Summer 2005

THE FORENSIC BULLETIN

ISSN 1447-6673

IN THIS ISSUE

A Note from the Editor

Welcome to the summer edition of the Forensic Bulletin, a little belated but well worth the wait. I could explain the delay in the production of this edition of the Bulletin – that like the summer itself, it is the climate's fault but I am unsure that many of our distinguished readers would swallow it, so allow me merely to say that the Bulletin has been competing with other demands!

It has been a very busy time for many Australian and New Zealand forensic practitioners. Not mentioned in this edition is the devastation caused by the tsunami on Boxing Day and the subsequent deployment of many forensic practitioners as part of ongoing assistance programs. The disaster and the forensic science community's response will be covered in depth in the next edition.

One of the big events on the NIFS calendar was our move to the gracious old building featured on the front cover. Our new premises were opened by the Honourable Senator Chris Ellison, Minister for Justice and Customs on March 2. A few photos have been included in the stop press on the back page and details and more photos will be featured in the next edition.

One of the exciting parts (not the deadlines) about being editor is the quality of the articles which are submitted for publication, so please make sure you read them and if you have something you wish to submit – long or short – send it in.

Happy reading (and writing).

Anna Davey
anna.davey@nifs.com.au

SMANZFL

The Australian and New Zealand Forensic Science Society Inc.

NIFS National Institute of Forensic Science Australia

other research currently being undertaken. At the Denver Medical Center we got to see some microchip technology and the MWG robot for liquid handling and thermocyling.

Our next visit was to the California Department of Justice (DOJ) DNA Laboratory in Berkeley. This laboratory is a state lab with new facilities, which serves a population of about 40 million people. They currently have about 95 staff and are broken up into Databanking, Missing Persons, Casework, Methods Development, Training, and CODIS units. They have recently validated a high throughput and cost effective system for databanking using the BODE buccal collector. We also had the opportunity to see the Hamilton and TECAN liquid handling platforms being used. Their training is very rigorous, consisting of a six month intensive training course, and they believe a mentoring system is the most beneficial method of training for introduction into casework.

Our trip to the US was extremely valuable to us both at a personal level as well as organisational. We have made many new contacts from the US Canada, New Zealand and Australia to help us with our future process developments. Please feel free to contact us for any further information you might be interested in.

ERROR RATES IN THE IDENTIFICATION SCIENCES

Dr Stephen Gutowski, Quality/R&D Manager, Crime Scene Division, Victoria Police Forensic Services Department, VPFSC, Forensic Drive, Macleod, Victoria, Australia 3085.

Parts of this paper were presented at the 17th ANZFSS International Symposium on the Forensic Sciences, Wellington 2004.

Introduction

Errors occur in all human activities from computer programming where it is estimated that 20-40 percent of all spreadsheets contain logical errors [1] to the clinical interpretation of laboratory test results where in one study, more than four percent of qualified nurses failed to recognise the symptoms of well developed diabetes [2]. Error in forensic sciences can be particularly embarrassing and costly both in human and monetary terms. Strategies have been and are being widely implemented to help manage this risk, most importantly laboratory accreditation by external bodies such as NATA [3].

There is evidence to suggest that accreditation does work at least with analytical laboratories [4]. Even in that study, however, six percent of accredited laboratories were rated as unsatisfactory (as opposed to 17 percent of unaccredited ones) and, to paraphrase Alexander Pope: To err is human, to forgive is not part of the legal system.

In the field and identification sciences in particular, error has been highlighted in recent years both in the UK by the McKie case [5] and in the USA by various Daubert hearings, culminating in the momentous rulings of Justice Pollak [6] in the case of USA vs. Plaza, Costa and Rodriguez. In the first instance it appears that a refusal of the Scottish Central Records Office to admit a possible fingerprint error nearly led to a serious miscarriage of justice while on the other hand the demonstration that "… there is no evidence to suggest that the error rate of certified FBI fingerprint examiners is unacceptably high" was crucial to Justice Pollak's eventual decision to allow fingerprint matches into evidence as expert testimony.

