Ref: Prosecutors fight DNA use for exoneration; Aug 29, 2003 Adam Liptak

need: 60 minutes on bullet analysis (now on DVD)

http://www.cbsnews.com/stories/2007/11/16/60minutes/main3512453.shtml

60 minutes 2022:
https://www.youtube.com/watch?v=H4g62cpRz7M

eyewitness demo: http://www.wimp.com/innocentmen/ (defunct)

https://www.wimp.com/a-roomful-of-people-condemn-five-innocent-men/

https://www.youtube.com/watch?v=D7EpYlKaBM8 Gary Wells

https://www.youtube.com/watch?v=oZPOHIndSVQ lineup for Wells

http://www.innocenceproject.org/dna-exonerations-in-the-united-states/

Science and the US criminal justice system: a revolution from DNA typing

Innocence Project

1) As of 2014, 318 people in US prisons have been released after new evidence showed that they could not have committed the crime (more than 15 of these were from death row) or that the evidence used to convict was seriously flawed. The most common cause of mis-conviction was mistaken identification by eyewitnesses.

as of 2001-2: 67 people had been freed because of DNA tests done after the conviction, which at the time was 2/3 of those tested after conviction.

In 2003, the Houston Crime lab lost its accreditation for DNA typing because of sloppy procedure. October 2007, there was another erroneous conviction identified from that lab's mistakes. Response across the state: make sure no one ever gets access to the labs again.

Most of these wrongful convictions can be attributed to bad data when identifying people or objects. How we ID people is often flawed, but so are most aspects of forensic 'science.'

Besides DNA, what other ways to we have of identifying you or of matching samples?

fingerprints	eye witness
hair	lie detectors
dog smell	bullet lead
blood type	bullet striations
shoe print	
bite marks	

What can go wrong with a matching method?

misleading claim of RMP (such as uniqueness or 'in my experience …')
lab error in doing the work

What do you want in a method – to know that you can trust testimony?

The expert can consistently get the right answer
Others can do the assay and get the same answer
No subjectivity in scoring
We actually know how rare the match would be by accident (RMP)

Definitions:

ref database: a large sample of the reference population used to judge the significance of a match between a suspect and a crime scene sample; for DNA is a sample of DNA types from many different people

proficiency test: the submission to a lab of samples of known identity to assess whether the lab is getting the right answer. Blind refers to whether the lab knows its being tested at the time. Think of it as a competency test.

The ideal features of a forensic procedure for **identifying people or things (bullets)**:

Feature	why needed	error red'n principle	flags- indicators of absence
1) labs subjected to and pass blind proficiency tests (high threshold)	provides assurance of the consistent accuracy of the methods in the hands of the lab providing the analysis	= standards to estimate overall H&T error rate	no error rate given; tests internal, not blind, undocumented
2) independent verification of match possible a) universal protocol b) characters permanent	someone else can challenge or repeat the results	= replication, to detect many types of error, especially bias, H&T and to avoid intentional falsification	methods of one expert cannot be evaluated by another; no explicit protocol; characters being measured are not permanent
3) reference database exists (large sample of reference population)	gives the population frequencies of the different characteristics, thus knowing the RMP	to calculate sampling error inherent in a match	not mentioned or claims that match is unique
4) characteristics measured are	removes subjectivity -- it is clear whether a person	avoid RPA error	no description of specific

discrete (not subjective)	has it or not, allowing consistent scoring		characteristics
5) blind processing and analysis	ensure unbiased interpretation of data (unless a machine is recording the data)	avoid bias	failure to specify blind; indication that samples were labeled with names

Go to 60 minutes bullet lead

Bullet lead video: **http://www.cbsnews.com/stories/2007/11/16/60minutes/main3512453.shtml**
or on Youtube: https://www.youtube.com/watch?v=H4g62cpRz7M

Let's think about the types of errors and where/how they can arise:

