Witnesses, Fingerprints, and Statistics in Criminal Trials

Marcello Di Bello – mdibello@stanford.edu
Department of Philosophy – Stanford University

Conference on Law, Governance and Development
May 18-19, 2012
Azim Premji University, Bangalore, India

My topic is a comparison between statistical and traditional evidence in Anglo-American criminal trials. Prompted by studies showing that traditional forms of evidence, such as testimonies and fingerprints, are highly subjective and unreliable, legal scholars have explored ways to counter this problem, e.g. by adopting more demanding standards of admissibility or rules of weight. I will suggest that it is difficult to restrict the role of traditional forms of evidence in criminal trials without creating further complications. I will also discuss some of the dangers that statistical evidence poses for criminal justice.

Traditional evidence in decline. In the last thirty years, we have seen a steady attack against traditional forms of evidence in criminal trials, with the main targets being witness testimony and forensic techniques such as fingerprints. The testimony of a honest witness is the bedrock of the Anglo-American trial system and fingerprint identification has for long been considered the gold standard of criminal evidence. Yet, research in psychology and cognitive science has shown that these forms of evidence are often unreliable,[1] easily prone to manipulation,[2] and too much driven by subjective considerations and matters of context.[3] A rethinking of their role in criminal trials is therefore needed, and a few policy recommendations have already made their way in the legal literature.

In order to rethink the role of testimonial and fingerprint evidence, three strategies can be identified: (1) enhancing the training of police investigators, forensic experts, and legal professionals, so that the overall quality of testimonial and fingerprint evidence will improve, and at the same time, inadequate instances of them will be more easily screened off during pretrial proceedings;[4] (2) making the admissibility rules for traditional forms

[1] Even when subjects focus on a series of events, they often fail to notice important details. Subjects whose attention is focused on watching people play basketball might not notice that a giant gorilla has made its appearance several times. See Simons and Chabris (1999).
[2] On how how easily false facts can be inserted in a witness memory, see Loftus (1996).
[3] On how context influences fingerprint identification, see Dror et al. (2006). See also Zabell (2005).
[4] The American Academy of Science has recently issued recommendations about forensic techniques; see NRC (2009). To improve the quality of eyewitness evidence, see the recommendations in Wells et al. (2006).

of evidence more demanding;[5] (3) introducing rules of weight, e.g. a rule prohibiting that a criminal conviction be based on a single piece of uncorroborated eyewitness or fingerprint evidence.[6]

The first strategy should obviously be pursued, but the benefits may come into being only after a long process of retraining of the entire legal profession. The second strategy can be implemented more easily, although—were it be unaided by the first—it would trigger the unintended result of declaring too many pieces of traditional evidence inadmissible, thereby leaving the criminal justice system in a situation of sudden evidential vacuum. Finally, the third strategy would call for a significant change in the Anglo-American trial system which has been resistant to the idea of rules of weight, as it leaves the task of weighing the evidence to the fact-finders. Also, the last two strategies might render the issuing of a conviction too difficult, and consequently, too many guilty defendants might go free.

With the exception of the first strategy, the other two restrict the role of traditional forms of evidence in criminal trials in important ways, so that unless the latter are replaced by other forms of evidence, an evidential gap will ensue. Interestingly enough, in the last thirty years we have witnessed the rise of evidence based on statistical and probabilistic methods.[7] We may then feel tempted to seek comfort, so to say, in these new forms of evidence. The good news is that they can be more objective, precise, and scientific, and thus, they should fall prey less frequently of the shortcomings that affect traditional forms of evidence. But as I shall show, the bad news is that they are troublesome in ways that traditional forms of evidence aren't. Legal scholars and legislators are therefore forced to walk a tight rope. On one hand, the importance of traditional forms of evidence must be diminished; on the other, we should not uncritically hope that statistical and quantitative methods will take their place without any harm. How to strike the right balance is a complex question for which I have no answer, but I hope to have something to say about the criteria to construct an answer after I have explained in what way statistical evidence is problematic.[8]

The trouble with statistical evidence. To make my point vivid, it is useful to look at a British case that has stirred much controversy, *R v. Clark*

[5]The Supreme Court of New Jersey in *New Jersey v. Henderson* (August 2011) has recently commented on the limits of eyewitness testimony and it has recommended more stringent admissibility standards.

[6]See Thompson (2008), Sangero and Halpert (2007).

[7]See e.g.Thompson and Shumann (1987).

[8]Attacks against statistical evidence have been voiced repeatedly in the legal literature. The earliest attacks against statistical evidence and probability in the courtroom have been Tribe (1971) and Nesson (1979). More recently, see Wasserman (1991), Colyvan et al. (2001), Stein (2005), Allen and Pardo (2007), Ho (2008), and Pundik (2009). My criticism of statistical evidence, however, has been relatively little developed in the existing literature.

