (1–6) is also evident in this type of task. Of interest, one FDE subject was able to achieve a very high score of 13 on this test (13 correct calls, three inconclusive calls, and zero errors), indicating that with the test signature under study, it was possible for the visual and cognitive system of a highly skilled FDE to very accurately discriminate between the forged and disguised signatures. The high score for this FDE subject could not have been due to the task being too easy, because from the control group, four subjects scored zero or less on the test, and overall the control group only showed a weak ability to perform the task.

KQ2: Do FDEs Perform Better than Control Subjects on Forged and/or Disguised Signatures?

To understand why FDE subjects performed better than the control group for KQ1, it is possible to evaluate the frequency with which the subject groups made errors on either disguised or forged signatures. The data were normalized for group size and the control group made 76.0% of the total number of erroneous opinions on forged signatures, and 71.8% of the total number of erroneous opinions on disguised signatures. Both groups made more of their errors when calling disguised signatures (68.8% and 73.3% for control and FDEs, respectively). The higher score by the FDE group (see KQ1) was possibly also influenced by a higher rate of inconclusive calls compared to the control group (2.6 and 1.8 times as many calls of inconclusive for forged and disguised signatures, respectively), indicating that expertise is likely to lead to clearer decisions about when the correct identification of a signature is too difficult to evaluate from the available information. Thus, FDEs performed better than control subjects on both disguised and forged signatures, and both groups found disguised signatures the more difficult type of signature to call.

KQ3: Do FDEs and Control Subjects Choose to Allocate Different Amounts of Time for Making Decisions About Disguised and Forged Signatures?

This analysis considered the two groups (FDEs or control) and the two types of signatures (disguised or forged) with a dependent variable of RT. Figure 2 shows the mean RT (±SEM) for making a decision about the signatures. Data were analyzed with a two-way between group ANOVA. There was no significant effect on RT depending upon whether subjects were making a decision about disguised or forged signatures [$F(1,36) = 0.068$, $p = 0.796$, Partial Eta squared = 0.002]. There was no interaction between groups [$F(1,36) = 0.069$, $p = 0.795$, Partial Eta squared = 0.002]. However, there was a significant difference on RT between the groups [$F(1,36) = 6.211$, $p = 0.017$, Partial Eta squared = 0.147] showing that for this type of task FDE subjects spent significantly longer making a decision than control subjects.

KQ4: Do FDEs or Control Subjects Look at Signature Simulations Differently When Making a Decision?

Figure 3*A* shows an example of the eye movements of a FDE viewing a disguised signature and the four reference genuine signatures. In this instance, a correct opinion was formed by the examiner. Also shown on this signature are the 15 different AOI that were considered in the study (AOI boxes were not seen by the subjects). Figure 3*B* shows a hotspot plot of the overall attention of FDE and control subjects to a forged signature. This qualitative data indicate a similar attention level to the different components of the signatures by both groups; however, to better understand if the subject group looked at stimuli differently during the test, 15 different AOI were selected to represent the different regions of a signature (6). Figure 3*C* shows the mean fixations to the different regions of the signature by FDEs and control subjects. These data also indicate that FDEs use a very similar search strategy to the control subjects, suggesting that expertise is not mediated by the particular features that subjects look at, but how the viewed information is processed by a subject. The data for this experiment are also very similar to data reported (6) for a genuine signature-simulated signature task (see Fig. 6B in Dyer et al. [6]), indicating that the search strategy used by subjects is largely independent of the different task requirements in the two studies.

