Implementing Blind Proficiency Testing in Forensic Laboratories: Motivation, Obstacles, and Recommendations

2021 Field Update

June 14, 2021

forensicstats.org
Why blind proficiency testing?

- 2009: National Academy of Sciences recommended that forensic proficiency testing programs include blind tests.
- 2016: National Commission on Forensic Science recommendation to the Attorney General recommended that all DOJ FSPs “seek proficiency testing programs that provide sufficiently rigorous samples that are representative of the challenges of forensic casework.”
- 2016: President’s Council of Advisors for Science and Technology wrote “test-blind proficiency testing of forensic examiners should be vigorously pursued, with the expectation that it should be in wide use, at least in large laboratories, within the next five years.”
Two 1970s studies in drug testing labs found that false negatives were higher in blind test compared to declared tests (La Motte et al, 1977).

A 2001 study comparing blind and declared proficiency tests in blood lead testing programs at two large state laboratories found error rates were higher in the blind tests and suggested that laboratories were making special efforts when analyzing known proficiency test samples (Parsons et al, 2001).

Today, Mandatory Guidelines for Federal Workplace Drug Testing Programs require participating laboratories to conduct blind testing.
Examination Result

Conforming
- Correct
  - Type I Error
  - Type II Error
Non-conforming
- Mistake
  - Malpractice
- Misconduct

Adapted from Mejia et al, FSI-Synergy, 2020
Examination Result

Conforming
- Correct
- Type I Error
- Type II Error

Non-conforming
- Mistake
- Malpractice
- Misconduct

Only blind testing can catch

Adapted from Mejia et al, FSI-Synergy, 2020
Which forensic labs use blind proficiency tests?

Burch et al, 2016
Meeting on blind proficiency testing

- October 2018, hosted at Allegheny County Office of the Medical Examiner (houses county crime laboratory)
- Participants
  - QA staff and lab directors from 7 forensic laboratory systems
  - Representatives of AFQAM
  - Faculty, graduate students, and a post doc from 2 universities representing fields ranging from psychology to statistics.
- Two days of presentations and discussions gauging interest in and obstacles to implementing blind proficiency tests at state and local laboratories

Mejia et al, FSI-Synergy, 2020
Developing blind proficiency tests

• Realistic tests cases can be complex to create
• Development of realistic submission materials is difficult
• Cost may be prohibitive

Images: Allegheny County Office of the Medical Examiner, South Dakota DC Forensic Laboratory websites
Submitting and managing blind proficiency tests

- Test must be submitted to the lab by an outside LEA
- Not all LIMS are equipped to easily flag and track test cases
- Labs must ensure results are not released as real cases
• Proficiency tests could impact metrics, so labs need to decide how/whether to include them
• Blind testing challenges the cultural myth of 100% accuracy
Forensic experts testify on day three of Hernandez murder trial

Audrey Mayer  Mar 17, 2021 Updated Mar 22, 2021
Proficiency matters to finders of fact

- Online survey of potential jurors (matched to demographics of US)
- Read mock transcript from burglary case with one piece of evidence – forensic bitemark analysis or latent print analysis
- Same evidence each time, varied the proficiency of the examiner.

<table>
<thead>
<tr>
<th></th>
<th>Unknown</th>
<th>Low</th>
<th>High</th>
<th>High-No Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guilty (%)</td>
<td>75.2</td>
<td>65.1</td>
<td>77.7</td>
<td>79.2</td>
</tr>
<tr>
<td>Likely committed (1-100)</td>
<td>79.36</td>
<td>75.23</td>
<td>80.97</td>
<td>81.44</td>
</tr>
<tr>
<td>Evidence persuasive (1-7)</td>
<td>6.16</td>
<td>5.89</td>
<td>6.25</td>
<td>6.3</td>
</tr>
</tbody>
</table>

Adapted from Crozier et al, Forensic Science Intl, 2020
## Stakeholders

<table>
<thead>
<tr>
<th>Laboratories</th>
<th>Clients</th>
<th>Professional Association</th>
<th>External Organizations</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Examiners</td>
<td>• Law Enforcement</td>
<td>• AFQAM</td>
<td>• Accreditation Bodies</td>
</tr>
<tr>
<td>• Quality Assurance Staff</td>
<td>• Prosecutors</td>
<td>• ASCLD</td>
<td>• Proficiency Test Providers</td>
</tr>
<tr>
<td>• Laboratory Management</td>
<td>• Defense Attorneys</td>
<td>• OSAC</td>
<td>• Researchers at CSAFE, RTI, and other academic institutions</td>
</tr>
<tr>
<td></td>
<td>• Judges</td>
<td>• AAFS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Jurors</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Houston Forensic Science Center