In the USA the Daubert judgement has been considerably embellished by numerous subsequent rulings, in particular those in the Joiner and Kumho cases [7]. In Australia, those extensions appear to be covered by the expertise and basis rules. Examples of Australasian decisions which employed criteria similar to those of Daubert have arisen, especially in relation to the area of expertise rule [8] and while not directly applicable, the threat of Daubert style challenges to forensic science in Australia is ever-present [9].

In its simple expression, the Daubert standard poses the following questions:

1. Can the technique be tested?
2. Has the technique been subject to peer review?
3. Is the real or potential error rate known?
4. Are there accepted standards and controls?
5. Has the technique been generally accepted in the scientific community?

Each of these areas is worth examination and is being pursued by a variety of interested groups in the field and identification sciences worldwide but the major purpose of this paper is to explore the often delicate area of error rates.

An estimate of actual or potential error rate is crucial to the probative value of all evidence. This is certainly true of the field and identification sciences where hard statistics on the frequency of occurrence of a particular pattern are impossible to come by and individuality is assumed but cannot be proven. A result given by an examiner using a technique and in a laboratory all of which have an extremely low or zero error rate will have tremendous probative value even if the other Daubert criteria are only partially satisfied. The importance of estimation of error rate is beginning to be appreciated even in the more academic areas of forensic science such as DNA profiling [10].

Most if not all forensic laboratories have employed quality assurance measures to help minimise errors for more than 20 years [see eg 11] but the first systematic study of error rates in the peer reviewed literature did not appear – at least to the author's knowledge - until 1995. In that study, by Peterson and Markham, which was published in two parts [12, 13], the authors examined the results of proficiency tests from most areas of the forensic sciences including firearms, toolmarks, fingerprints and footwear, carried out between 1978 and 1991. The reports refer to previous studies by the US LEAA published in 1978 but these studies appear to be about setting up a proficiency program rather than its execution [12].

Peterson and Markham made several general comments on the use of proficiency tests in the determination of error rates [12, 13]. Firstly they admit the possibility of error or contamination in setting up the tests. Secondly they outline the limitations of the tests: actual tools/shoes etc are often not available, the trials are declared and may be treated differently from normal casework and the level of difficulty may not be representative. The authors recommend caution and indeed, CTS, the major suppliers of proficiency tests state on their reports that "…the results compiled in the summary report are not intended to be an overview of the quality of work in the profession and cannot be interpreted as such."[14]. Haber and Haber [15] make similar comments about the need for caution.

What constitutes an incorrect result and how an error rate is then defined are in themselves somewhat moot points. For example, in a recent fingerprint case in the Republic of Ireland, the examiner took a traditional line that there is human error due to incorrect procedures being followed but the error rate in fingerprint examination is zero [16]. It is clear, however, from personal, anecdotal and survey evidence that there is an error rate in field and identification sciences. The situation is well described by Grzybowski *et al* in their seminal 2003 paper [17] in which the authors state "The court is not interested in the 'theoretical error rate', which assumes that everything has been done correctly and the correct answer obtained but it is interested in the real life potential error rate that is reflective of all human endeavours". These authors also say that "To proffer a firearm and toolmark identification as 'infallible' is simply not true and will be met with immediate suspicion" [25].

Grzybowski et al [17] do not count inconclusive results as incorrect results. Incorrect results comprise false positives (inclusions), where an expert asserts that two impressions have the same origin when in fact they arise from different objects, and false negatives (exclusions) where the expert states that two impressions come from different items whereas in fact they have

a common origin. This appears to be a logical approach in that an inconclusive result does not lead to a miscarriage of justice: an inconclusive result is the same as no result- another scene without prints of value, another smudged shoe impression etc. A wrong result is far more damaging, to the wrongly convicted person, the offender acquitted and left to re-offend, and the forensic scientist with the wrong result on their conscience. This paper will count only false inclusions and false exclusions as incorrect results.

While it is clear that all field and identification sciences have an error rate, it is also clear that, by and large, that rate is small. All measures of error rate are less than ideal but some form of estimate is vital. There is some anecdotal evidence from casework in fingerprint examination and validity studies in firearms identification and knife identification. As for the rest, summary reports of proficiency tests remain the major, systematic source of information on error rates within the profession for the period to 2004 and must therefore serve as "best evidence", if only as used by Justice Pollak [6] to show that "… there is no evidence to suggest that the error rate of … is unacceptably high". In addition, Haber and Haber, while urging caution [15], seem to have little problem in using similar data from 1995 to 2001 to criticise fingerprint examination.