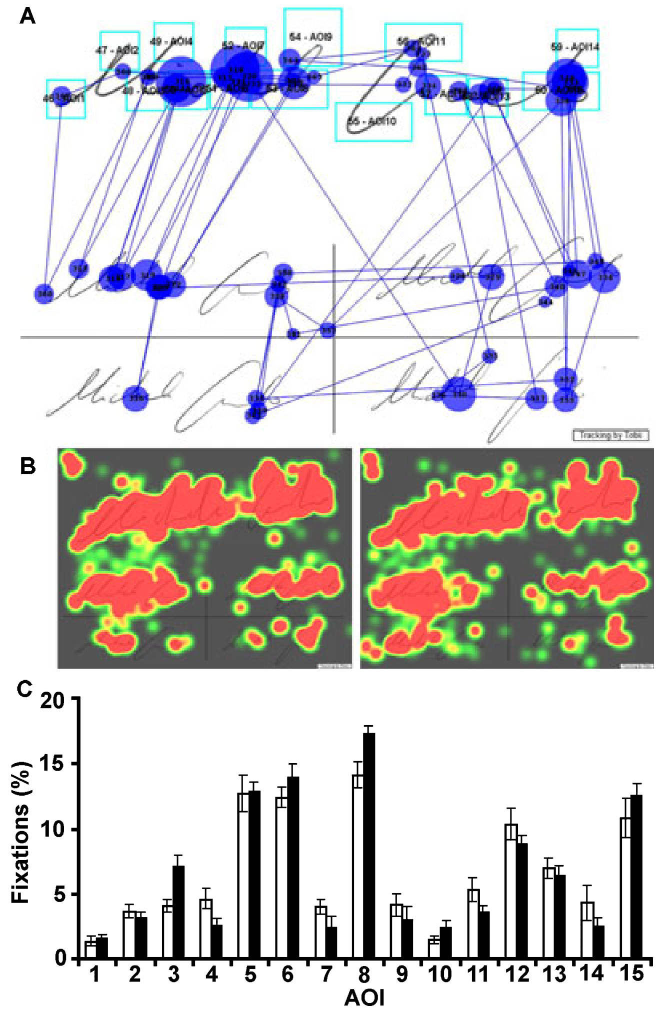

FIG. 3—*(A) Example of a disguised signature, and the four genuine specimen signatures made available to subjects for each questioned signature. The 15 different areas of interest (AOI) considered in this study are superimposed as rectangles over the forged signature (these AOI were not visible to subjects). Eye movements of a forensic document examiner who correctly called this signature are also shown (lines show saccades, plots are fixations, size of plot indicates normalized fixation time). (B) Hotspot plot of the mean fixations of control subjects (left) and forensic document examiners (right) when viewing a forged signature. (C) Mean percentage fixations to AOI1-15 of control subjects (white) and forensic document examiners (black) when making decisions about signatures (see text for details).*

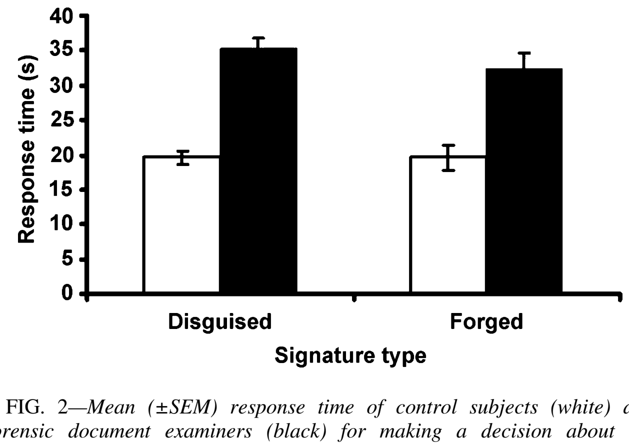

FIG. 2—*Mean (±SEM) response time of control subjects (white) and forensic document examiners (black) for making a decision about the signatures.*

Discussion

This study shows that, under controlled conditions, qualified FDEs were able to make correct decisions about whether one individual's signature was either the product of a disguise or forgery process. There was also evidence of an expertise effect as the mean score of FDEs was significantly greater than the control group. It should be noted that the significant differences in performance were unlikely to be due to differences in motivation, as it was shown (6) that when control subjects were fully briefed as to the importance of making calls about signatures, that both FDE and control subjects exhibited similar behavior when making inconclusive calls (i.e., both groups spent significantly longer evaluating difficult signatures, indicating a high level of motivation to do well in the control group).

Evidence from the RT data (Fig. 2) shows that the expertise effect in FDEs was due to these subjects taking into account multiple sources of visual information about signature stimuli. Dyer et al. (6) showed that the expertise in FDEs for a genuine-simulated signature type task was due to FDEs "parallel processing" (taking multiple pieces of information into account before arriving at a decision [14]), while control subjects showed evidence of making decisions on the basis of single salient features. The longer RT recorded in this study indicates that FDEs were looking for multiple sources of information to process so as to make an accurate decision. Interestingly however, the data on the AOI to which FDE and control subjects paid attention during their evaluation of a signature were remarkably similar (Figs. 3*B* and 3*C*); these eye movement data are also very similar to data for the genuine-simulated task presented in Dyer et al. (6). This indicates that both subject groups are viewing the necessary features of a signature required to make accurate decisions, and that expertise lies in the domain of the cognitive processing of the information rather than the sourcing of the information. Consistent with data reported in Dyer et al. (6), the current study supports the view that FDEs do not use any particular topdown (knowledge based) scan-path strategy of searching stimuli (e.g., Fig. 3*A* shows a typical but rather anarchic search of signature by a FDE subject), but FDEs use bottom-up (i.e., stimulus driven) anarchic or reflexive type eye movements to search for evidence. Recent work on information processing in humans suggests that this is likely to be a more efficient method of searching for visual information compared to preplanned volitional search strategies (15).