Error	Where it can arise	consequence	example of or basis of concern	Fix
H&T	all tests, evidence records; analysis	false incrimination; false exoneration	sample mixup; bad protocol	proficiency test, replication by other labs
bias	interpretation of subjective evidence; choice of evidence or where to test; falsification of evidence	data can be 'made' to fit a suspect	eyewitness interviews suggestive, lab procedures not blind	BLIND proficiency test, replication by other labs
RPA	some types of lab analyses	matching criteria can become subjective	bullet lead video	discrete characters
sampling	occasional match by chance	match would be misinterpreted	recent fingerprint match	large ref database

DNA – good on all of above but +/- on blind proficiency tests

DNA is a good model in all respects, though mistakes are still made to the tune of 1-2%

History of DNA

Fill out first column in table below, then ask: Over last 2 decades, which methods used by prosecutions against defendants have been found to have error rates of 10% or more in proficiency tests or experiments? all but DNA

Summary of forensic Identification methods and characteristics **(put on web)**

(1) labs pass proficiency tests (routinely)
(2) independent verification possible
(3) reference database
(4) characteristics measured are discrete

Method	(1)	(2)	(3)	(4)	discredited?
DNA	varies	+	+	+	no
bullet lead	-	-	?	-	yes
fingerprints pre-1990	-	-	- §	-	yes
fingerprints post 2000	?	+?	+	+/-	no
hair matching	-	-	-	-	yes
bite marks	-	-	-	-	yes
shoeprint ID	-	-	-	-	yes
dog sniffing	-	-	-	-	yes
eyewitness	-[1]	-	-*	-	still used

§ - a database existed but was difficult to screen and could not be used to give a RMP
* a photo catalog can be considered a ref database but is not used properly as one bc/ the witness is only allowed to choose one
[1] – individuals fail proficiency tests, but this is not the same as a lab being subjected to a prof test

Fingerprints: –

3) ref database exists (for over half a century) but not searchable for RMP

4) initially, no discrete characteristics for declaring a match; there is now (2000 on)

1) independent analysis possible but was no explicit protocol for declaring a match before ~2000. Thus no independent verification in practice.

2) voluntary proficiency tests offered to labs in 1995-8, using multiple fingerprint comparisons. Not all labs responded, but for those that did respond, the false positive error rate of labs was at least a few percent and was as high as 22% !!!!!!!!!!!!

Surprisingly, the main international association of fingerprint experts (which, incidentally, consisted mostly of US experts) resisted the establishment of criteria for demonstrating a match into the 1990s. They instead proclaimed that each decision about a match was to be made on a case-by-case basis and should be left up to the expert reviewing the case. (There was disagreement about this point between the two main fingerprint organizations, and the British adopted a minimal set of criteria for declaring a match.)

once the most trusted method of identification in forensics, fingerprints fell from grace in the 1990s. Long considered the icon of personal identification, fingerprint experts were finally subjected to proficiency tests in the mid-late 1990s and found to make false matches 10%-20% of the time.
There has long been (1) reference databases; failures due to problems with (2), (3), (4)

history:

1911: first successful introduction of fingerprint evidence in US court, but not as the sole evidence.
1941: legal precedent was established for convictions based on fingerprint evidence alone.

then or later: acceptance of the general assertion that a single fingerprint was also unique and could be used to establish identity.

So now, fingerprint ID is considered OK, but a recent erroneous match indicates its fallibility

Hair matching: bad protocols, bad standards.
Hair-matching methods, often used in court to establish that a single hair came from a particular person, have been declared nonsense. In 2003, Canada not only abandoned use of (non-DNA) hair matching methods, but began reviewing convictions that used hair matching, to see if the convictions should be overturned.

21 of 130 (2007) wrongful convictions listed by the Innocence Project involved hair matching.