A final complication was mentioned by Bruce Budowle in 1999 "An error rate is a wispy thing like smoke, it changes over time…"[18]. Accreditation of laboratories has gained momentum over the last 15 years and training of examiners has become more formal in many jurisdictions, to mention but two factors. This may have reduced error rates although it is also possible that by encouraging a more open and accountable culture it may have increased the reporting rate of errors. It may therefore be of value to compare error rates from the early studies with those being experienced today.

An attempt is made below to collect known data on error rates and supplement it with an analysis of the proficiency testing carried out by the Victoria Police Forensic Services Department over the period 2000-2004. Where applicable, VPFSD results are compared with worldwide results over this period and global figures for error rates calculated for comparison.

Background

1. Error rates to 1991 (from Peterson and Markham 1995 [13])

The analysis carried out by Peterson and Markham on latent print proficiency tests was complex but several conclusions can be drawn. Firstly examiners rated only 98 percent of the manufacturer's prints of value as indeed being prints of value and conversely rated approximately eight percent of the manufacturer's no value prints as having value. Similarly, examiners rated as identifiable only 92 percent of those prints which, according to the manufacturer should have been identifiable. Such results are often called false negatives but in fact are really only false negatives if definitely described as mismatches with all provided exemplars. If a print cannot be matched definitely to the reference prints provided, this may be due, in the minds of some experts, to a lack of sufficient observable matching detail rather than to observed mismatching details which cannot be explained. Such prints should be classed as inconclusive, not exclusions/false negatives. On the other hand experts rated only two percent of the (manufacturer's) unidentifiable prints as identifiable, thus showing a natural conservatism [13].

0.45 percent of matches were ascribed to the wrong person; more were ascribed to the wrong digit of the correct person. These two classes of error are both false positives but may have been typographical or administrative rather than technical errors. The false positive rate was approximately two percent (75/4735), in total, in the studies between 1978 and 1991 [13].

The treatment of toolmark results was also not straightforward. In most tests, results are classified as correct positive, correct negative, false positive if reported as matching when they in fact come from different items, false negative if reported as excluded when they in fact come from the same item, and inconclusive if no firm conclusion is drawn. In three tests, however, a further category of "unjustified exclusion" had to be introduced [13]. In these tests, tools were not provided and marks made by one side or area of the tool were compared with marks made by the other side or different area of the same tool. In the absence of the tool, the different areas could not be checked and the correct response was inconclusive not exclusion.

With footwear and firearms, the results are classified as correct positive, correct negative, false positive, false negative and inconclusive and interpretation is straightforward. Results are tabulated below.

Table 1: External Proficiency Test Results 1978-1991

Test type	# of comparisons	True positive	True negative	False positive	False negative	Inc.
Footwear	1745	484	1033	6 (0.3%)	6 (0.3%)	216 (12%)
Firearms	2106	905	954	12 (0.6%)	17 (0.7%)	218 (10%)
Toolmarks*	1551	538	604	30 (2%)	40 (2.6%)	339 (21.9%)
Fingerprints+	6000~	4658	871	123 (2%)	?	430~ (7-8%)

* *Excluding those tests with unjustified exclusions*
\+ *Treating the difference between the manufacturer's identifiable prints and those rated as identifiable by the examiners as a measure of the number of inconclusive prints. Adapted by Gutowski. Data from Peterson and Markham 1995 [13]*

2. Published Error Rates 1991-2004

Until recently, there has been little to add to Peterson and Markham's study in the field and identification sciences. In document examination, the work of Dr Bryan Found has pioneered an approach to evidential value based on error rate [19] which is beginning to find acceptance [20]. Studies of bitemark and ear analyses have been carried out [21, 22]. In an FBI study in 1999, 50,000 fingerprints were matched by computer simulation with themselves and matches were only obtained with the correct prints [18]. This is support for a low rate of false positives. A better study would have been to run 50,000 prints against 50,000 prints obtained from the same 50,000 fingers but at different times. This would have given more information on false positives, false negatives and inconclusive results.