Both subject groups made more of their errors when calling disguised signatures (68.8% and 73.3% for control and FDEs, respectively); however, the overall rate of FDEs errors (28.2% of total errors) was considerably lower than the control group (71.8%). This was partially because FDEs formed a greater number of inconclusive opinions. The finding that FDEs had a higher number of inconclusive calls than control subjects is in line with previous reports on expertise such that the skill is thought to be best characterized by what FDEs do not say rather than what they do say (5). These data thus suggest that to make accurate calls, subjects need to be able to learn how to build a cumulative model of the features in a signature where the significance of evidence from individual features is weighted in parallel by evidence from other features that either support or refute an authorship proposition.

Although a number of studies have reported the type and frequency of occurrence of handwriting features that may be used as predictors of whether questioned signatures are genuine, forged, or disguised (16–20), there is still a relative dearth of published information regarding how reliably FDEs can discriminate between forged and disguised signatures. This current study involved a task requiring discrimination between disguised and forged signatures. It provides preliminary evidence that FDEs can, under the conditions described, carry out the task (albeit with some associated error) and that FDEs showed an expertise effect over well-motivated control subjects. Limitations of the study exist. This study is not a validation trial; it involved only eight FDEs and only one individual's signature (and their disguise strategy) was the subject of this study. We cannot rule out the possibility that an expertise effect observed in this limited study may not be shown in a study involving larger numbers of examiners and control subjects or where the disguised signatures were provided by another writer. It can be observed from inspection of the disguised signature in Fig. 1 that a large proportion of the signature used in this study remained relatively unchanged. This made the task for the FDEs much more straightforward than it may have been if a heavier disguise strategy had been used. It is clear that many more studies, using different subjects' signatures, would need to be investigated to sample the myriad of disguise behaviors that could potentially be used and the ability of FDEs to detect them.

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. Kam M, Wetstein J, Conn R. Proficiency of professional document examiners in writer identification. J Forensic Sci 1994;39:5–14.
2. Kam M, Fielding G, Conn R. Writer identification by professional document examiners. J Forensic Sci 1997;42:778–86.
3. 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.
4. Kam M, Gummadidala K, Fielding G, Conn R. Signature authentification by forensic document examiners. J Forensic Sci 2001;46:884–8.
5. Sita J, Found B, Rogers D. Forensic handwriting examiners' expertise for signature comparison. J Forensic Sci 2002;47:1117–24.
6. Dyer AG, Found B, Rogers D. Visual attention and expertise for forensic signature analysis. J Forensic Sci 2006;51:1397–404.
7. Yarbus AL. Eye movements and vision. New York: Plenum, 1967.
8. Duchowski AT. Eye tracking methodology: theory and practice. London: Springer-Verlag, 2003.
9. Martinez-Conde S, Macknik SL, Hubel DH. The role of fixational eye movements in visual perception. Nature Neurosci 2004;5:229–40.
10. Kustov AA, Robinson DL. Shared neural control of attentional shifts and eye movements. Nature 1996;384:74–7.
11. Henderson J, Falk R, Minut S, Dyer F, Mahadevan A. Gaze control for face learning by humans and machines. Adv in Psych 2001;130:463–81.
12. Busey TA, Vanderkolk JR. Behavioral and electrophysiological evidence for configural processing in fingerprint experts. Vision Res 2005;45:431–48.
13. 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.
14. Townsend JT, Wenger M. The serial-parallel dilemma: a case study in linkage of theory and method. Psychonomic Bullet and Rev 2004;11:391–418.
15. Wolfe JM, Alvarez GA, Horowitz TS. Attention is fast but volition is slow. Nature 2000;406:691.
16. Leung SC, Cheng YS, Fung HT, Poon NL. Forgery I. Simulation. J Forensic Sci 1993;38:402–12.
17. van Gemert AWA, van Galen GP. Dynamic features of mimicking another person's 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.
18. Herkt A. Signature disguise or signature forgery. J Forensic Sci Soc 1986;26:257–66.
19. Wendt GW. Statistical observations of disguised signatures. J Am Soc Quest Doc Exam 2000;3:19–27.
20. 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.

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