
THE SKILL OF A GROUP OF FORENSIC DOCUMENT EXAMINERS IN EXPRESSING HANDWRITING AND SIGNATURE AUTHORSHIP AND PRODUCTION PROCESS OPINIONS

Bryan Found^{1,2} Doug Rogers¹ and Allan Herkt³

Abstract. *Between March 1998 and June 2001, the six members of the New Zealand Police Document Examination Section completed six blind forensic handwriting and signature trials where the identity of the questioned writings were known by the experimenters but not by the document examiners. A total of 7494 opinions were expressed by the examiner group regarding the authorship of questioned handwriting and signature samples. Of these opinions, 2700 were correct, 11 were erroneous and 4783 were inconclusive. This translates into an overall raw error score of 0.1% of opinions, and a 'called error' score (one calculated by excluding the inconclusive data) of 0.4% of opinions. In addition, a total of 2982 opinions were expressed by the group on whether questioned signatures were the product of a simulation process. The group did not express any opinions that a simulated signature was the product of a genuine writing process, nor did the examiners express any opinions that a genuine signature was the product of a simulation process. Thus, for process opinions there was a zero error rate for the group. This paper overviews the individual and group opinion data associated with each of the trials. The results provide significant support to the validity and strong probative value of the skill that has been claimed by this group of examiners.*

Reference: Bryan Found, Doug Rogers, Allan Herkt (2001, Vol. 14 – reprinted and reformatted). The Skill of a Group of Forensic Document Examiners in Expressing Handwriting and Signature Authorship and Production Process Opinions. J. Forensic Document Examination, Vol. 29, pp. 73 - 82.

Keywords: Forensic Expertise Profiling Laboratory, handwriting, error rates, document examiners' skill

1. Introduction

The desire for the New Zealand Police Document Examination Section (NZPDES) to expose themselves to extensive and sustained blind testing of their claimed skill in forensic handwriting identification can be sourced originally to the concerns raised within the now historically significant Risinger, Denbeaux and Saks (1989) publication on the topic. Specifically

of interest to the current study were the following criticisms raised by these authors:

- No court anywhere has ever explicitly considered and passed on its (handwriting identification) claim to validity.
- There exist almost no studies of its claims in any academic literature.
- Such studies as have been conducted, published and unpublished, raise serious questions as to its validity.
- The law has resisted requiring presentation of the asserted expertise in ways that would expose its validity problems.

It appeared obvious from the lack of published validation trials internationally that the criticisms raised were valid and, more importantly, could largely be addressed through the administration of

1 Forensic Expertise Profiling Laboratory, School of Human Bio.sciences, La Trobe University, Victoria, 3086, Australia.

2 Document Examination Team, Victoria Forensic Science Centre, Forensic Drive, Macleod, Victoria, 3085, Australia.

3 Document Examination Section, Wellington Central Police Station, Victoria Street, PO Box 693, Wellington, New Zealand

blind tests. The Special Advisory Group (Document Examination), which represents police and government document examiners in Australia and New Zealand, has been collaborating with the now Forensic Expertise Profiling Laboratory (La Trobe University, Australia) to carry out such trials since the mid 1990's. The NZPDES is one of the participants in this process and, in addition to the direct testing component of the trials, has participated in researching other aspects of forensic handwriting examination such as method development and documentation (Found & Rogers, 1999).

Since our collaborative research interest in validation testing of forensic handwriting examiners has commenced, a number of relevant events have occurred and important studies have been published on the topic. Judge McKenna in the *Starzecpyzel* decision (United States v. *Starzecpyzel*, 1995) stated that, "The Daubert hearing established that forensic document examination which clothes itself with the trappings of science, does not rest on carefully articulated postulates, does not employ rigorous methodology, and has not convincingly documented the accuracy of its determinations. Forensic handwriting identification was, in spite of this statement, recognized as a practical skill and, therefore, held to be admissible in evidence. Risinger and Saks (1996) argued that an implication of this decision was the potential for plummeting validation standards for admissibility, which may result in the burden falling on the opponent, "to prove affirmatively that the skilled witnesses cannot do what they claim they can do". Clearly this is not a position that forensic examiners and legal specialists would be comfortable with. These authors then open the window to allowing some resolution of the concerns by stating that, "science can examine the dependability of such a process (handwriting identification) even when the process is not a science." Science has commenced to do so.