Recently a validity study has been published for the examination of cartridge cases by FBI examiners. In this study there were 360 comparisons and no false positives or false negatives were obtained [23].

A knife identification project was carried out by the Washington State Patrol Crime Laboratory which found an error rate of approximately 0.8 percent [32].

Firearms and tool marks error rates in 2002 CTS external proficiency tests were discussed at the AFTE annual training seminar in Philadelphia in 2003 [24] (0.6 and 8.3 percent respectively but see also below). Grzybowski et al [17] calculate false positive rates for firearms and toolmarks of

1.0 percent and 1.2 percent respectively for the period 1998-2002.

A very courageous article from the UK 2001 National Fingerprint Conference admits that two errors by the Greater Manchester Police Fingerprint Unit [25] were recently discovered despite the use of ISO procedures and identifications requiring three experts to agree. Further error was discovered with one of the Manchester examiners on further investigation[25].

Results

Results of external and internal proficiency tests are tabulated below. VPFSD participants were all authorised to carry out the tests required: the tests were not used as training exercises. All tests were declared trials. External proficiency tests were treated as tests of the VPFSD Quality System and were therefore technically and administratively reviewed in a similar fashion to casework. An exception was encountered with fingerprints external proficiency test CTS #00-516. Part of the corrective action resulting from the errors in this test was to emphasise the reviews which should have taken place in this test but did not.

Internal proficiency tests were treated as tests of the individual. There was no administrative or technical review of case notes or reports although participants were expected to consult with other experts in the formulation of their opinions in a similar way to casework determinations.

1. Shoe impressions

VPFSD results for the comparison of shoe impressions are given in Table 2.

General error rates are shown in Table 3.

Table 2: VPFSD proficiency test results – shoe impressions

Test	No. Examiners	No. Comparisons	Correct Pos	Correct Neg	False Pos*	False Neg+	Inc#
CTS 00-533	6	42	12	27	0	0	3
CTS 01-533	1	16	5	11	0	0	0
Internal 2000/01	7	26	22	3	0	0	1
Internal 20002/03	7	42	5	35	0	0	2
CTS 03-533	5	25	6	11	0	0	8
Total	-	151	50	87	0	0	14

* *Non-match reported as match* + *Match reported as non-match* # *Match or non-match reported as inconclusive including match reported as class match*

Table 3: Overall error rate in CTS proficiency tests- shoe impressions

Test	No. Examiners	No. Comparisons	Correct Pos	Correct Neg	False Pos	FalseNeg	Inc#
CTS 00-533	219	1533	438	1000	2	0	93
CTS 01-533	217	3468*	1081	2022	3	0	362
CTS 02-533	232	2552	682	1707	1	2+	160
CTS 03-533	231	1153	445	602	6	1	99
Total		8706	2646	5331	12	3	714

* *Two shoes for comparison purposes – left and right* + *Error in transcription in final report leading to at least one of these false negatives*

The VPFSD error rate (0/151) can be contrasted with the general error rate in the CTS proficiency tests nos 00, 01, 02 and 03-533. This rate was 15/8706, approximately 1/600 or 0.17 percent. It comprised a false positive rate of approximately 0.14 percent and a false negative rate of around 0.03 percent.

The VPFSD rate for inconclusive results (9.3 percent) was slightly higher than the average in the profession (8.2percent). Given the small number of results involved, this return is not likely to be of significance.

2. Vehicle number restoration

Only internal VPFSD tests are available: a proposed CTS research test in 2002 did not eventuate. Reports of tests were therefore not reviewed before submission. Results are tabulated below (Table 4).

The error rate is therefore approximately 1/105 or approximately one percent. No comparative figures from other organisations are currently available, to the author's knowledge. A proposal by CTS to introduce a number restoration proficiency test in 2003 did not eventuate.

