

# Defense Forensic Science Center

## DNA Mixture Interpretation Study: *DNA Examiner Assessment (DEAT) Tool*



R Aranda IV, Emily Rogers, K. Mereus, P. Spencer, and R. Tontarski  
International Symposium on Forensic Science Error Management

Office of the Chief Scientist

23 July 2015



# Disclaimer

---

The opinions or assertions contained herein are the private views of the author and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

Names of commercial manufacturers or products included are incidental only, and inclusion does not imply endorsement by the authors, DFSC, OPMG, DA or DoD.

Unless otherwise noted, all figures, diagrams, media, and other materials used in this presentation are created by the respective author(s) and contributor(s) of the presentation and research.



# Briefing Overview

---

- DFSC mixture study structure
  - Background, goals and composition
- The DNA Examiner Assessment Tool
  - In-house Metrics
  - Data Visualization
- Results and Applications
  - Uncovering interpretation issues
  - Providing examiner and lab training benchmarks
  - Measuring changes in SOP
  - Revealing accuracy limitations depending on mixture complexity



# DFSC Mixture Study



- Purpose:
  - To assess the inter- and intra-laboratory variation in DNA examiners' generated genotype interpretations
  - To better understand the current state and potential **limitations** of mixture interpretation in the forensic community
- Participation
  - Initiated Summer 2014
  - 55+ participating labs
  - n=185 returned datasets



# Study Datasets:

---

- Examiners asked to deconvolute 6 identical mixtures:
  - Use their laboratory's SOP
  - Stochastic and analytical thresholds set by DFSC
  - Genotype interpretations recorded on Excel-based worksheet provided
  - Survey collected from each participant covering education, number of years experience, time spent on deconvolution, difficulty of interpretation, weekly caseload, etc.



# Examiner Worksheet Template



Excel interface showing a worksheet template for DNA analysis. The worksheet is titled "Test - Excel" and contains data for two datasets (1 and 2) with columns for Mixture ID, Number of Contributors (NOC), and various genetic profiles (Major, Minor 1, Assumed) with their proportions and genotypes across multiple loci (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317).

Dataset	Mixture ID	Number of Contributors (NOC)	NOC: Consistent With/At Least						
1	121	2	Consistent With						
Profile	Proportion		D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317
Major	0.74	Genotypes	13,15	29,29	8,9	11,13	16,20	6,9	9,11
		Analysis Comment	restricted (rRMP)	restricted (rRMP)	restricted (rRMP)	restricted (rRMP)	restricted (rRMP)	restricted (rRMP)	restricted (rRMP)
Minor 1	0.26	Genotypes	12,12	31,32.2	10,11		14*,16	6,7*,9	12,Any
		Analysis Comment	restricted (rRMP)	restricted (rRMP)	restricted (rRMP)	restricted (rRMP)	restricted (rRMP)	restricted (rRMP)	modified (mRMP)
		Genotypes							
		Analysis Comment							
		Genotypes							
		Analysis Comment							
2	202	2	Consistent With						
Profile	Proportion		D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317
Major	0.74	Genotypes	12,13	28,29	8,11	10,11	15,16	8,9	8,11
		Analysis Comment	restricted (rRMP)	restricted (rRMP)	restricted (rRMP)	restricted (rRMP)	restricted (rRMP)	restricted (rRMP)	restricted (rRMP)
Minor 1	0.26	Genotypes	13,16	28,32.2	11,11	10,12	14,18	8,9	11,12
Assumed		Analysis Comment							
		Genotypes							
		Analysis Comment							
		Genotypes							
		Analysis Comment							



# Automatic Analysis of Genotypes



7	<b>ID#</b>		
8			
9	<b>Number of Contributors (NOC)</b>	<b>NOC: Consistent With/At Least</b>	
10	<b>2</b>	<b>Consistent With</b>	
11			
12		<b>D8S1179</b>	<b>D21S11</b>
13	<b>Genotypes</b>	<b>13,15</b>	<b>29,29</b>
14	<b>Analysis Comment</b>	restricted (rRMP)	restricted (rRMP)
15		<b>12,12</b>	<b>31,32.2</b>
16	<b>Analysis Comment</b>	restricted (rRMP)	restricted (rRMP)



H:\OCS File Backup\Karl Mereus Files 2Jul2015\Examiner Temp... H:\OCS File Backup\Karl M... x

Allele : D19S51 Annotation: '12,10' True Value: '10,10' Scores: (0.267076707670767, 0, 1, 1)

Allele : D5S818 Annotation: '11,13' True Value: '11,13' Scores: (1.0, 2, 2, 0)

Allele : FGA Annotation: '21,23' True Value: '21,23' Scores: (1.0, 2, 2, 0)

Mixture 5

Profile 1

Allele : D8S1179 Annotation: '11,13,14,15' True Value: '11,15' Scores: (0.777777777777778, 2, 2, 0)

Allele : D21S11 Annotation: '27,29,30,30.2,32' True Value: '30,30.2' Scores: (0.9122807017543859, 2, 2, 0)

Allele : D7S820 Annotation: 'inc' True Value: '11,12' Scores: (0.0, 1, 0, 0)

Allele : CSF1PO Annotation: '9,10,11,13' True Value: '9,11' Scores: (0.818181818181818, 2, 2, 0)

Allele : D3S1358 Annotation: '14,16,18' True Value: '16,18' Scores: (0.8909090909090909, 2, 2, 0)

Allele : TH01 Annotation: '6,8,9,9.3' True Value: '8,9.3' Scores: (0.6428571428571428, 2, 2, 0)

Allele : D13S317 Annotation: '10,12,13' True Value: '10,13' Scores: (0.833333333333334, 2, 2, 0)

Allele : D16S539 Annotation: '11,12,13' True Value: '11,13' Scores: (0.833333333333334, 2, 2, 0)

Allele : D2S1338 Annotation: '18,19,20,23,24' True Value: '18,23' Scores: (0.8351648351648352, 2, 2, 0)

Allele : D19S433 Annotation: '13,14,15' True Value: '13,15' Scores: (0.9428571428571428, 2, 2, 0)

Allele : vWA Annotation: '15,16,17,18,19,20' True Value: '16,17' Scores: (0.533333333333333, 2, 2, 0)

Allele : TPOX Annotation: '8,11' True Value: '8,11' Scores: (1.0, 2, 2, 0)

Allele : D18S51 Annotation: '14,15,16,18' True Value: '14,18' Scores: (0.8901098901098901, 2, 2, 0)

Allele : D5S818 Annotation: '10,11,12,13' True Value: '12,13' Scores: (0.777777777777778, 2, 2, 0)

Allele : FGA Annotation: 'inc' True Value: '22,26' Scores: (0.0, 1, 0, 0)

Profile 2

Allele : D8S1179 Annotation: '11,13,14,15' True Value: '13,14' Scores: (0.777777777777778, 2, 2, 0)

Allele : D21S11 Annotation: '27,29,30,30.2,32' True Value: '27,30' Scores: (0.9122807017543859, 2, 2, 0)

Allele : D7S820 Annotation: 'inc' True Value: '8,11' Scores: (0.0, 1, 0, 0)

Allele : CSF1PO Annotation: '9,10,11,13' True Value: '10,11' Scores: (0.818181818181818, 2, 2, 0)

Allele : D3S1358 Annotation: '14,16,18' True Value: '14,18' Scores: (0.8909090909090909, 2, 2, 0)

Allele : TH01 Annotation: '6,8,9,9.3' True Value: '6,8' Scores: (0.6428571428571428, 2, 2, 0)

Allele : D13S317 Annotation: '10,12,13' True Value: '12,13' Scores: (0.833333333333334, 2, 2, 0)

Allele : D16S539 Annotation: '11,12,13' True Value: '11,12' Scores: (0.833333333333334, 2, 2, 0)