Forensic validation studies have been reported by Kam, Fielding and Conn (1997), Kam, Wetstein and Conn (1994) and Found, Sita and Rogers (1999). These studies have provided some support for the expertise claimed by practitioners, or at least those that have been tested within the trials, in terms of it being real and demonstrable. In each case an error score has been reported. It is this error score that is relevant

to document examiner client groups, particularly the judicial system. It is the magnitude of the error score that best dictates the probative value of the evidence being presented.

In order to assess the magnitude of any error, the Forensic Expertise Profiling Laboratory has adopted a philosophy of testing, largely based on the criticisms of this field historically. Specifically, our philosophy attempts to address the following guiding statements:

1. "The level of correctness of the assertions made by examiners from day to day casework is not likely to prove to be a credible source for the (validation) data needed" (Huber & Headrick, 1999).
2. "A process such as handwriting identification presents a number of potential subtasks dealing with variables such as writing instruments, forgery of various sorts, age, health and so forth. No single test can map the abilities of any one practitioner, or any group of practitioners" (Risinger & Saks, 1996).
3. "A great many tests... would be necessary to know what, if anything, (examiners) can do accurately, and under what conditions" (Risinger & Saks, 1996).
4. "A complete testing regime would have tests which covered the entire spectrum of conditions and difficulties" (Risinger & Saks, 1996).

The results presented in this paper represent the NZPDES results on trials completed between March 1998 and June 2001. It should be noted that this laboratory has historically recorded one of the lowest error scores amongst the groups participating in our trials. In spite of this, these examiners were keen to bring into evidence issues surrounding the probative value of the skill that they had traditionally claimed. This overview of their testing results does not contain the minutia of details associated with the construction of each of the trials, a task we felt was best left to reports concerning each of the trials independently with the inclusion of all participants' data.

2. Overview of the trials included in this report

2.1. Trial 1

This trial was an upper-case handwriting trial. Examiners were provided with original samples of questioned and specimen writings. The specimen material was produced by three individuals. Examiners were required, amongst other tasks, to compare specimen writings with a total of 134 questioned samples. The questioned samples were requested normal (written by specimen writer and other writers), disguised (written by the specimen writer and other writers) or simulated writings (written by specimen writer and other writers). For each of the questioned samples examiners were required to express a '*direction of identification*' opinion, a '*direction of exclusion opinion*', or an *inconclusive opinion*. Opinions were marked as either correct, incorrect or inconclusive. The opinions of each of the examiners were not subjected to a peer-review process.

2.2 Trial 2

This trial incorporated both questioned signatures and handwriting. Examiners were provided with 30 questioned documents (withdrawal slips), each with a signature, and 5 distinct samples of handwriting for opinion. Samples from 2 individuals were provided for comparison purposes. Each distinct sample of the questioned handwriting was written by one or other of the specimen writers. Each questioned signature was either a genuine signature by a specimen writer or a simulation. For each of the questioned handwriting samples examiners were required to express a direction of identification opinion, a direction of exclusion opinion, or an inconclusive opinion. For each of the questioned signatures examiners were required to express a direction of identification opinion, a simulation opinion, or an inconclusive opinion. Opinions were marked as either correct, incorrect or inconclusive. The opinions of each of the examiners were subjected to a peer-review process.

2.3 Trial 3

This trial was a handwriting trial. Examiners were provided with original samples of questioned and specimen writings. One individual produced

the specimen material. Examiners were required, amongst other tasks, to compare the specimen writings with a total of 250 questioned samples. The questioned samples were requested normal (written by specimen writer and other writers), disguised (written by specimen writer and other writers), or simulated writings (written by specimen writer and other writers). For each of the questioned samples examiners were required to express a direction of identification opinion, a direction of exclusion opinion, or an inconclusive opinion. Opinions were marked as either correct, incorrect or inconclusive. The opinions of each of the examiners were subjected to a peer-review process.