Table 4: VPFSD Impressed Number Restoration - proficiency test results

Test	No. Examiners	No. Comparisons	Correct Pos	Correct Neg	False Pos	False Neg	Not restored
1-4/001	4	32	11	NA	1+	NA	20
1-4/012	4	24	1	NA	0	NA	23
1-4/023	4	24	23	NA	0	NA	1
1-4/034	5	25	20	NA	0	NA	5
Total	-	105	55	NA	1	NA	49

+ *Error in transcription from case notes to final report leading to apparent false positive. Internal test, therefore by design not reviewed.*

3. Fingerprint identification

VPFSD participation in CTS external proficiency tests between 2000 and 2004 is outlined below (Table 5):

Table 5: VPFSD Fingerprint Proficiency Test Results

Test	No. Examiners	No. Comparisons	Correct Pos	Correct Neg	False Pos	False Neg	Inc#
CTS 00-516	9	90	88	0	2*	0	0
CTS 01-516	12	132	115	12	0	0	5
CTS 01-517	1	11	11	0	0	0	0
CTS 02-517	12	120	107	12	0	0	1
CTS 03-516	12	120	108	12	0	0	0
CTS 03-517	1	9	6	2	0	0	1
CTS 04-516	1	12	9	3	0	0	0
Total	-	494	444	41	2	0	7

** Errors in transcription from job cards to final report leading to false positive results.*

Note: In 2000, external proficiency tests were not technically or administratively reviewed. This has now changed to reflect normal casework procedures.

The VPFSD error rate for false positives in external proficiency tests was therefore 2/494 or approximately 0.4 percent. The rate for inconclusive results was 7/494 or approximately 1.4 percent. These results can be compared with overall responses to CTS testing over the same period (Table 6):

The overall error rate for false positives shown in Table 6 below is 68/20873 or approximately 0.33 percent. Inconclusive results comprise 295/20873 or approximately 1.4 percent.

Table 6: Overall CTS Fingerprint Proficiency Test Results

Test	No. Examiners	No. Comparisons	Correct Pos	Correct Neg	False Pos	False Neg	Inc#
CTS 00-516	278	2780	2745	0	13	NA	22
CTS 01-516	296	3256	2873	296	10	NA	77
CTS 01-517	120	1320	1290	0	2	NA	28
CTS 02-516	303	3333	2922	303	15	NA	93
CTS 02-517	146	1460	1153	292	7	NA	8
CTS 03-516	336	3360	2985	336	5	NA	34
CTS 03-517	188	1692	1481	376	1	NA	22
CTS 04-516	306	3672	2733	914	15	NA	11
Total	-	20873	18182	2517	68	-	295

NA- Not applicable

4. Firearms identification

The VPFSD participation in internal and external proficiency tests is outlined below (Table 7) and this is compared with the overall rate of error in external proficiency tests internationally together with the FBI validation study [23] (see Table 8).

No errors (0/38) were made by VPFSD firearms examiners: there were no false positives nor were there any false negatives. The FBI validation study also revealed a zero error rate. This zero error rate contrasts with a small but real error rate in external proficiency tests internationally where there were 18/4113 false positives (0.44 percent) though only one false negative (0.02 percent). Inconclusive results comprised approximately 600/4113 or 14.6 percent.

Table 7: VPFSD proficiency test results- Firearms

Test	No. Examiners	No. Comparisons	Correct Pos	Correct Neg	False Pos	False Neg	Inc#
CTS 00-526	1	3	2	1	0	0	0
CTS 01-526	1	3	1	2	0	0	0
1,2-001	2	6	2	4	0	0	0
1,2,3-002	3	9	3	6	0	0	0
CTS 02-526	1	3	2	1	0	0	0
1,2,-023	2	8	7	1	0	0	0
CTS 03-526	1	3	2	1	0	0	0
CTS 03-527	1	3	2	1	0	0	0
Total	-	38	21	17	0	0	0

Table 8: Overall CTS proficiency test results – firearms

Test	No. Examiners	No. Comparisons	Correct Pos	Correct Neg	False Pos	False Neg	Inc#
CTS 00-526	230	690	460	230	0	0	0
CTS 01-526	235	705	228	261	12	0	204
CTS 01-527	80	240	160	63	1	0	16
CTS 02-526	249	747	498	191	0	0	58
CTS 02-527	95	285	186	49	1	1	48
CTS 03-526	246	738	492	221	0	0	25
FBI Study [23]	8	360	70	118	0	0	172
CTS 03-527	116	348	215	52	4	0	77
Total	-	4113	2309	1185	18	1	600

5. Toolmarks

The VPFSD's error rate for toolmark examination proficiency tests over the past three years is 0/23 and the rate of inconclusive results was 3/23 or 13 percent.