Allele : D2S1338 Annotation: '18,19,20,23,24' True Value: '19,20' Scores: (0.8351648351648352, 2, 2, 0)

Allele : D19S433 Annotation: '13,14,15' True Value: '14,14' Scores: (0.9428571428571428, 2, 2, 0)

Allele : vWA Annotation: '15,16,17,18,19,20' True Value: '19,20' Scores: (0.533333333333333, 2, 2, 0)

Allele : TPOX Annotation: '8,11' True Value: '8,8' Scores: (0.972222222222222, 0, 1, 1)

Allele : D18S51 Annotation: '14,15,16,18' True Value: '15,15' Scores: (0.8901098901098901, 2, 2, 0)

Allele : D5S818 Annotation: '10,11,12,13' True Value: '10,13' Scores: (0.777777777777778, 2, 2, 0)

Allele : FGA Annotation: 'inc' True Value: '22,26' Scores: (0.0, 1, 0, 0)

Profile 3

Allele : D8S1179 Annotation: '11,13,14,15' True Value: '13,15' Scores: (0.777777777777778, 2, 2, 0)

Allele : D21S11 Annotation: '27,29,30,30.2,32' True Value: '29,32' Scores: (0.9122807017543859, 2, 2, 0)

Allele : D7S820 Annotation: 'inc' True Value: '11,12' Scores: (0.0, 1, 0, 0)

Allele : CSF1PO Annotation: '9,10,11,13' True Value: '10,13' Scores: (0.818181818181818, 2, 2, 0)

Allele : D3S1358 Annotation: '14,16,18' True Value: '16,16' Scores: (0.8909090909090909, 2, 2, 0)

Allele : TH01 Annotation: '6,8,9,9.3' True Value: '9,9' Scores: (0.6428571428571428, 2, 2, 0)





# What is the Ideal Deconvolution?

---



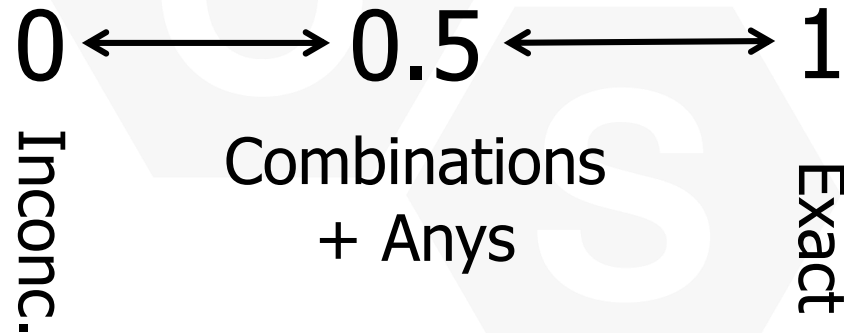
- Correctly determine the number of contributors (NOC) in the sample
- Generate the correct genotypes for each contributor in the sample
- Narrow down as much as possible the number of possible combinations
- Analyze via metrics:
  - Genotype Interpretation Metric (GIM)
  - Allelic Match scoring (AM)





# Metrics: GIM

$$\text{GIM} = \frac{(1 - C/C_{\text{STR}})}{2^A}$$



***How many answers did I provide at each locus?***



# Metrics: AM (Allelic Match)

Known	Generated	AT	AF	Inc
11, 12	11, 12	2	0	0
	11, Any	1	0	0
	11, 13	1	1	0
	10, 13	0	2	0
	Inc.	0	0	2