2.4 Trial 4

This trial was a signature trial. Examiners were provided with examples of a specimen signature and were required to compare the specimen signatures with a total of 80 non-original (photocopied) questioned signatures. The questioned signatures comprised requested normal signatures and simulated signatures (written by the specimen writer and other writers). For each of the questioned signatures examiners were required to express a '*direction of identification*' opinion, a '*simulation*' opinion, or an inconclusive opinion. Opinions were marked as either correct, incorrect or inconclusive. The opinions of each of the examiners were not subjected to a peer-review process.

2.5 Trial 5

This trial was a signature trial. Examiners were provided with examples of a specimen signature and were required to compare them with a total of 260 original questioned signatures. The questioned signatures comprised requested normal signatures and simulated signatures (written by the specimen writer and other writers). For each of the questioned signatures examiners were required to express a direction of identification opinion, a simulation opinion, or an inconclusive opinion. Opinions were marked as either correct, incorrect or inconclusive. The opinions of each of the examiners were subjected to a peer-review process.

2.6 Trial 6

This trial was a signature trial. Examiners were provided with examples of a specimen signature and were required to compare them with a total of 250 questioned signatures. All signatures were high resolution scanned images, printed using a laser printer. The questioned signatures comprised requested normal signatures and simulated signatures (written by the specimen writer and other writers). For each of the questioned signatures examiners were required to express a direction of identification opinion, a direction of exclusion opinion, a simulation opinion, or an inconclusive opinion. Opinions were marked as either correct, in correct or inconclusive. The opinions of each of the examiners were subjected to a peer-review process.

3. Definition of scores used in this report

The development of methodology (Found & Rogers, 1999) was occurring during the administration of the trials described in this report. Incorporated in this process were changes in the definition of terms used to express opinions to more closely align to the reporting philosophies articulated in Evett (1998). To facilitate the compilation of results in this study, opinions were either treated as correct (in spite of the level of support for the proposition), incorrect (in spite of the level of support for the proposition), or inconclusive.

Examiners' authorship responses (opinion units) were marked as correct, incorrect or inconclusive. These marks were then analyzed to produce scores for each of the different questioned handwriting types (normal writing by the specimen writer, disguised writing by the specimen writer, simulated writing by the specimen writer, simulated writing not by the specimen writer, normal writing not by the specimen writer, and disguised writing not by the specimen writer). The scores are presented as numbers of opinions or as percentages, the latter representing opinion rates. The following definitions of the score categories are used in subsequent results tables in this report.

3.a # Correct

The number of authorship opinions that were correct.

3.b # Error

The number of authorship opinions that were incorrect.

3.c # Inconclusive

The number of authorship opinions that were inconclusive.

3.d % Correct

The number of correct authorship opinions divided by the total number of authorship opinions (expressed as a percentage).

3.e % Error

The number of incorrect authorship opinions divided by the total number of authorship opinions (expressed as a percentage).

3.f % Inconclusive

The number of inconclusive authorship opinions divided by the total number of authorship opinions (expressed as a percentage).

3.g % Correct called

The number of correct authorship opinions divided by the sum of the correct and erroneous authorship opinions (expressed as a percentage).

3.h % Error called

The number of incorrect authorship opinions divided by the sum of the correct and erroneous authorship opinions (expressed as a percentage).

The called scores do not include inconclusive opinions and, therefore, equate to a number that reflects the opinion rate when an examiner is expressing an opinion that is other than inconclusive.

4. Results

The results of all six of the authorized forensic document examiners with the NZPDES are included in this report. A total of 7494 authorship opinions

Writing Type	Opinion Scores							
	# correct	# error	# inc.	% correct	% error	% inc.	% correct called	% error called
Signatures	899	0	2791	23.4	0	76.6	100	0
Handwriting	1801	11	1992	47.3	0.3	52.4	99.4	.06
Handwriting and Signatures	2700	11	4783	36.0	0.1	63.8	99.6	.04

TABLE 1. Summary of authorship opinion unit scores for all opinions expressed in the trials for signature, handwriting and combined signature and handwriting samples.