• Started blind quality control program in 2015
• September 2015-Dec 31, 2018
  • 978 blind QC samples submitted
  • 901 fully analyzed
  • 51 discovered as blinds by analysts
• Benefits
  • Tests full pipeline
  • More realistic gauge of effectiveness
  • Enables multi-disciplinary tests

Hundl et al, Journal of Forensic Sciences, 2019
## Houston Forensic Science Center

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Blind QCs submitted per month*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toxicology</td>
<td>14</td>
</tr>
<tr>
<td>Firearms – blind verification</td>
<td>1</td>
</tr>
<tr>
<td>Firearms – blind QC</td>
<td>1</td>
</tr>
<tr>
<td>Seized drugs</td>
<td>30</td>
</tr>
<tr>
<td>Forensic biology</td>
<td>10</td>
</tr>
<tr>
<td>Latent print processing</td>
<td>3</td>
</tr>
<tr>
<td>Latent print comparison</td>
<td>10</td>
</tr>
<tr>
<td>Multimedia – digital forensics</td>
<td>1</td>
</tr>
<tr>
<td>Multimedia – audio/visual</td>
<td>1</td>
</tr>
</tbody>
</table>

* Approximately 5% of casework completed during 2017

Hundl et al, Journal of Forensic Sciences, 2019
## HFSC – proficiency testing costs in USD

<table>
<thead>
<tr>
<th></th>
<th>Blind 2017</th>
<th>Blind 2018</th>
<th>External/Declared 2017</th>
<th>External/Declared 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toxicology</td>
<td>16,716</td>
<td>28,901</td>
<td>1,720</td>
<td>1,765</td>
</tr>
<tr>
<td>Firearms</td>
<td>0</td>
<td>0</td>
<td>2,300</td>
<td>2,245</td>
</tr>
<tr>
<td>Seized drugs</td>
<td>5,300</td>
<td>165</td>
<td>3,200</td>
<td>3,060</td>
</tr>
<tr>
<td>Forensic biology</td>
<td>1,840</td>
<td>0</td>
<td>8,606</td>
<td>8,262</td>
</tr>
<tr>
<td>Latent prints</td>
<td>0</td>
<td>20</td>
<td>6,130</td>
<td>6,060</td>
</tr>
<tr>
<td>Digital forensics</td>
<td>0</td>
<td>378</td>
<td>2,786</td>
<td>2,490</td>
</tr>
<tr>
<td>Audio/visual</td>
<td>NA</td>
<td>221</td>
<td>4,550</td>
<td>4,125</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>1210</td>
<td>334</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Hundl et al, Journal of Forensic Sciences, 2019
HFSC blind QC for latent print comparison

- Over 2.5 years, 376 latent prints submitted as part of 144 cases
- Used LQ metrics to classify prints
- 92% of prints submitted of sufficient quality to enter into AFIS
- Of those, for prints with a source present in AFIS, 47% of print searches generated candidate list with source present

Gardner et al, Forensic Science International, 2021
Results

<table>
<thead>
<tr>
<th>Examiner conclusions</th>
<th>True association</th>
<th>Not associated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preliminary association</td>
<td>123 (correct association)</td>
<td>0 (false positive)</td>
</tr>
<tr>
<td>Not associated</td>
<td>154 (false negative)</td>
<td>68 (correct exclusion)</td>
</tr>
<tr>
<td>No AFIS search</td>
<td>24 (potential false inconclusive)</td>
<td>6 (potential false inconclusive)</td>
</tr>
</tbody>
</table>

- **Recall:** tests the pipeline – not just examiner comparisons
- **Additional analyses in paper**

Gardner et al, Forensic Science International, 2021
What next?

• Results of survey of latent print examiners on blind proficiency testing
• Survey of laboratory managers on blind proficiency testing
• Compile a directory of laboratories currently doing or planning blind proficiency testing
• Resume discussions of collaborations over test materials, studies (meetings/presentations)
• Engage with proficiency test providers about materials
Questions?

• In order of appearance:
  • Robin Mejia: rmejia@andrew.cmu.edu
  • Brandon Garrett: bgarrett@law.duke.edu
  • Callan Hundl and Maddisen Newman at HFSC, quality@houstonforensicscience.org
  • Brett Gardner: BG2DD@Virginia.edu
  • Or anyone else at CSAFE