***Did my genotypes include the “correct answer”?***

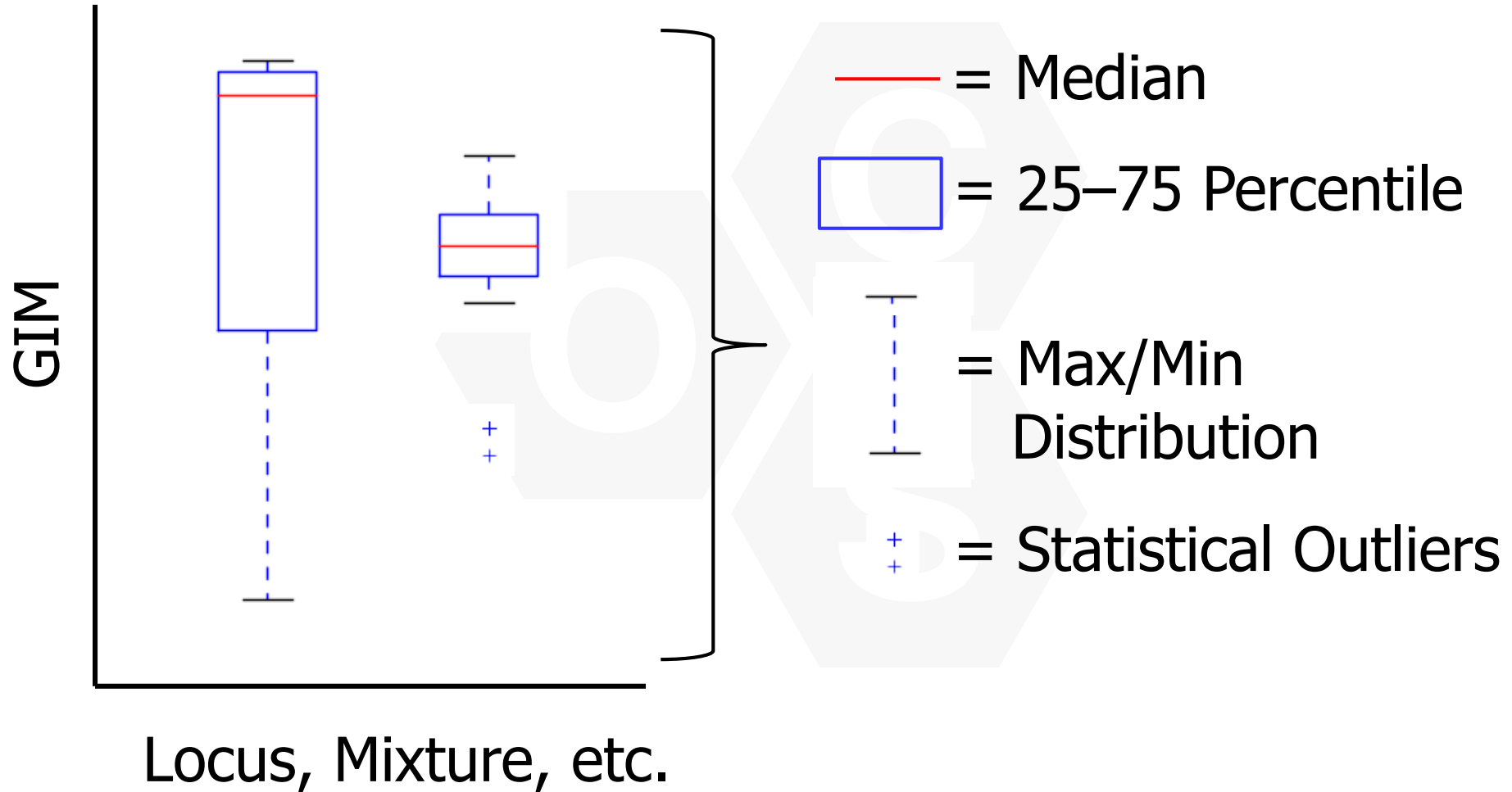


## In total

---

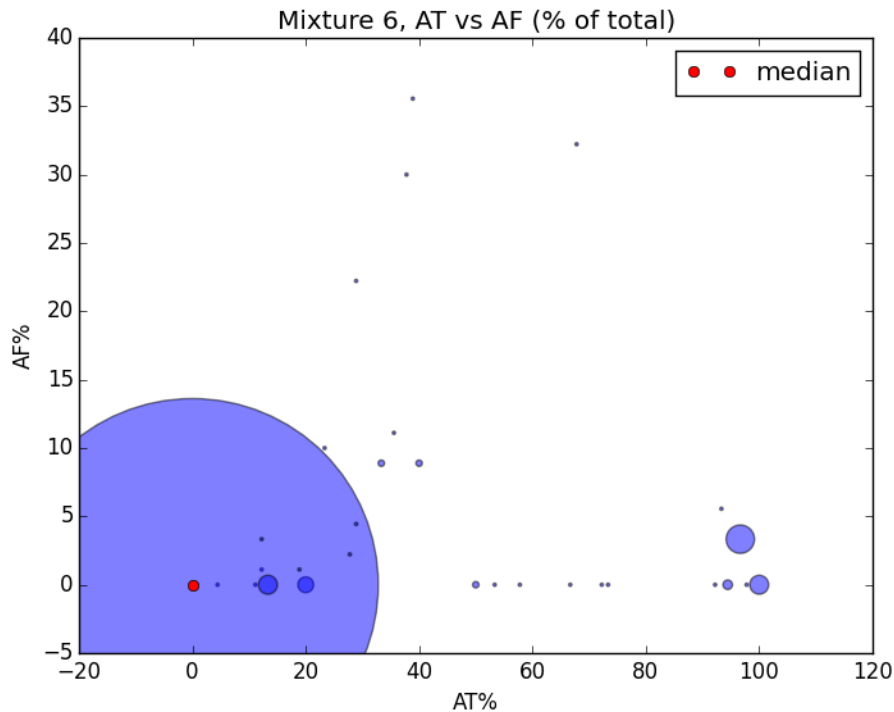
- Six mixtures
- Four 2-person mixtures, with two profiles Major and Minor
- Two 3-person mixtures, with three profiles Major, Minor 1 and Minor 2
- Each profile has 15 loci
- Each locus receives a GIM score out of 1, and a AM score out of 2
- => total GIM score out of 210
- => total AM score out of 420

# Visualize Variation: GIM Box Plots





## Visualize AM Variation: Scatterplots



- AT vs. AF
- AF vs. GIM
- AF vs. Inc.
- Each dot represents a single examiner
- Larger radius, increased number of examiners with same score



# Applications of the DEAT

---



1. Uncovering interpretation issues
2. Providing examiner and lab training benchmarks
3. Measuring changes in SOP
4. Revealing accuracy limitations depending on mixture complexity



# Interpretation Issues: by locus



## Mixture 1: Major and Minor\*

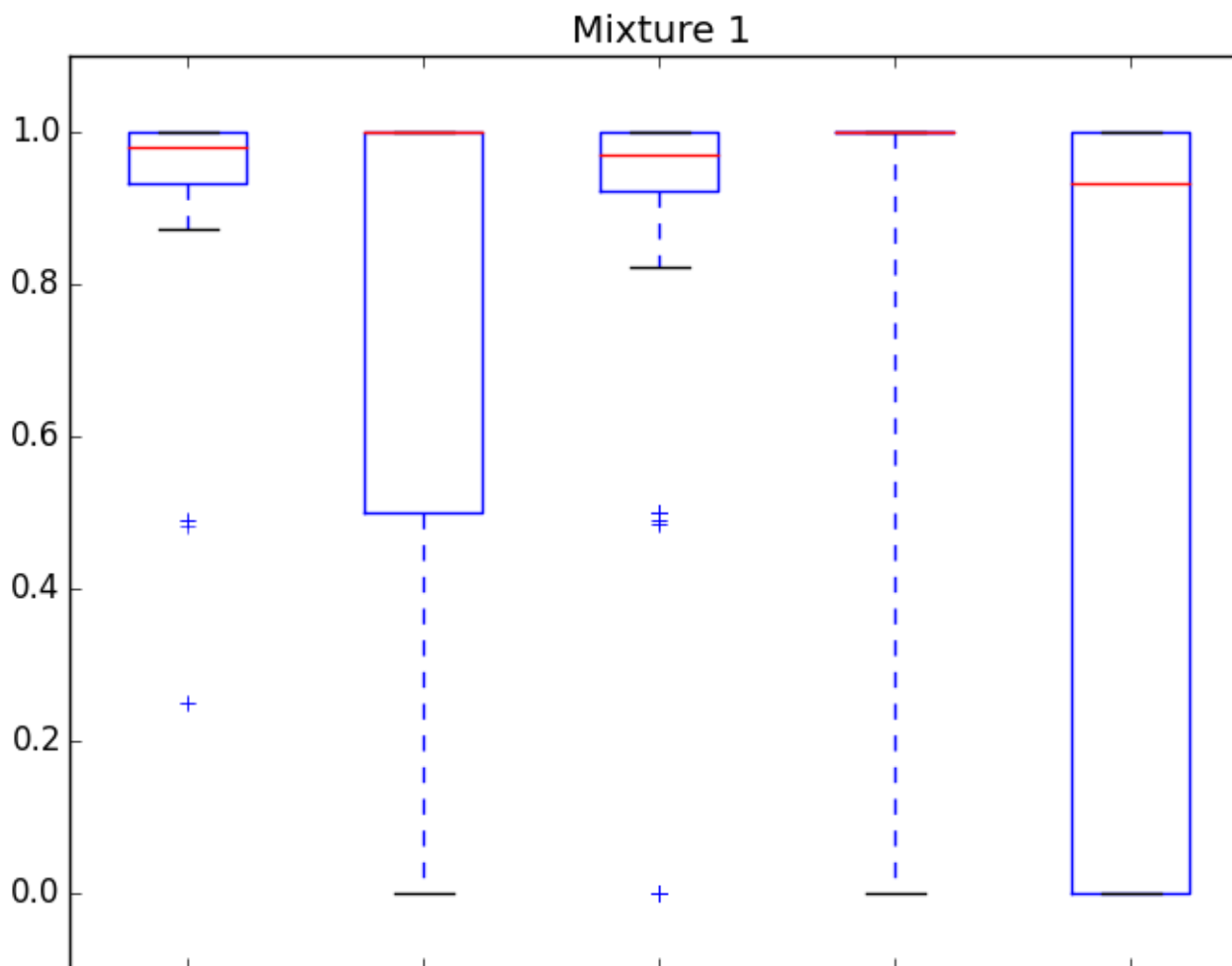
Locus	AT%	AF%	INC%		Locus	AT%	AF%	INC%
D8S1179	65.8%	0.8%	31.6%		D2S1338	89.0%	0.3%	10.3%
D221S11	89.4	0.3	10.3		D19S433	77.3	1.0	12.6
D7S820	88.2	0.5	11.3		vWA	76.6	0.7	11.9
CSF1PO	59.0	0.3	37.4		TPOX	63.2	0.8	32.9
D3S1358	87.1	1.3	8.1		D18S51	86.1	1.0	11.6
THO1	83.7	2.4	8.4		D5S818	81.0	0.5	11.9
D13S317	70.7	0.8	16.5		FGA	88.5	0.8	10.7
D16S539	76.9	0.5	12.9					
<b>MIXTURE AVG</b>	<b>78.8</b>	<b>0.8</b>	<b>15.9</b>					