3) no reference databases for hairs
4) characteristics not discrete
1) independent verification should have been possible
   a. no objective criteria for a match or RMP (no way of coding hairs)
2) proficiency tests in early 1970s found error rates of 28%-68% , and different labs made different mistakes (as expected if there is no uniform protocol for doing the matches)

Now discredited (is basis of Quiz 5)
In hindsight, there were many major problems with it that should have kept it from ever seeing the light of day.

Dog sniffing identification: (2 recent articles on webpage)

3) no reference database
4) no discrete characteristics
1) independent verification
   a. no explicit protocol (can't control a dog's brain)
   b. characteristics not permanent (odors change)

2) proficiency tests not done (informal ones revealed problems)

http://www.statesman.com/news/texas-politics/appeals-court-frees-man-convicted-of-murder-based-932506.html?cxtype=rss_news&viewAsSinglePage=true

is effectively a method lacking in protocols. There is no way to know what method a dogs is using for odor identification and matching. Tests using trial dogs have found that dogs are not very good at matching odors from different parts of the same person (e.g,, hand versus neck). Fails on (1)-(4); proficiency tests ARE possible but not attempted.

Bite marks: also junk for all of the above reasons.

Eyewitness identification:

3) photo catalog can be considered a reference database but is not used correctly – witness not allowed to select more than one person.

4) no discrete characteristics (people don't mentally record individual features; yes or no for whole thing)

(1)

i. not possible to specify a protocol for what people remember
ii. memories are dynamic, so what is remembered is not permanent (hence susceptible to bias) (<u>see below</u>)

(2) consistent failures by normal people in 'proficiency tests' going back 100 yrs; furthermore, eyewitnesses are often positive of their ID when it's wrong. (<u>see below</u>)

To help avoid mistaken ID, the protocol for a lineup (including instructions) is critical (as per our class demonstration). Innocence Project page has recommended protocol

<u>See Gary Wells page for an eyewitness ID demo</u>
Give option of 7) not sure (one class) and 7) None or not sure (other class)

Eyewitness identification of a suspect is the most powerful evidence there is for swaying a jury. And it is among the most fallible of all evidence: **173 of 225 (2009) wrongful convictions** tabulated by the Innocence Project involved mistaken ID, the most common error attributed to wrongful conviction.

(2) independent verification when multiple witnesses? the same errors by different witnesses

Errors by eyewitness testimony in court have been documented for decades, and some of them are profound. There have been several cases in which *half a dozen* or more eyewitnesses identified the same person, and it was the *wrong person*. In many cases, there is not even a close resemblance between the right and the wrong person. When several people all make the same mistake, it is a clear indication that police protocols are bad – it generally means that the police are influencing witnesses, or witnesses are influencing each other.

(1) proficiency tests? 100 year history of demonstrated eyewitness fallibility

It has been known for over a century that eyewitness accounts are far from perfect. A 1902 experiment conducted in class (involving a gun – not something we'd do now) revealed that the best witnesses were wrong on 26% of important details; the worst had an 80% error rate. In a more recent California State University experiment with a staged attack on the professor, only 40% of the class later identified the attacker, and 25% attributed the attack to a bystander.

Why so easy to be wrong?

Two points:

1) **memories are dynamic**, recreated to make things fit together in a consistent picture
2) **biased police procedures affect eyewitnesses** so that different witnesses are directed toward the same suspect

Let's start with a single eyewitness. Some psychologists use a 3-sequence model of memory:

i) acquisition – the events are recorded in your brain
ii) retention – the acquired events remaining in your brain are lost at some rate
iii) retrieval – the events are recalled by you

The acquisition phase is known to be imperfect – you never record all aspects of an event. That is, your memory starts out like a photograph with gaps. The more distracted or stressed you are at the time, the less you acquire. Thus, a person being raped or facing a gun/knife will have a much faultier acquisition than a person facing a non-threatening situation.