The overall error rate in the CTS toolmark proficiency tests listed below was 51/2816 or 1.8 percent being made up of 35/2816 or 1.2 percent false positives and 16/2816 or 0.6 percent false negatives. The rate of inconclusive results in these tests was 75/2816 or approximately 2.7 percent.

Table 9: VPFSD proficiency test results – toolmarks

Test	No. Examiners	No. Comparisons	Correct Pos	Correct Neg	False Pos	False Neg	Inc#
CTS 00-528	1	3	1	1	0	0	1
1-3/001	3	7	3	3	0	0	1
CTS 02-528	1	3	1	2	0	0	0
1-2/023	2	6	1	4	0	0	1
CTS 03-528	1	4	1	3	0	0	0
Total	-	23	7	13	0	0	3

Table 10: Overall CTS proficiency test results – toolmarks

Test	No. Examiners	No. Comparisons	Correct Pos	Correct Neg	False Pos	False Neg	Inc#
CTS 00-528	198	594	330	192	6	10	56
CTS 01-528	208	624	197	404	12	2	9
CTS 02-528	214	640	209	423	3	1	4
CTS 02-529	48	142	42	86	9	3	2
CTS 03-528	204	816	200	607	5	0	4
Total	-	2816	978	1712	35	16	75

6. Synopsis of VPFSD and CTS error rates July 2000- June 2004

Table 11 shows a synopsis of error rates from proficiency testing by the VPFSD over the period July 2000 – June 2004 and compares this with error rates in the same or similar tests worldwide.

Table 11:

Test type	VPFSD/ CTS participants	Casework/ Proficiency Test	No. of Comparisons	False Pos	False Neg	Total Error (%)	Inc#
Footwear	VPFSD	Proficiency Test	151	0	0	0	14 (9.3%)
Footwear	CTS	Proficiency Test	8706	12 (0.14%)	3 (0.03%)	0.2	714 (8.2%)
Number restoration	VPFSD	Proficiency Test	80	1 (1.25%)	0	1.25	44 (55%)
Fingerprints	VPFSD	Proficiency Test	473	2 (0.4%)	0	0.4	7 (1.5%)
Fingerprints	CTS	Proficiency Test	15509	53 (0.3%)	-	0.3	284 (1.7%)
Firearms	VPFSD	Proficiency Test	35	0	0	0	0
Firearms	CTS	Proficiency Test	3765	14 (0.4%)	1 (0.03%)	0.4	523 (13.9%)
Toolmarks	VPFSD	Proficiency Test	23	0	0	0	3 (13%)
Toolmarks	CTS	Proficiency Test	2816	35 (1.2%)	16 (0.6%)	1.8	75 (2.7%)

7. Comparison of error rates 1978-91 and 2000-2003

Table 12 shows the error rates obtained from the final reports of CTS proficiency tests carried out by all participants as published in the study of Peterson and Markham [12, 13] for the period 1978-1991 and compared with the results of similar tests circulated by CTS for the period 2000-2004.

Table 12:

Test type		# comparisons	False positives	False negatives	Inconclusive
Footwear	1978-91	1745	6 (0.3%)	6 (0.3%)	216 (12.4%)
	2000-04	8706	12 (0.14%)	3 (0.03%)	714 (8.2%)
Firearms	1978-91	2106	12 (0.6%)	17 (0.7%)	218 (10.4%)
	2000-04	4113	18 (0.4%)	1 (0.02%)	600 (14.6%)
Toolmarks	1978-91	1551	30 (1.9%)	40 (2.6%)	339 (21.9%)
	2000-04	2816	35 (1.2%)	16 (0.6%)	75 (2.7%)
Fingerprints	1978-91	6000~	123 (~2%)	-	430~ (~7%)
	2000-04	20873	68 (0.3%)		295 (1.4%)

Discussion

Errors occur, that much is certain. A number of errors was detected in this report; many may remain. Even CTS itself is not immune from typographical error [see eg ref. 26, p3, para 1]. What is not certain is the rate of error.