\*ID+, n=155



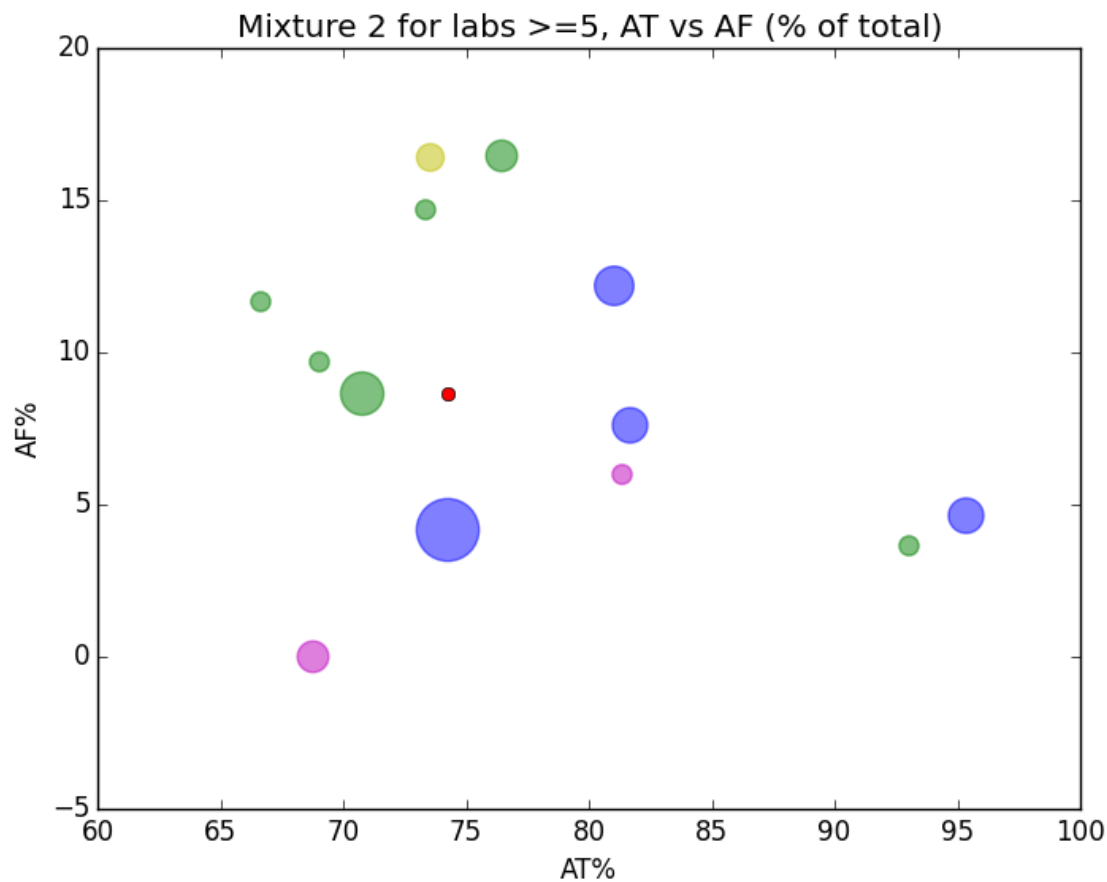


# Interpretation Issues: by lab





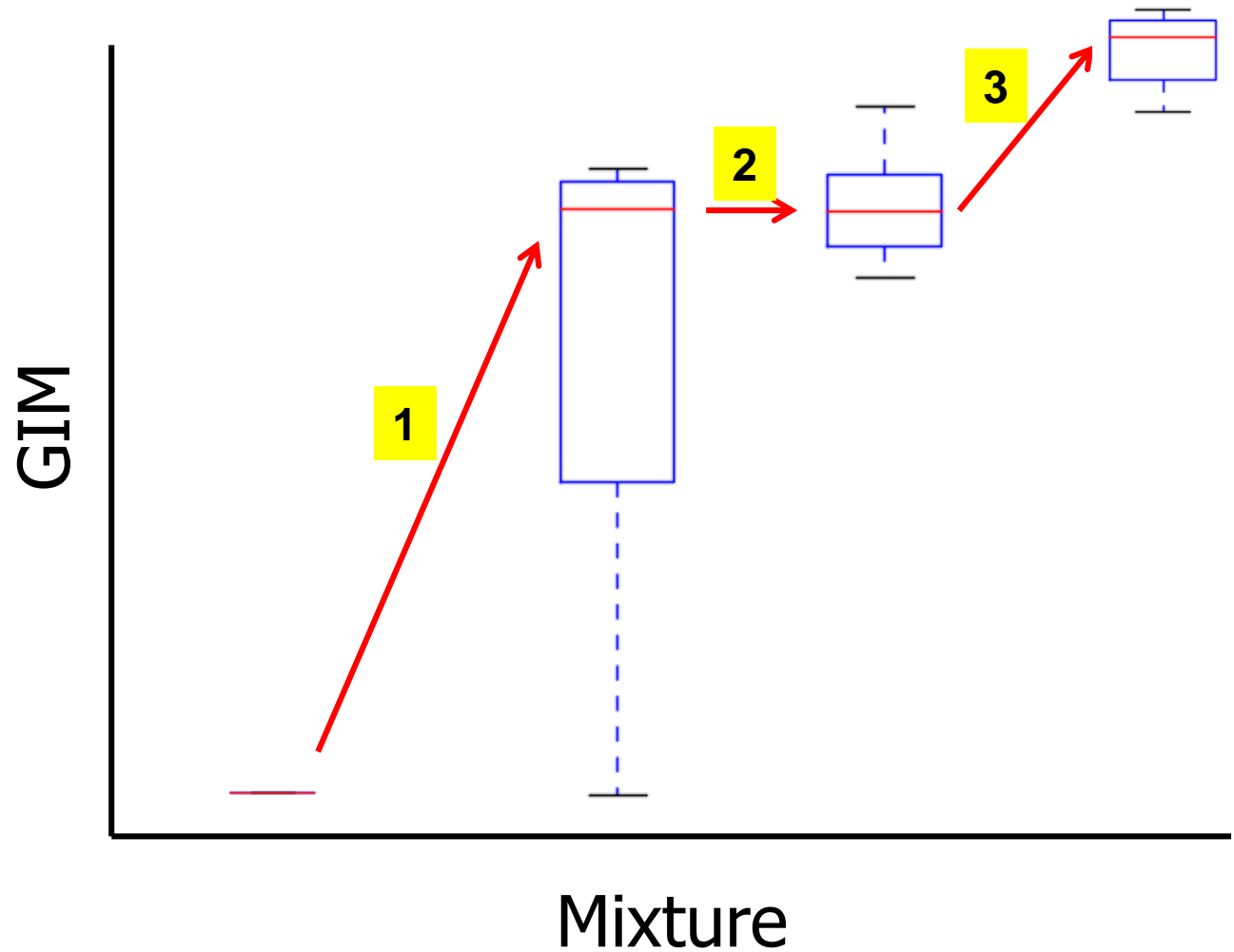
# Interpretation Issues: by region





# Providing training: Benchmarks per lab

1. Interpretation
2. Variation
3. Deconvolution