Writing Type	Opinion Scores							
	# correct	# error	# inc.	% correct	% error	% inc.	% correct called	% error called
Normal by specimen writer	707	1	474	59.8	.01	40.1	99.9	.01
Normal not by specimen writer	525	1	614	46.1	.01	53.9	99.8	0.2
Disguise by specimen writer	301	1	94	76.0	0.3	23.7	99.7	.03
Disguise not by specimen writer	82	0	266	23.6	0.0	76.4	100.0	0.0
Simulated not by specimen writer	154	3	383	28.5	0.6	70.9	98.1	1.9
Simulated by specimen writer	32	5	161	16.2	2.5	81.3	86.5	13.5
Handwriting totals	1801	11	1992	47.3	.03	52.4	99.4	0.6

TABLE 2. The authorship opinion scores for the examiner group across all handwriting types represented across each of the three handwriting text trials.

have been expressed by the six examiners in the group. Five of the six examiners completed all six of the trials. One examiner did not examine one of the trials.

5. Handwriting text results

There were 3804 authorship opinions expressed by the group on handwriting text comparisons. Table 2 provides the authorship opinion scores for the examiner group across all handwriting types represented across each of the three handwriting text trials. As can be observed, the ‘potential or estimated error rate’ for handwriting types varies according to the questioned writing type. The % error is <1% for all handwriting text types except those samples that are simulated by the specimen writer, where the error is found to be 2.5% (a called error rate of 13.5%).

Although the two simulation writing types have the highest error rates of the handwriting types, this must be balanced with the corresponding % inconclusive scores. These two categories of writing exhibit high % in conclusive scores, which indicates that examiners are more conservative when expressing opinions regarding samples of this type. In addition, the 5 errors made calling a simulated sample of writing (by the specimen writer) as not written by the specimen writer, were all made on a non peer reviewed trial, and 4 of the 8 errors were made by one individual.

Table 3 provides the scores for authorship opinions expressed for each examiner across all handwriting types represented in each of the three handwriting text trials. Note that the handwriting types are represented by the codes SP (by specimen writer), NSP (not by specimen writer), DSP (disguised by specimen writer),

examiner number	Writing Type	Opinion Scores							
		# correct	# error	# inc.	% correct	% error	% inc.	% correct called	% error called
1	DNSP	10	0	48	17.2	0.0	82.8	100	0
2	DNSP	12	0	46	20.7	0.0	79.3	100	0
3	DNSP	7	0	51	12.1	0.0	87.9	100	0
4	DNSP	10	0	48	17.2	0.0	82.8	100	0
5	DNSP	23	0	35	39.7	0.0	60.3	100	0
6	DNSP	20	0	38	34.5	0.0	65.5	100	0
1	DSP	50	0	16	75.8	0.0	24.2	100	0
2	DSP	50	0	16	75.8	0.0	24.2	100	0
3	DSP	50	0	16	75.8	0.0	24.2	100	0
4	DSP	50	1	15	75.8	1.5	22.7	98.0	2.0
5	DSP	51	0	15	77.3	0.0	22.7	100	0
6	DSP	50	0	16	75.8	0.0	24.2	100	0
1	NSP	95	1	119	44.2	0.5	55.3	99.0	1.0
2	NSP	34	0	31	52.3	0	47.7	100	0
3	NSP	93	0	122	43.3	0	56.7	100	0
4	NSP	98	0	117	45.6	0	54.4	100	0
5	NSP	106	0	109	49.3	0	50.7	100	0
6	NSP	99	0	116	46	0	54.0	100	0
1	SNSP	24	0	66	26.7	0	73.3	100	0
2	SNSP	22	0	68	24.4	0	75.6	100	0
3	SNSP	22	0	68	24.4	0	75.6	100	0
4	SNSP	25	0	65	27.8	0	72.2	100	0
5	SNSP	31	2	57	34.4	2.2	63.3	93.9	6.1
6	SNSP	30	1	59	33.3	1.1	65.6	96.8	3.2
1	SP	128	0	94	57.7	0	42.3	100	0
2	SP	62	0	10	86.1	0	13.9	100	0
3	SP	129	0	93	58.1	0	41.9	100	0
4	SP	130	0	92	58.6	0	41.4	100	0
5	SP	131	0	91	59.0	0	41.0	100	0
6	SP	127	1	94	57.2	0.5	42.3	99.2	0.8
1	SSP	6	0	27	18.2	0	81.8	100	0
2	SSP	6	0	27	18.2	0	81.8	100	0
3	SSP	3	0	30	9.1	0	90.9	100	0
4	SSP	1	4	28	3.0	12.1	84.8	20	80
5	SSP	10	0	23	30.3	0	69.7	100	0
6	SSP	6	1	26	18.2	3.0	78.8	85.7	14.3
	Totals	1801	11	1992	47.3	0.3	52.4	99.4	0.6