The retention phase is also imperfect. However, not only can you lose the memory of an event you had once acquired, you can also add events that never happened. Your memory is dynamic, and you are constantly building it, often filling in the holes. This rebuilding of the memory is where problems arise with eyewitnesses. In particular, your memory is very prone to subtle suggestions which eventually bias what you remember. Here are two problems that confound witness identification.

- Subtle hints can influence a witness to choose a particular suspect, even if that suspect is not the right one. These hints can be as innocuous as the police merely asking the witness to "take a careful look at #3" for example. Anything that makes one suspect stand out from the others can be a subtle influence on the witness (the way a picture of a suspect is taken, what they are wearing, etc.). Bad protocols and an absence of blind testing contribute to this bias. (The absence of blind exists if the police are influencing the witness and know who is their preferred suspect). The witness should not experience *any* outside factors that will push their choice toward a particular suspect.

- Reinforcement. Once a witness has been exposed to a set of suspects, any subsequent exposure to one of the suspects reinforces the witnesses memory of THAT suspect. For example, if a witness is shown two lineups (at different times), and one suspect is common to both lineups, that suspect is likely to be chosen because of the familiarity. This problem is difficult to eliminate completely, because the familiarity may have been obtained before the witness saw the crime, so police procedure could not, in that case, prevent it.

- Familiarity transfers from one setting to another. If an eyewitness has had some previous exposure to a suspect in a setting unrelated to the crime, it is common for that familiarity to be transferred over in the witness's mind to the new context. In one case, the witness mistakenly identified a man who lived on her block (but whom she had only seen at a distance on a couple of occasions). In an experiment listed above, 25% of witnesses chose the person who had been a bystander to the crime – no doubt because that person was familiar to them, even though not committed the act.

Other methods (not identification)

gather evidence, interviews, lab tests, polygraph ...

ideal data problems
rampant absence of blind, standards/proficiency tests
use of protocols that are biased toward incriminating prime suspect (read letters)

reinforcement or selective pursuit of evidence: determine a prime suspect and then collect evidence against them. This method merely reinforces their views rather than helping them identify mistakes.

Since the routine use of DNA testing was implemented, 25% of the time the prosecution's prime suspect has been cleared by DNA *before* trial. This means that the prosecution's initial stage of gathering evidence led them to the wrong person. Because of the biases built into the prosecutorial practices, many of these 25% would have been convicted if DNA typing had not been available.

Other methods:

what	blind	consistent protocol	approp stds	recorded	discredited
polygraph	-	-	-	+	yes
suspect interviews (allowed to lie)	-	-	n/a	-/+	
lab tests	-	+	?	-/+	

ideal data problems:

i) Failure to gather and analyze data blindly. Bias
ii) Bad standards: failures to conduct proficiency tests, inadequate (non-existent) reference databases
iii) Independent verification not attempted

Main types of error: Human & technical, bias

bias is difficult to correct, however, because many aspects of prosecutorial duties *cannot* be done blindly.

other problems: prosecution gets first crack at most of the evidence; may "forget" to turn over evidence to defense that is useful to defense because it hurts the prosecution's case

false confessions: encouraged through psychological encouragement of getting witness to help solve the case; there is no incentive to check their validity (15 of 62 wrongful convictions)

Other specific methods used that have problems

Polygraph (lie detector):

no reference database (for an individual's responses)
lack of blind
no objective protocol for interpretation of the data

A polygraph records indirect measures of anxiety (heart rate, sweat, ...). The first problem is that there is no unambiguous way of connecting lying to anxiety. Second, anxiety is affected by many factors, even how aggressively the person asks the questions. Higher levels of anxiety are interpreted as lies, but a person's responses may be affected by many things other than whether they are lying (no ref. database for this). The interpretation is also subjective, and the questioning itself can influence a person's anxiety level, both of which are not done in a blind fashion, so the person conducting the interview can influence the anxiety level. On top of it, people can be trained to 'beat' the polygraph, and the method is at best only designed to detect people who are aware they are lying.