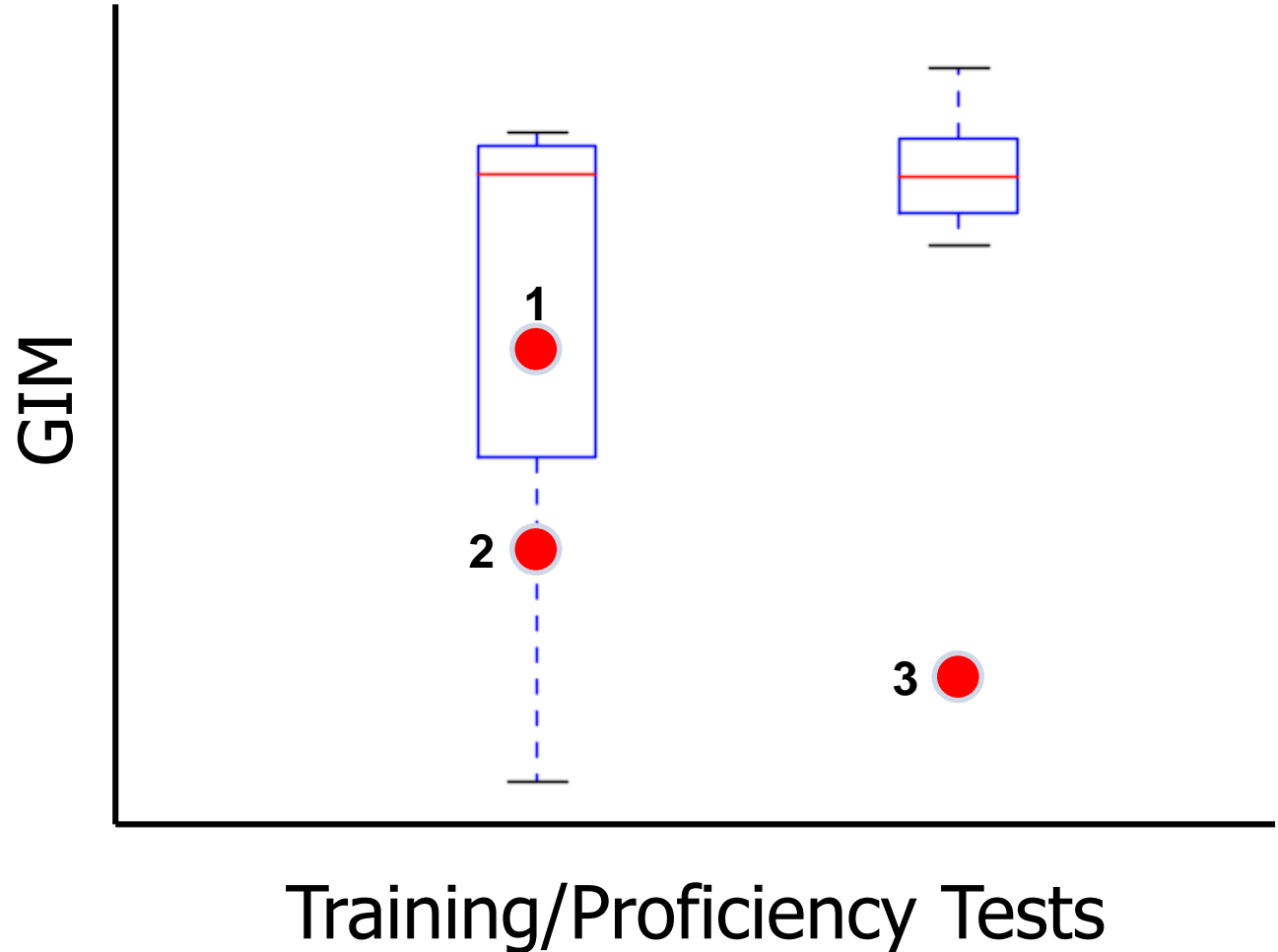




# Providing training: Benchmarks by examiner

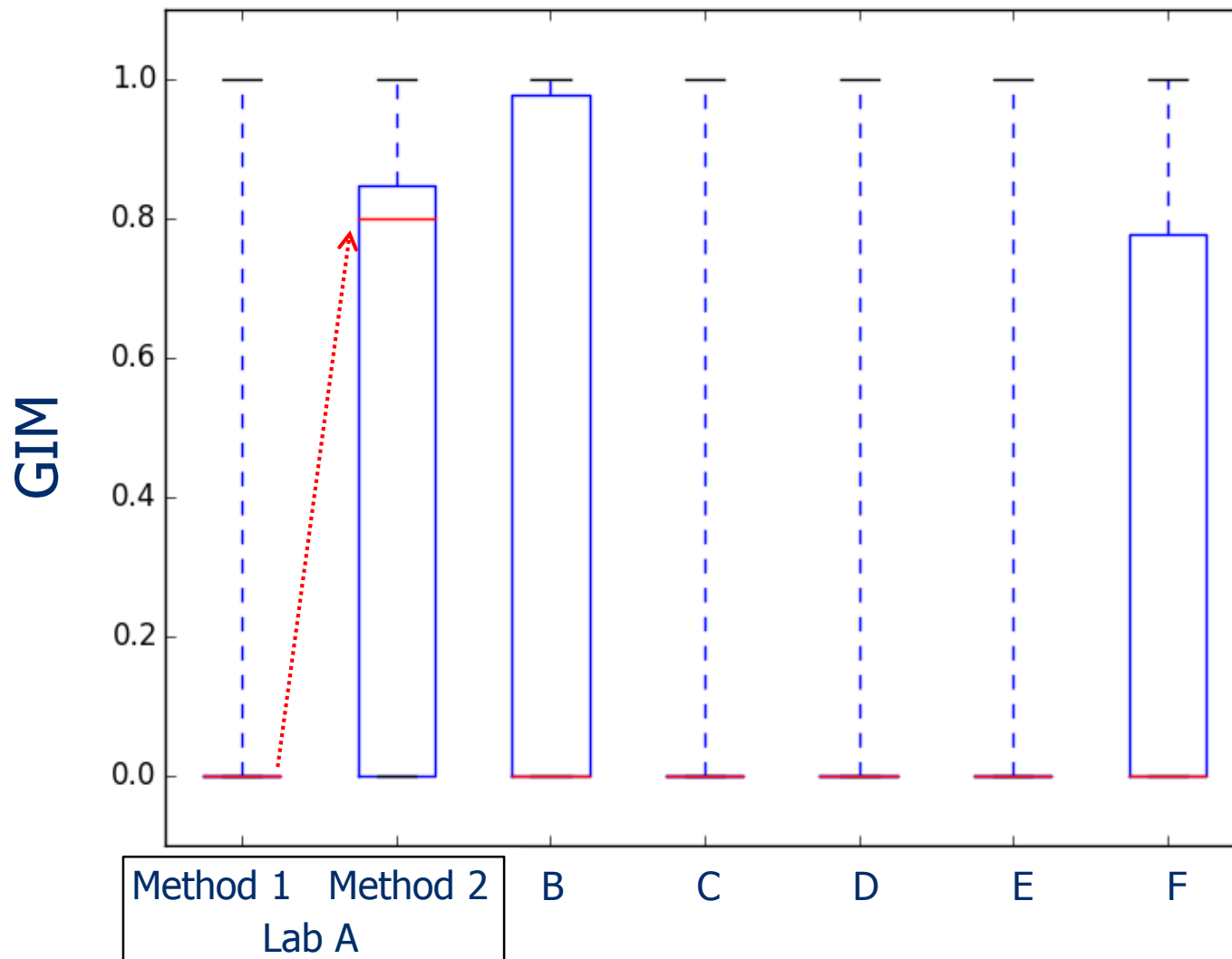
Training of new and existing DNA examiners:

- 1. Within  $\Delta$  range
- 2. Minor  $\Delta$ :  
review/retest?
- 3. Major  $\Delta$ :  
retrain?