TABLE 3. The authorship opinion scores expressed for each examiner across all handwriting types represented in each of the 3 handwriting text trials.

Signature Type	Opinion Scores							
	# correct	# error	# inc.	% Correct	% error	% Inc.	% correct Called	% error called
Normal by specimen writer	712	0	27	96.3	0	3.7	100	0
Disguise by specimen writer	122	0	286	29.9	0	70.1	100	0
Simulated not by specimen writer	65	0	2322	2.7	0	97.3	100	0
Simulated by specimen writer	0	0	156	0	0	100	n/a	n/a
Totals	899	0	2791	24.4	0.0	75.6	100	0.0

TABLE 4. The opinion scores for all signature types represented in the trials.

DNSP (disguised not by specimen writer), SNSP (simulated not by specimen writer and SSP (simulated by specimen writer).

6. Signature results

There were 3690 authorship opinions expressed by the group on signature comparisons. Table 4 provides the grouped authorship opinion unit scores for all signature types represented in the trials. The scores are for the group of examiners as a whole, where all of the same questioned signature types from the different trials have been combined. As can be observed, no error has yet to be recorded by the group regarding the authorship of questioned signatures. It should be noted, however, that the group has not recorded any opinions where the specimen writer was excluded from having written a particular signature.

Table 5 provides the opinion scores for each examiner across all signature types represented in each of the signature trials. Note that the signature types are represented by the codes SP (by specimen writer), DSP (disguised by specimen writer), SNSP (simulated not by specimen writer and SSP (simulated by specimen writer). Although there were no errors in the direction of identification or exclusion, examiners were 100% inconclusive as to whether or not the specimen writer wrote any of the signatures that were the product of a simulation process.

7. Signature process

The determination of a writing process is not about whether or not the writer of the specimens did or did not write a particular entry, but is an opinion

regarding the writing behaviour itself. From trials 2, 4, 5 and 6 it is possible to extract opinions by the group on whether or not examiners believed that questioned signatures were genuine (where it can be assumed that the examiners were of the opinion that the signatures were not the product of a simulation process), or produced using a simulation (or imitation) process. In many instances the authorship of simulated signatures is not determinable due to the difficulty in excluding the proposition that the specimen writer did not simulate his or her own signature for the purposes of denial at a later date. An opinion that a signature was produced using a simulation process can, however, be of assistance to the judiciary.

A total of 2982 opinions were expressed by the group on genuine and simulated signature samples. The scores for these process opinions are shown in Table 6. As can be observed, the group did not express any opinions that a simulated signature was the product of a genuine writing process, nor did the examiners express any opinions that a genuine signature was the product of a simulation process.

Table 7 provides the process opinion scores for each examiner for genuine and simulated signatures.

8. Discussion

There are many aims in conducting skill research of this type. Examples include whether the skills claimed by a particular group are real, what the error rate in decision making by individuals and groups is, and what the relationship is between results' profiles from different laboratories, including all the variables associated with qualifications, training programs, experience etc. At this stage in the documentation of

