To err was forbidden: the changing culture of error exploration in forensic pattern evidence

Bryan Found PhD

Chief Forensic Scientist
Office of the Chief Forensic Scientist
Forensic Services Department

Adjunct Professor
Faculty of Law
University of New South Wales
You Want to Look Smart at Exposition
The pattern evidence experiential positive feedback model

Image features

Class characteristics

Exclusion

Individual characteristics

Identification

+ +
70’s
Critics
LOVE IS IN THE AIR?

FALSE.

NITROGEN, OXYGEN, ARGON AND CARBON DIOXIDE ARE IN THE AIR.


• Validation testing
• Method development
• Reporting procedure
• Error rate estimates
• Proficiency testing
• Context and bias
The Questioned Document Culture

- 5 year on the job training
- Mentored casework
- Context information
- Identification paradigm
Early culture

• Courts for acceptance

• Case outcomes as validation

• Error forbidden
Forbidden error

- 100 years of acceptance as expert evidence internationally
- Defining textbooks
- Hierarchy based on experience
- Professional societies
- Existed within government multidisciplinary laboratories
- Documented approaches, peer review
- Acceptance
Pattern evidence
- Validity & reliability
- Underpinning theory
- Reporting practices
- Error rate estimates
- Proficiency testing
I highly recommend graduate school as a way to avoid the real world for 2 to 5 additional years.
• Neuroscience of handwriting movements and application to forensic examination of handwriting

• Software tools for handwriting measurement

• Complexity theory to replace class/individual characteristics belief

• Signature complexity statistical modelling

• Method development

• Expertise testing and characterisation
Perception, cognition and blind testing

- State the beliefs of the practitioners
- Test the beliefs of practitioners where the ground truth is known
- Define the character of the evidence in terms for it potential to be misleading
• “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)

• “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” (United States v Hines, 1999)

• “A great many tests... would be necessary to know what, if anything, (examiners) can do accurately, and under what conditions” (United States v Hines, 1999)

• “A complete testing regime would have tests which covered the entire spectrum of conditions and difficulties” (United States v Hines, 1999).
‘...science can examine the dependability of the results of such a process (handwriting identification) even when the process is not a science’.

Risinger & Sacs (1996)
Science and Non-science in Courts
Iowa Law Review
The 90’s years
• The Special Advisory Group (Document Examination)

• All individual Government document examiners who were prepared to offer the soft bit of their professional tummies to enormous independent scrutiny for the good of the Justice system
A little background
Although not stated, practitioners must be considering competing propositions

- Normal writing by the specimen writer
- Normal writing not by the specimen writer
- Simulation of the specimen writing not by the specimen writer (forgery)
- Disguised writing not by the specimen writer
- Disguised writing by the specimen writer
Signature propositions

- Genuine by the specimen writer
- Simulated by someone other than the specimen writer
- Disguised writing by the specimen writer
Case emulation trial structure

Population of exemplar handwritten images

Opinions regarding authorship & process of production

Population of random single questioned handwritten images containing all possibilities
Dear Joyce,

You are a rip off bastard. I will not come over there to fill my car as that cock-eyed person is ripping everyone else off and don’t say anything to him because you must be good lovers. When I see your face it looks like my white bum!

Anonymous

Dear Joyce,

You are a rip off bastard. I will not come over there to fill my car as that cock-eyed person is ripping everyone else off and don’t say anything to him because you must be good lovers. When I see your face it looks like my white bum!

Anonymous

Specimen handwriting

Opinion

Opinion type and strength for each questioned writing type
Limitations

- Non original
- Traditional K to Q structure
- Line up
- Declared
- Fatigue
- Cognitive confusion
- External validity issues
What was expected?