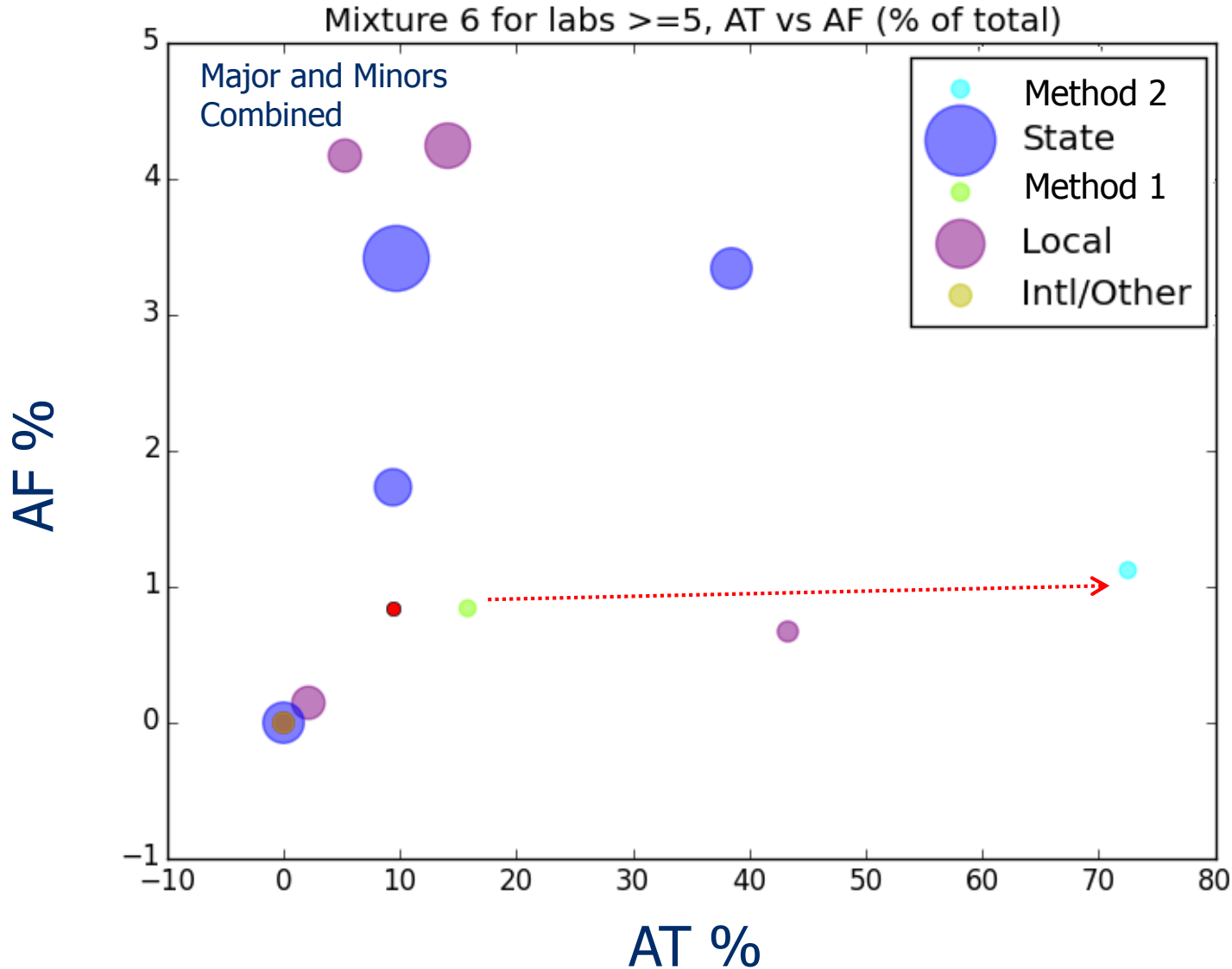


# Adopting new methods: Changes to SOP



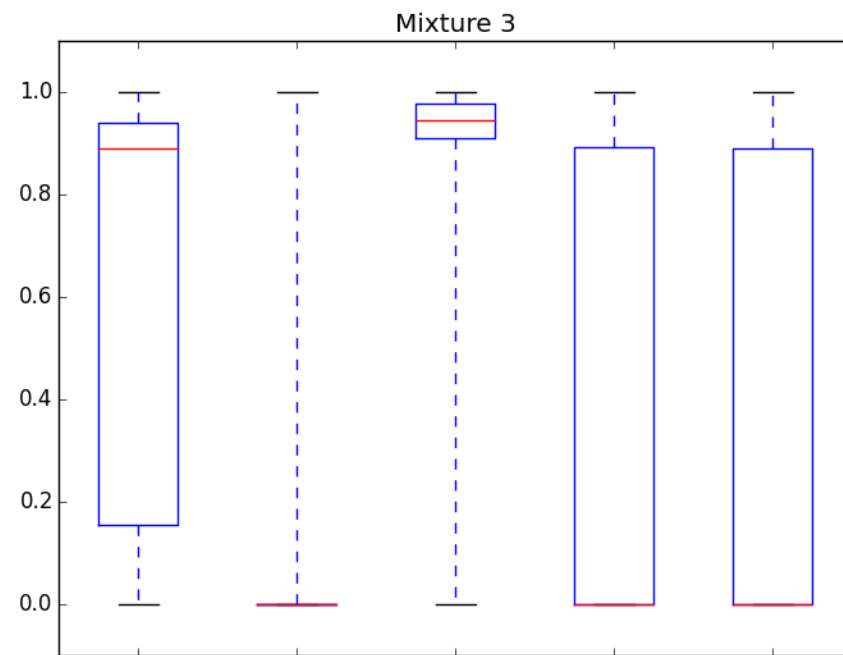
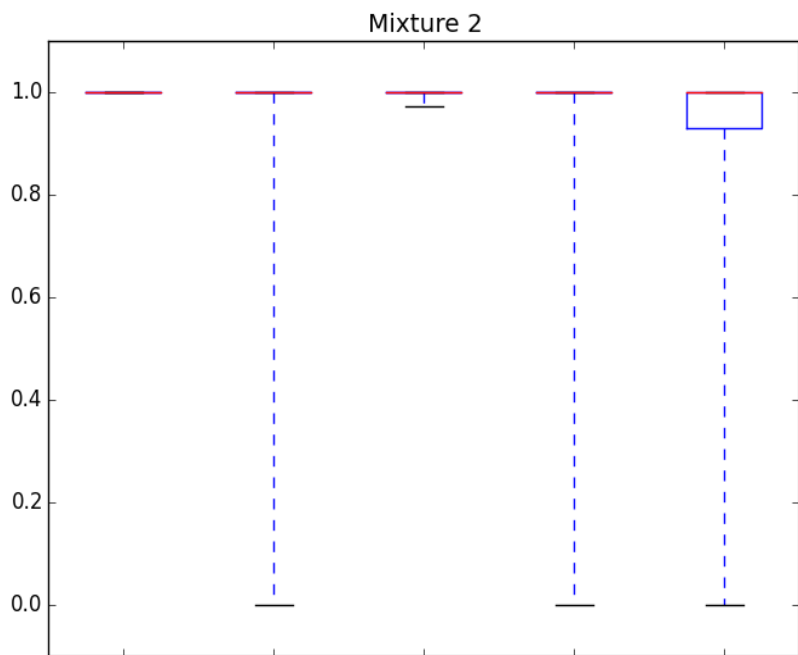