A report released 8 October, 2002 by the National Academy of Sciences described polygraph testing as little more than junk science. Although a 1988 federal law banned the use of such tests for employment screening in most private businesses, and polygraph data had been inadmissible in nearly all state courts, the method has been widely used in government agencies concerned with national security. There was a time when polygraph data were used in court, and the polygraph has been used for unofficial purposes in criminal investigations to help prosecutors decide who to rule out as suspects. Thus, the fact that it has been inadmissible in court has not prevented it from assuming an important role in criminal investigations.

Interviews with suspects:

inconsistent protocols;
not blind;
no formal records of the entire interview process

Interviews have commonly **not** been videotaped or transcribed, so accounts of what was said have been based on recollections; the conduct of interviews has also been variable. Nonetheless, law enforcement officials often use their "recollections" of what a suspect said during an interview. Claims of confessions or incriminating statements may thus have been in error. Information may also have been passed to the suspect that was then used to indicate intimate knowledge of the crime. The use of physical force during interviews was banned by the Supreme Court only by 1936.

Lab tests:

lack of blind

Allows fabrication of results (deliberate bias);
allows honest mistakes to persist and encourages over-interpretation of results.
see Innocence Project web site on fraud

Knowing which samples belong to which people allows fabrication of evidence (of which there have been many cases, see Innocence Project page (Fred Zain's string of fraud resulted in so

many convictions that prosecutors sought him out, as did the news program “60 minutes” when he was exposed). For honest technicians, the absence of blind encourages honest mistakes and selective replication of results (you repeat the test if the results don’t fit your preconceived ideas about guilt). And as revealed in the Castro Case (early DNA testing), the lack of blind causes people to overinterpret results and make them fit preconceived notions.

As an example of the absence of blind analysis by labs, here are two letters sent from the Chicago Police Department to the FBI, requesting DNA typing. All names are omitted from our text; where names were included in the letters, a description is given in square brackets [].

Letter 1: From Chicago Police Crime Lab to F.B.I. DNA Laboratory Division, 10 August, 1989

Dear [name of Director of F.B.I. lab],
I am writing to request that DNA typing be performed on several items of serological evidence. The names of the people involved are: [name of female victim] F/W (the victim) and [name of male suspect] M/B (the suspect). The evidence I am sending you consists of the following:

- Blood standard from [name of victim]
- Blood standard from [name of suspect]
- Extract from swab
- Extract from victim's pants
- Extract from victim's bra

All three of these extracts were found to be semen/spermatozoa positive and the two extracts from the clothing were found to have ABO, PGM and PEP A activity consistent with that of the suspect. I am also enclosing a copy of my laboratory report stating these results.
The facts of the case are that on 25 May 1989, the victim was grabbed from behind, pulled into the woods and sexually assaulted. The victim never got a good look at her offender and therefore is not able to make a positive I.D. of the suspect. The suspect [name] had just been released from the ILLINOIS DEPARTMENT OF CORRECTIONS after serving time for the same type of crime in the same area. At this time the suspect has not been charged.
Thank you very much for your assistance in this matter. Please feel free to contact me if you need more information.

Sincerely,
[name]
Criminalist II
Chicago Police Crime Lab

Letter 2: From Chief of Detective Division, Chicago Dept. of Police to F.B.I. DNA lab

Dear [name, Commanding Officer, F.B.I. DNA lab],
In early January, 1990, detectives assigned to the Chicago Police Department's Detective Division, Area Three Violent Crimes Unit were assigned to investigate the particularly brutal Aggravated Criminal Sexual Assault, Robbery and Kidnapping of one [name of victim], recorded under Chicago Police Department Records Division Number N-005025. On January 10, 1990, one [name