• Low error distributed amongst the authorship opinions

• Similar profiles of responses from qualified FDEs
What appeared
Inter-examiner differences

![Bar chart showing the number of misleading opinions for each examiner. The x-axis represents the examiner numbers from 1 to 17, and the y-axis represents the number of misleading opinions ranging from 0 to 30. Examiner 17 has the highest number of misleading opinions, while examiners 1 to 6 have significantly fewer.]
Results for examiners

Number of errors

Years experience
Results for examiners

Years since qualified vs. Number of errors
Ramping up
2000 – 2008 Blind trials

- 45850 OU 2000 -2008 signatures
- 32050 OU 2001-2005 handwritten text

77900 blind opinions by practicing FDEs
The things about investigations of error that I reflect on
Studying FDE behavior

Similar problems as with normal laboratory animals...........

68% comply either because they think it is a good idea or they are told to

22% are annoyed about it and don’t want to participate but are controlled by the 68%

5% are angry and just won’t participate in spite of whether they consider the experiment a good idea or not

5% attack the experimenters
Expectations around experience

Skill Development

- Youngling (Beginner)
- Apprentice (Intermediate)
- Jedi (Advanced)
- Master Jedi (Expert)
- Yoda (900 years)

- 10 hours
- 100 hours
- 1,000 hours
- 10,000 hours
- 100,000 hours

adobegeek.ca
The relationship between experience and specialists’ behavior

N = 19
The character of expertise
The data


Comparison FDE to lay people

Examiners
Controls

Mean number of opinion units

- correct
- misleading
- inconclusive

*
All studies to date have shown that the ‘expertise’ is real and demonstrable.
Did we find a global error rate?
The easy answer: Signatures
(45850 OU 2000 - 2008)
The easy answer: Handwritten text
(32050 OU 2001-2005)
The more sophisticated answer
The more sophisticated answer

Handwriting

% Misleading called
% Correct called
% Inconclusive
% Misleading
% Correct

2001 2002 2003 2004 2005
The much more sensible more sophisticated answer

<table>
<thead>
<tr>
<th>Questioned signature type</th>
<th>Genuine</th>
<th>Disguise</th>
<th>Simulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Score</td>
<td>% Correct</td>
<td>% Misleading</td>
<td>% Inc.</td>
</tr>
<tr>
<td>% Score</td>
<td>74.5</td>
<td>21.5</td>
<td>3.9</td>
</tr>
</tbody>
</table>
The much more sensible more sophisticated answer

<table>
<thead>
<tr>
<th>Questioned handwriting type</th>
<th>% Score</th>
<th>% Correct</th>
<th>% Misleading</th>
<th>% Inconclusive</th>
<th>% Correct called</th>
<th>% Misleading called</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulations (N)</td>
<td></td>
<td>41.7</td>
<td>28.3</td>
<td>10.6</td>
<td>0.1</td>
<td>0.3</td>
</tr>
<tr>
<td>Disguise (N)</td>
<td></td>
<td>42.2</td>
<td>57.7</td>
<td>0.3</td>
<td>0.6</td>
<td>0.8</td>
</tr>
<tr>
<td>Normal (N)</td>
<td></td>
<td>77.4</td>
<td>22.0</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Normal (Sp)</td>
<td></td>
<td>99.7</td>
<td>0.3</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Disguise (Sp)</td>
<td></td>
<td>99.2</td>
<td>96.6</td>
<td>99.6</td>
<td>62.0</td>
<td>6.0</td>
</tr>
</tbody>
</table>
How do blind trials help?

- Revision & corrective action package for each trial
- Group and individual analysis
- Problem samples
AN EXPERT IS A MAN WHO HAS MADE ALL THE MISTAKES WHICH CAN BE MADE, IN A NARROW FIELD.

Niels Bohr
"I thought I felt a paradigm shift, but it was just my undershorts riding up."
Validity & reliability of a human skill
Pattern evidence practitioners should be treated as ‘instruments’

All efforts should be made to calibrate and monitor these instruments, using appropriate testing sets prior to and during their use in forensic investigations
Like before, through the application of the ‘principles processes of science’, we provided practitioners with the opportunity to make mistakes, and provided real opportunity for skill development.

Skill profile across the international community is still largely unknown.
‘Look at the largest source of error first. These are likely to be human processes’

Is there a global error (misleading) rate?
Error mitigation
In the interim
Early microscope