# Adopting new methods: Improved Accuracy?





# Uncover accuracy/interpretation limitations



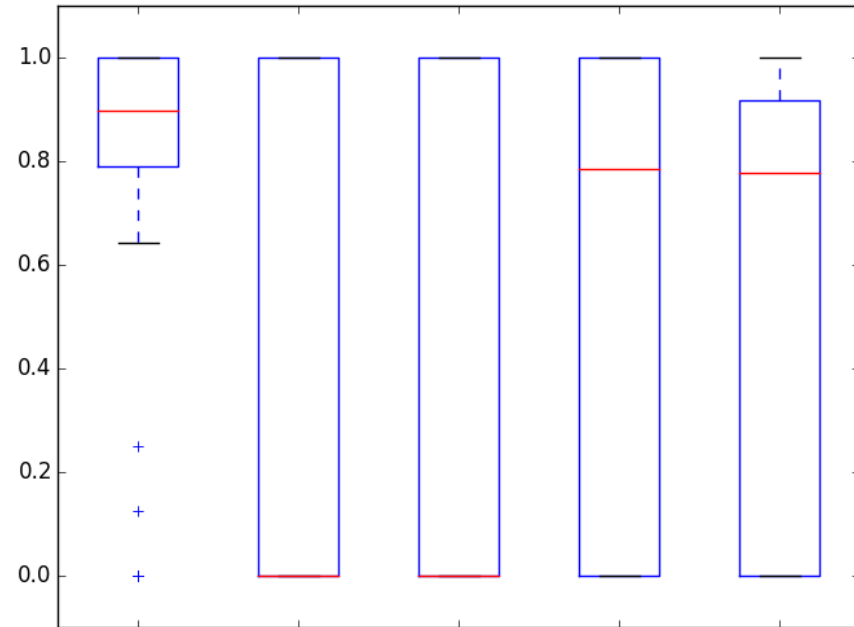




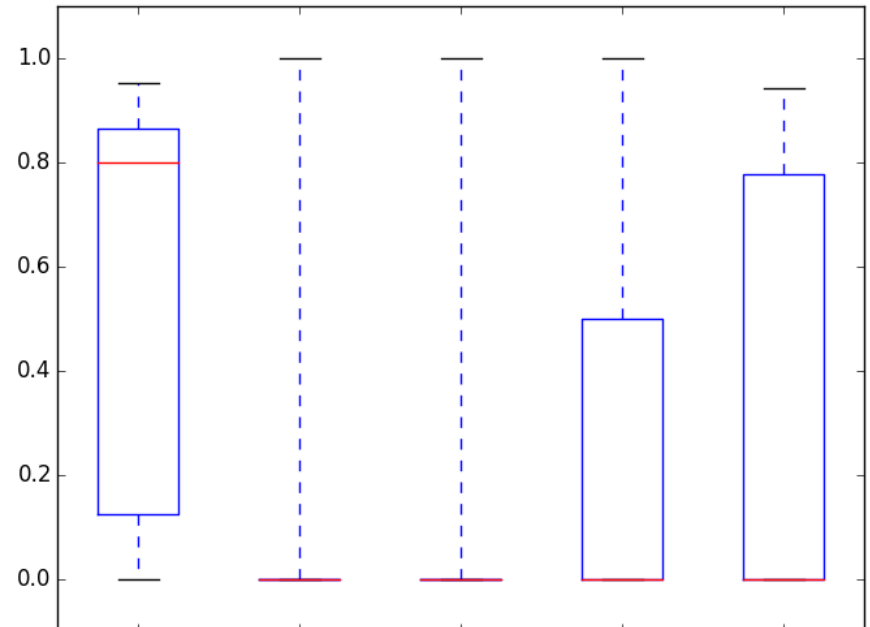
# Uncover accuracy/interpretation limitations