of suspect] M/N/31 years, FBI [#], C.P.D. Record Number [#], was arrested and charged with this and other offenses.
Blood and saliva samples of the offender and victim were obtained and tendered to Technician [name of technician] of the C.P.D. Criminalistics Unit. A sexual assault kit (Vitullo Kit) was also completed and submitted for the victim.
The undersigned requests that the recovered specimens and evidence be evaluated and subjected to DNA comparison testing. Although the offender has been identified and charged, we feel this comparison would greatly enhance the prosecution of [name of suspect], who was arrested after a week long crime spree.
If any additional information is needed, kindly contact Detective [name], star [#], Area Three Violent Crimes Unit, 3900 South California, Chicago, Illinois 60632, Telephone #(312)-744-8280, or the office of the undersigned.

Sincerely,
[name]
Detective Division
Room 501
1121 South State Street
Chicago, Illinois 60605

In addition to an absence of blind testing, no standards are included, which would offer quality control assurances as well as guard against sample mixup. As far as we know, these letters are typical. The law does not require blind testing or standards, and the prosecution units may not even recognize the possible consequences of omitting these design features.

The CJS is currently undergoing a lot of changes because of bad science. It can really help to know the scientific issues if you become a suspect, but even a good lawyer can't ensure your freedom just because you are innocent.

LOS ANGELES TIMES

Wednesday, October 9, 2002

A long-awaited research report released Tuesday by the National Academy of Sciences describes polygraph testing for national security screening as little more than junk science.

The United States' premier scientific organization said such tests, a key counterespionage tool for 50 years, are so inaccurate as to be counterproductive: promoting false confidence that spies and other national security threats have been ferreted out and likely producing false accusations against innocent people.

The academy's National Research Council prepared the report based on 19 months of study, and it confirms long-standing doubts about the validity of polygraph testing. A 1988 federal law bans the use of such tests for employment screening in most private businesses. Polygraph results are also inadmissible as evidence in nearly all state courts.

"If logic has anything to do with it, then the report will have a major policy impact," said Steven Aftergood, an intelligence analyst with the Federation of American Scientists. "I don't think

federal agencies stop and ask themselves how many spies have we caught with this, because the answer is `none,' or how many people have been unfairly denied employment, because the answer is `many.' "

Federal security agencies would not discuss the report's conclusions, saying they needed time to review the report in detail.

The federal government subjects thousands of job applicants and employees in sensitive positions to lie detector tests each year. The CIA and the National Security Agency give polygraph tests to all job applicants and employees. The FBI and the Defense Department also test extensively, particularly since Sept. 11. Such screening also is common at large police departments nationwide.

Linton Brooks, acting head of the National Nuclear Security Administration, said the agency will reassess its use of lie detectors in light of the new report.

A Defense Department spokesperson said that agency has valued the polygraph "as an investigative tool" for half a century but agrees that further research would be valuable.

Some experts say the wide-ranging and authoritative report could trigger profound changes in security practice.

"It is going to be a watershed" that shifts the burden of proof from polygraph skeptics to its advocates, said Paul Giannelli, a law professor at Case Western Reserve University and a consultant to the National Academy panel.

The study was underwritten by the Department of Energy, which was embarrassed by contradictory interpretations of polygraph tests given to Wen Ho Lee, who had been accused of being a nuclear spy. Lee was ultimately exonerated, and the furor over his case indirectly stimulated the polygraph study.

Which methods used by prosecutions against defendants have been found to have error rates of 10% or more in proficiency tests or experiments? (MTF)

A) DNA typing
B) Hair matching
C) Fingerprint matching
D) Lie detectors (polygraphs)
E) Eyewitness testimony

What does it mean if the random match probability (RMP) is 1/billion but the lab error rate is 1%? In other words, how should a jury deal with those two numbers in assessing the significance of a match. (one only)

(b) The significance calculation is not affected.

(c) The proper calculation is 1/billion x probability that no error was made (0.99)
(d) Or the proper calculation is 1/billion PLUS 1%
(e) You cannot calculate the significance of a match if there is a chance of human & technical error.