examiner number	Writing Type	Opinion Scores							
		# correct	# error	# inc.	% correct	% error	% inc.	% correct called	% error called
1	DSP	21	0	47	30.9	0.0	69.1	100	0.0
2	DSP	20	0	48	29.4	0.0	70.6	100	0.0
3	DSP	20	0	48	29.4	0.0	70.6	100	0.0
4	DSP	20	0	48	29.4	0.0	70.6	100	0.0
5	DSP	20	0	48	29.4	0.0	70.6	100	0.0
6	DSP	21	0	47	30.9	0.0	69.1	100	0.0
1	SNSP	0	0	400	0.0	0.0	100	n/a	n/a
2	SNSP	0	0	387	0.0	0.0	100	n/a	n/a
3	SNSP	0	0	400	0.0	0.0	100	n/a	n/a
4	SNSP	0	0	400	0.0	0.0	100	n/a	n/a
5	SNSP	0	0	400	0.0	0.0	100	n/a	n/a
6	SNSP	0	0	400	0.0	0.0	100	n/a	n/a
1	SP	123	0	3	97.6	0.0	2.4	100	0.0
2	SP	107	0	2	98.2	0.0	1.8	100	0.0
3	SP	118	0	8	93.7	0.0	6.3	100	0.0
4	SP	121	0	5	96	0.0	4.0	100	0.0
5	SP	124	0	2	98.4	0.0	1.6	100	0.0
6	SP	119	0	7	94.4	0.0	5.6	100	0.0
1	SSP	0	0	26	0.0	0.0	100	n/a	n/a
2	SSP	0	0	26	0.0	0.0	100	n/a	n/a
3	SSP	0	0	26	0.0	0.0	100	n/a	n/a
4	SSP	0	0	26	0.0	0.0	100	n/a	n/a
5	SSP	0	0	26	0.0	0.0	100	n/a	n/a
6	SSP	0	0	26	0.0	0.0	100	n/a	n/a

TABLE 5. The opinion scores expressed for each examiner across all signature types represented in the trials.

forensic handwriting examiners' skills, the most critical factors being investigated were the characterisation of examiners' skill at providing identification/exclusion evidence on different categories of writing, and the potential error rate associated with expressing opinions on those writing types. The determination of the potential error rate of the technique is important, such that the client group can choose whether the result-generating system is appropriate to that claimed and has probative value suitable for judicial use.

The results generated by the NZPDES as a group are characterised by low error rates (< 1% overall), which provide significant support to the validity of the

skill that has been claimed by this group. Larger error rates are associated with opinions regarding samples of handwriting text that have been 'simulated'. The errors associated with the two 'simulation' writing types are, however, not shared by all members of the group. Eight of the eleven authorship opinion errors were made on non-peer reviewed trials and it is not unreasonable to expect that errors of this type would be significantly reduced through the normal quality peer-review practices used by this group. In addition, the continued participation in expertise profiling trials, which offer a revision and corrective action component, should maximize the opportunity for

The Skill of a Group of Forensic Document Examiners in Expressing Handwriting - 81

Sample Type	Opinion Scores							
	# correct	# error	# inc.	% correct	% error	% inc.	% correct called	% error called
Genuine Samples	536	0	29	94.9	0	5.1	100	
Simulated not by specimen writer	2356	0	61	97.5	0	2.5	100	0

TABLE 6. The process opinion scores expressed by the group in the trials.

examiner number	Writing Type	Opinion Scores							
		# correct	# error	# inc.	% correct	% error	% inc.	% correct called	% error called
1	SP	94	0	3	96.9	0	3.1	100	0
2	SP	78	0	2	97.5	0	2.5	100	0
3	SP	89	0	8	91.8	0	8.2	100	0
4	SP	92	0	5	94.8	0	5.2	100	0
5	SP	94	0	3	96.9	0	3.1	100	0
6	SP	89	0	8	91.8	0	8.2	100	0
1	SNSP	384	0	21	94.8	0	5.2	100	0
2	SNSP	384	0	8	98.0	0	2.0	100	0
3	SNSP	378	0	27	93.3	0	6.7	100	0
4	SNSP	405	0	0	100	0	0.0	100	0
5	SNSP	403	0	2	99.5	0	0.5	100	0
6	SNSP	402	0	3	99.3	0	0.7	100	0

TABLE 7. The process opinion scores for each examiner for signatures by the specimen writer (SP) and simulations of the specimen writer's signature not by the specimen writer (SNSP).

perceptual and cognitive revision where the system has not produced the correct response.