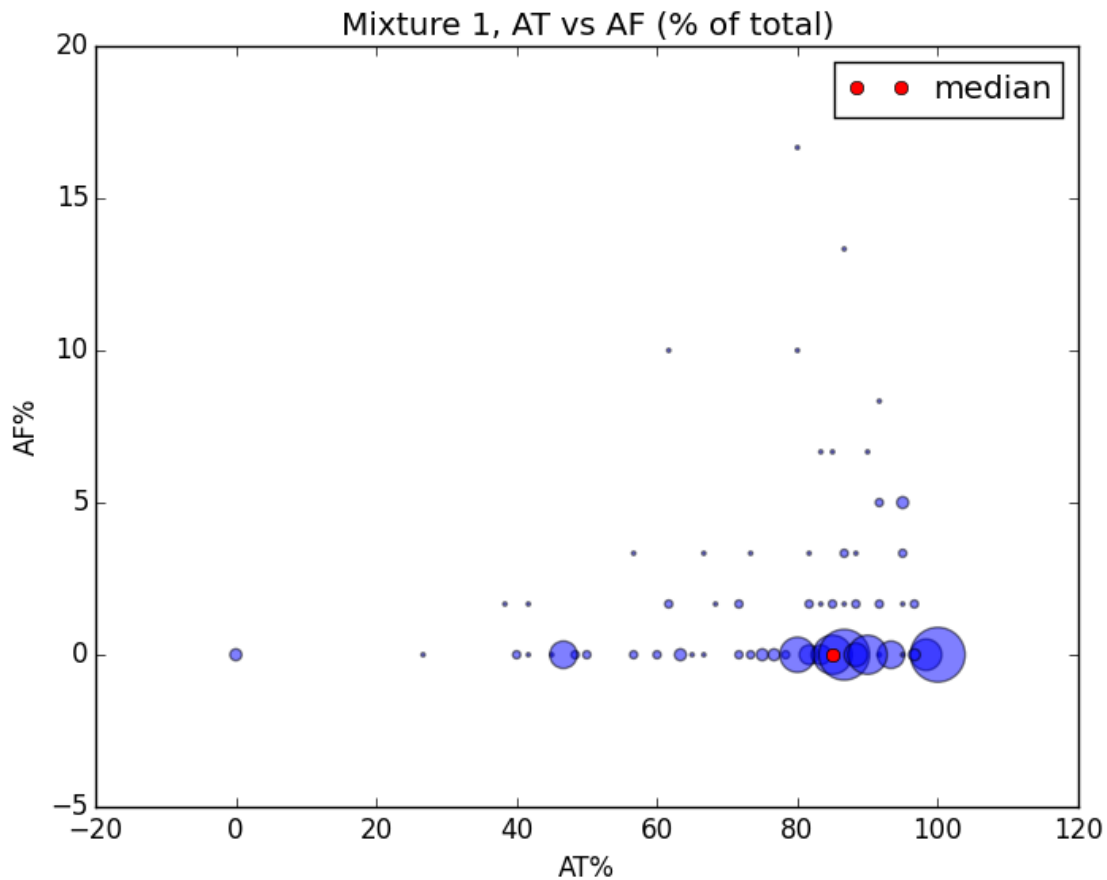
Mixture 5



Mixture 6

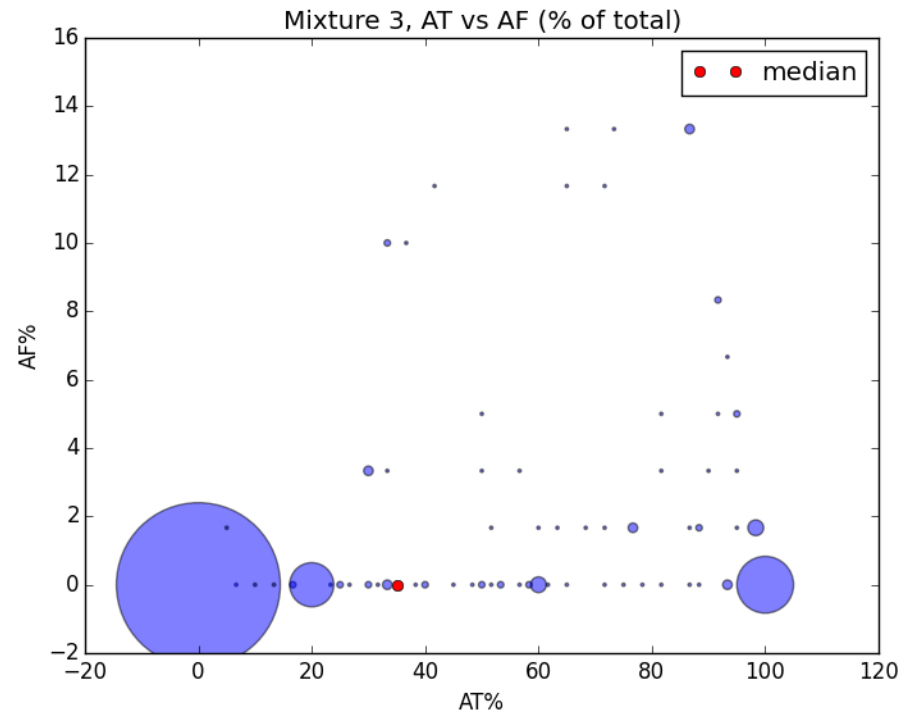
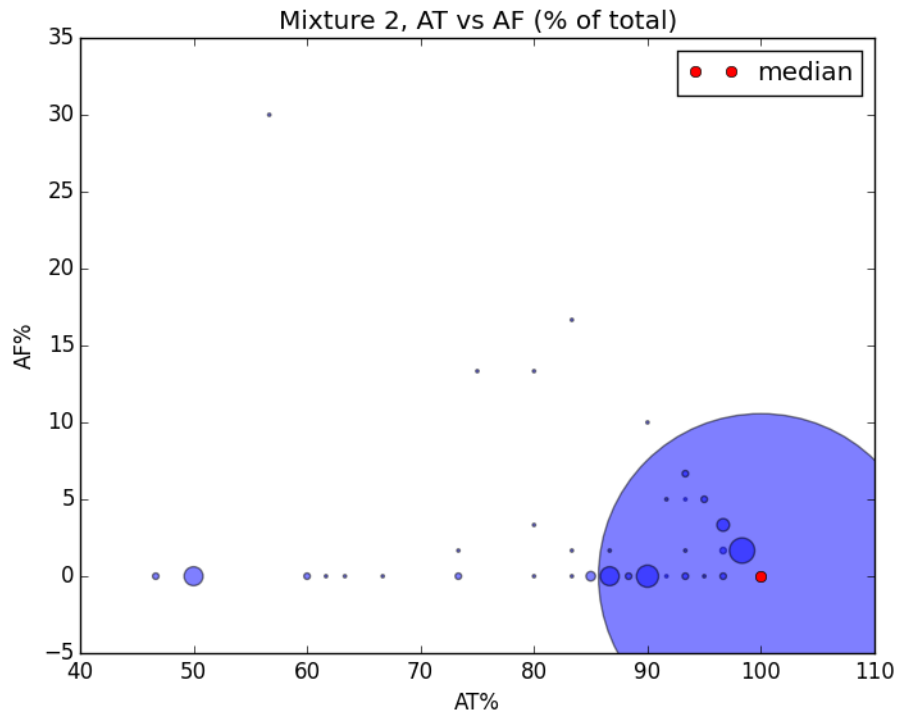


# Baseline AF vs. AT Scoring per lab



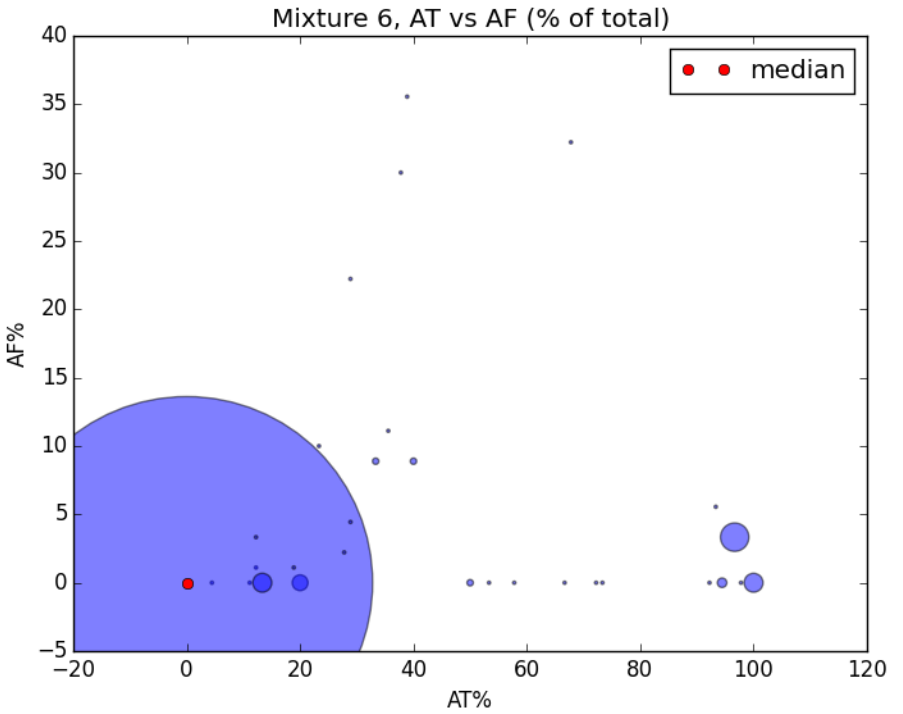
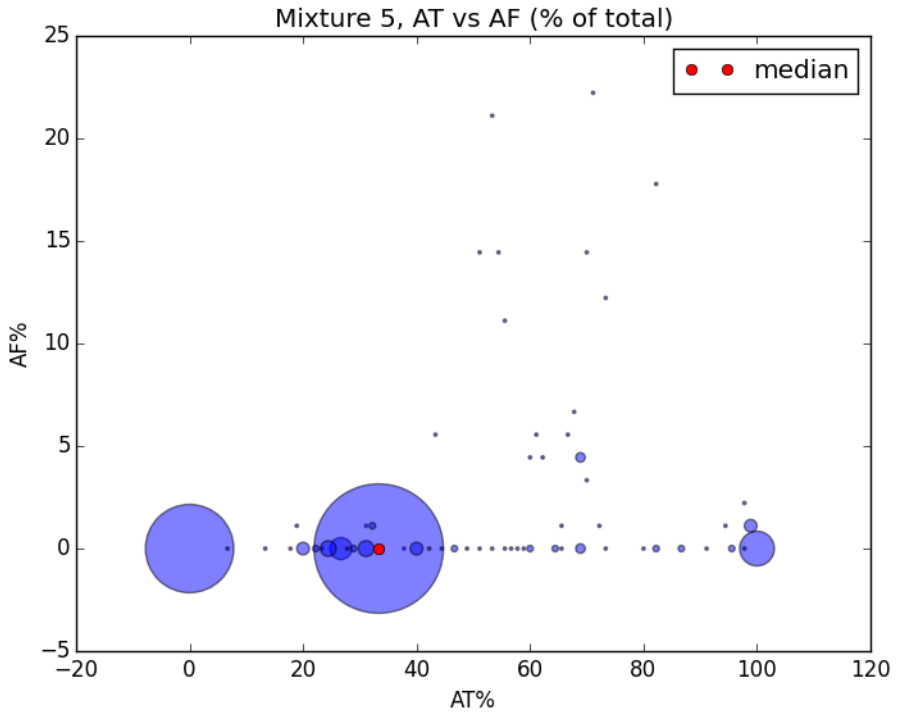


# Comparing samples W/ and W/O Reference





# 3-person Mixtures: Changes to accuracy?





# Summary

---

- Potential use of DEAT:
  - In uncovering areas of interpretation weaknesses
  - In training, assessment, and proficiency testing of DNA examiners individually and of labs as a whole
  - In providing a measure of effectiveness to changes in SOP or protocols
  - In revealing limitations of accuracy depending on mixture complexity
- In the future: an online version of DEAT available to laboratories for training of new and assessing existing DNA examiners

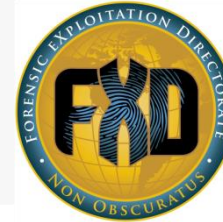


# Acknowledgement

Office of the Chief Scientist, DFSC



Forensic Exploitation Directorate, DFSC



National Institute of Justice for partial funding



Defence Science and Technology Laboratory (Dstl, UK)





**Roman Aranda IV**  
**Research Scientist**  
**Office of the Chief Scientist**  
roman.aranda3.ctr@mail.mil  
404-469-7253