Discovered errors in casework are, due to good practice, extremely rare and are a minimum estimate of error rate. A better estimation of error rate in casework would be most rigorously achieved by the re-examination of several thousand cases where each case was examined by a panel of experts to achieve consensus. In the absence of a massive increase in funding, this is unlikely to happen. Estimates of error rates in casework might be able to be made in the future from systems of near-miss reporting which encourages practitioners to report errors and near misses, which learns from mistakes, and rewards staff when potentially serious mistakes are avoided. Such systems are in place in some medical facilities [27] and could be encouraged by the processes of internal audits but in many institutions this would require a change in culture. What are available now are point estimates resulting from diagnostic actions and the results of proficiency tests.

The arguments against using proficiency tests as estimates of error rate are well outlined by Peterson and Markham [12] and most of those arguments still stand. In particular, checking procedures used in casework may not be followed in proficiency tests. In addition, not all CTS tests over the period July 2000-June 2004 are tabulated here. Results are limited to those in which the VPFSD participated and those where the final CTS reports are available on the internet. Nevertheless, the test results remain of value as general indicators of error rates [6, 17] and one argument against them does not apply to the VPFSD results: all these results were submitted by qualified examiners, they were not used as training exercises.

A comparison of false positive rates from 1978-91 and 2000-2004 (Table 12) shows that these rates have declined in all classes of test. Due to the small numbers of such results in some classes the results in those classes may not be statistically significant. In other classes such as fingerprints though, the trend is clear and extremely heartening. As Grzybowski et al comment [17] this may reflect the increased number of accredited laboratories or the greater importance now being given to proficiency tests such that the error rate is now approaching the "real" error rate in casework or both. VPFSD results support this idea. Three errors were detected in the VPFSD proficiency results. All three errors were transcription errors which would have been expected to have been picked up in the usual technical and administrative reviews of the completed case. Such reviews are mandatory in accredited facilities.

The drop in the percentage of false positive results is accompanied, in some classes of test, by

an even greater drop in false negatives This may also be due to improved procedures but may also be due to changes in reporting – results given in the past as negative may now be being more correctly reported as inconclusive. There is also a general decline in the number of inconclusive results. This is particularly marked for toolmark examination and may reflect changed practices by the test supplier: CTS now generally provides the tool along with the questioned marks.

A comparison of VPFSD and CTS results shows that, in general, the VPFSD error rates in proficiency tests are in line with or somewhat better than results achieved by the profession.

Error rates in proficiency tests in the identification sciences can be compared with rates in other areas of forensic science. Using similar methodology to that used above, a false positive error rate of around two percent can be calculated for one handwriting test concerning simulated writing [28], a false negative rate of around one percent was encountered in fibre analysis [29] and an error rate of approximately 1.2 percent for architectural paint analysis [30, only includes those laboratories with a full range of techniques available]. Even areas where testing is comparatively well automated, such as DNA analysis, display measurable error rates in proficiency testing. For example, while no false assignments were made in CTS Test 03-571 [31], mismatches were obtained in eight out of approximately 24,000 DNA results. Of these mismatches, six were of a type known as allele dropout which can conceptually be considered as similar to differences in fingerprints due to distortion of the finger or differences in shoe impressions due to subsequent wear. Two DNA typings, however, were in error and in different circumstances could have led to false negative results (the source of this DNA was not meant to match either of the known sources so an incorrect mismatch did not alter the conclusion).

This single proficiency test for DNA typing therefore shows two laboratories in error and an error rate of around 2/1700 or approximately 0.12 percent. These observations are in no way meant to be a comprehensive analysis of errors in these fields but they do point out that whatever the area of forensic science, errors will occur. Accreditation or at least greater attention to quality assurance does appear to have helped to reduce error rates over time. It is expected that standardisation and automation will reduce the rate further in the future. Forensic scientists and technicians should continue with their improvement programs to provide the best service available for the budget provided.