Although it appears that the judiciary invests strongly in examiner experience to gauge the reliability of opinion, studies conducted at the Forensic Expertise Profiling Laboratory, incorporating the data presented here, have yet to find a simple correlation between experience (that is the number of years that an examiner has been practising forensic handwriting examination), and their correct, error and conservatism scores. Given this reality, it is proposed that the only mechanism by which the judiciary can assess the value of examiner opinion is through examiner results on independent blind trials of the types presented.

9. Utilizing potential or estimated error rates

Because of the number of varied trials undertaken by this group, we consider that the error shown is a

good estimate of the group's potential error rate that can be considered when applying the technique in the casework setting. This error rate can, therefore, be reported as the group's potential error rate. It is important to consider that, although a potential or estimated error rate of < 1% is appropriate to discuss, this rate is associated with examiners making decisions on blind validation trials and then grouping the results. The grouping of results does dilute the data, as the overall data set contains a number of distinct categories of writing, and examiners' relative skill in expressing opinions about these categories does vary between the group and between examiners. A single trial, or even a series of trials, is unlikely to capture all of the variables associated with the routine presentation of forensic casework. Forensic handwriting examination involves an enormous number of tasks prior to a final opinion being expressed. In addition,

questioned and specimen writing can vary with respect to quantity, quality, complexity, skill etc. The error quoted is, therefore, without question an estimate based on the application of the same cognitive skill set to different types of blind trials that is used to examine handwriting and signatures in the casework environment. Since we observe an enormous amount of casework variables, the only approach available to examiners at this time is constant exposure to blind trials that emulate casework as closely as possible.

It is still the case that most examiners internationally have not been exposed to the rigours of testing of the magnitude described in this paper. For courts to take holistic comfort in error scores generated by blind trials, if in fact they take comfort at all, would be a precarious position. To take this position would be to embrace an underlying assumption that the error scores generated by the individuals taking part in the reported trials are representative of error rates over larger groups of document examiners. There is, at this point in time, no clear evidence to support this proposition. It is, therefore, in no way possible to suggest that individuals not covered by this report (that is, outside the New Zealand Police Document Examination group), should be attributed with a similar skill profile and associated error rate.

10. References

- Evetts, I. (1998). Towards a uniform framework for reporting opinions in forensic science casework. *Science & Justice*, 38(3), 198-202.
- Found, B., & Rogers, D. (Eds.). (1999). Documentation of forensic handwriting comparison and identification method: A modular approach. *Journal of Forensic Document Examination*, 12, 1-68.
- Found, B., Sita, J., & Rogers, D. (1999). The development of a program for characterising forensic handwriting examiners' expertise: Signature examination pilot study. *Journal of Forensic Document Examination*, 12, 69-80.
- United States v. Starzeczpyzel, 880 F.Supp.1027 (S.D.N.Y. 1995).
- Huber, R.A., & Headrick, A.M. (1999). *Handwriting Identification: Facts and Fundamentals*, Boca Raton, CRC Press.
- Kam, M., Wetstein, J., & Conn, R. (1994). Proficiency of professional document examiners in writer identification. *Journal of Forensic Sciences*, 39, 5-14.
- Kam, M., Fielding, G., & Conn, R. (1997). Writer identification by professional document examiners. *Journal of Forensic Sciences*, 42, 778-786.
- Risinger, D. M., Denbeaux, M.P., & Saks, M.J., (1989). Exorcism of ignorance as a proxy for rational knowledge: The lessons of handwriting identification "expertise". *University of Pennsylvania Law Review*, 137, 731-792.
- Risinger, D. M. & Saks, M.J., (1996). Science and nonscience in the courts: Daubert meets handwriting identification expertise. *Iowa Law Review*, 82, 21-74.

Acknowledgement

The trials have received financial sponsorship from the National Institute of Forensic Science, Australia and the Senior Managers of Australian and New Zealand Forensic Science Laboratories. We acknowledge the following examiners: David Boot, Allan Herkt, Trish James, Gordon Sharfe, John Walker and Delwynne Walsh, for their participation in this project.