References

1. "Spreadsheet development in laboratories". Wilson J, NATA News 99, March 2001, 4-7
2. "Post analytical 'error' rates in point of care testing: use of a quality assurance programme", 2001, Accreditation and Quality Assurance, 6, 402-404.
3. "Accreditation for Forensic Science". Ross A, N tification and classification of physical evidence". Peterson J and Markham P, Journal of Forensic Sciences, 1995, 40/6, 994-1008.
4. "Laboratory accreditation: Proof of performance". Tighe P, NATA News, March 2002, 8.
5. "There's nane ever fear'd that the truth should be heard save they whom the truth would indite". McKie I, Science and Justice, 2003, 43/3, 161-165.
6. Order in the US District Court for the Eastern District of Pennsylvania, Cr. No. 98 362-10, 11, 12. Pollak L, March 13th 2002.
7. "The 'what' and the 'how' of Daubert challenges to expert testimony under the new Federal Rule of Evidence 702". Gebaur M, at www.productdefenders.com /articles/meg0501.html.
8. "The Area of Expertise Rule". Freckleton I in "Expert Evidence", Freckleton I and Selby H, Eds., LBC Sydney, 1993.
9. "'Daubert' enhancements and the Australian reaction". Thatcher P, NIFS Bulletin, December 2002, 23-27.
10. "How the probability of a false positive affects the value of DNA evidence". Thompson W, Taroni F and Aitken C, Journal of Forensic Sciences, 2003, 48/1, 47-54
11. "Quality assurance in forensic science". Pereira M, Forensic Science International, 1985, 28, 1-6
12. "Crime laboratory proficiency test results 1978-1991 I: Identification and classification of physical evidence". Peterson J and Markham P, Journal of Forensic Sciences, 1995, 40/6, 994-1008.
13. "Crime laboratory proficiency test results 1978-1991 II: Resolving questions of common origin". Peterson J and Markham P, Journal of Forensic Sciences, 1995, 40/6, 1009-1029.
14. Toolmarks Examination Test No. 03-528 Summary Report, CTS 2003.
15. "Error rates for human latent fingerprint examiners". Haber L and Haber R, submitted to Advances in Automatic Fingerprint Recognition, Springer-Verlag, New York, 2003.
16. http://breaking.tcm.ie/2003/02/26story89734.html
17. "Firearm/toolmark identification: Passing the reliability test under Federal and State evidentiary standards". Grzybowski R, Miller J, Moran B, Murdock J, Nichols R and Thompson R, AFTE Journal, 2003, 35/2, 209-241.
18. Government's Post-Daubert Hearing Memorandum re USA vs B Mitchell, Criminal No 96-00407 in the District Court of Eastern Pennsylvania, Stiles M and Sarmousakis P, 1999
19. "The skill of a group of forensic document examiners in expressing handwriting and signature authorship and production process opinions". Found B, Rogers D and Herkt A, Journal of Forensic Document Examination, 2001, 14, 15-30.
20. "Characterising the science of forensic document handwriting and signature examinations". Part 1 of the Fundamental Principles of Handwriting Identification Seminar at the 3rd European Academy of Forensic Scien ces Meeting, Istanbul, September 22nd 2003.
21. "Digital bite mark overlays- an analysis of effectiveness". Pretty I and Sweet D (sic), Journal of Forensic Sciences, 2001, 46/6, 1385-1391.
22. "Ear identification based on surveillance camera images". Hoogstrate A, Van den Heuvel H and Huyben E, Science and Justice, 2001, 41/3, 167-172.
23. "A comprehensive validity study for the forensic examination of cartridge cases". Bunch S and Murphy P, AFTE Journal, 2003, 35/2, 201-203.
24. "Proficiency test overview". Stanley G, International Bulletin for Shoeprint/ Toolmark Examiners, 2003, 9/2, 29 (Abstract only).
25. "An update on some recent Manchester cases "Presentation Notes"". Ritchie A, Fingerprint Whorld, 2002, 28/107, 11-17.
26. CTS Firearms Examination Test No. 03-526 Summary Report, CTS 2003.
27. "Medical errors deliver valuable lesson", T Noble, The Age, Tuesday October 21st 2003, News 3.
28. CTS Handwriting Examination Test No. 01-524 Summary Report, CTS 2001.
29. CTS Fiber Analysis Test No. 02-539 Summary Report, CTS2002.
30. CTS Paint Analysis Test No. 03-545 Summary Report, CTS 2003.
31. CTS Forensic Biology Test No. 03-571 Summary Report, CTS 2003.
32. "Knife identification project (KIP)", E Thompson and R Wyant, Association of Firearm and Tool Mark Examiners Journal, 2003, 35/4, 366-370.