(1999). Sally Clark's children, a few months old, both died at the presence of their mother. The first baby died in 1996 and the second in 1998. Sally was arrested and charged with murder. At trial, the pediatrician Roy Meadow testified that the probability of two accidental consecutive cot deaths, in a family similar to the Clarks, was 1 in 73 million. These statistics were thought to be impressive enough to rule out the possibility that the two deaths occurred accidentally or because of natural causes.[9] Sally was convicted to life imprisonment in 1999 and the conviction was upheld on appeal in 2000. On a second appeal, in 2003, Sally was released on the ground that the statistics were flawed[10] and had been misinterpreted.[11]

The undeniable yet often overlooked moral is that statistical data should be supplied to the courts with extreme care and by high qualified professionals. But this is not the issue that interests me here, however important. My concern is on what to do when the statistics are sound and correctly interpreted. Suppose that good statistical data would render the probability of Sally Clark intentionally killing her babies extremely high. The question arises of whether such high probability would have been enough to justify a murder conviction against her—this is the sort of question I am concerned with here. My answer is that convictions based on only statistical evidence of the kind brought against Sally Clark should make us worry, even when the evidence is untainted by misinterpretation and unsound data.

The statistical evidence against Sally Clark lacked any significant degree of specificity and individualization. It did not provide any information about how Sally killed her babies; why she did it; when exactly the killing occurred; etc. Even at its best, it did not provide any information about who committed the killing. Since it only asserted that it was unlikely that the two babies died accidentally or by natural causes, no conclusion should have been drawn about the identity of the perpetrator.

As a result of the minimal degree of specificity and individualization of the evidence brought against her, Sally Clark could hardly defend herself. She only had two options: showing that the statistics were flawed or that other evidence exculpated her. Sally could not find other weighty evidence in her favour, and only on a second appeal, she managed to convince the court that the statistics against her were in fact flawed. After all, she was relatively lucky. If the statistics against her were not flawed, she

[9] This is a misinterpretation of the statistics, known as the *prosecutor's fallacy*. The figure "1 in 73 million" is the probability that two consecutive cot deaths occurred given the hypothesis that they were accidental, i.e. $\textit{1/73 million} = P(\textit{two coth deaths}|\textit{accidental death})$. This is different from the probability that the deaths were accidental given that they occurred, i.e. $P(\textit{accidental death}|\textit{two coth deaths})$.

[10] Roy Meadow assumed that the two deaths were independent events: he calculated the probability of two consecutive deaths by multiplying the probability of each. But the independence assumption is questionable, and it it is more likely that a first cot death raises the probability of a second one.

[11] See footnote 9.

would have had nearly no opportunity to defend herself, despite her possible innocence.

The reason why statistical evidence is troublesome is that it typically lacks any significant degree of specificity and detail. A prosecutor's case based on it will tend to be very vague and unspecified, although it might make the defendant's guilt highly probable. Confronted with an underspecified case, a defendant will have a hard time to respond to it. In contrast, if a prosecutor's case is richer in details about the crime, the defendant, especially if innocent, will more easily be able to attack it, by targeting the different propositions which the prosecutor committed himself to. When eyewitness and fingerprint evidence are used to incriminate defendants, more details about the crime will be usually available, so that defendants will have more avenues to rebut to the prosecutor's case; and this is why statistical evidence raises problems that traditional forms of evidence do not.

(To clarify my claim, in the appendix, I've added a hypothetical scenario which makes my argument more precise and general. I've also added a discussion of DNA evidence, which in my view cannot count as simply statistical evidence, so that my criticism of statistical evidence should not apply to it.)

Competing interests. To conclude, statistical evidence is troublesome not because it hampers the trial goal of error minimization or accuracy; on the contrary, the accuracy of criminal trials might be enhanced by a more extensive use of statistical evidence.[12] The trouble with it is that, when it is the only piece of incriminating evidence, it consigns defendants to unfair situations in which they have nearly no opportunity to exercise a defense. How much space we should accord to statistical evidence depends on how extensively we want to protect defendants' possibility to exercise a defense, and on how we intend to balance this protection with accuracy as a trial goal. A conflict between diverging values and interests is at stake here. On the one hand, we want to protect the interests of defendants *qua* individuals under trial, and it is intuitive enough that they should be given an opportunity to tell their own side of things, whether they are factually innocent or guilty. On the other hand, the pursuit of the long-term societal goals of error minimization and accuracy may be hampered if restrictions on the introduction of statistical evidence are put in place.

Further, if the restrictions in question were to affect both statistical and traditional evidence in criminal trials, we would face the risk of an unfilled evidential gap. In order to assess how harmful and problematic this gap will be, an empirical research is needed which would show how many wrongful convictions have resulted because of unreliable traditional evidence and how often statistical evidence would be used and needed in

[12] See Koehler and Shaviro (1990).

criminal trials.[13] Until empirical data are available, specific policy recommendations would be premature.

Appendix 1: A hypothetical scenario. The hypothetical I want to consider was originally invented by Charles Nesson[14] and it is here reported in a slightly modified version:

> *Prisoner.* One hundred prisoners are in a prison yard. Suddenly, 99 of them execute an attack against a guard on duty, who is brutally murdered. The rate of participation in the murder is 99:100, as attested by a video recording showing that all prisoners, except one, actively and deliberately participated in the murder.[15] The innocent prisoner, however, cannot be identified from the recording and no other evidence is available about the murder.[16] After the fact, all prisoners manage to escape except one, call him Eschaton, who is captured and tried for murder.
>
> The prosecutor argues that, given the 99:100 rate of participation, Eschaton is guilty of murder with a 0.99 probability. Only one prisoner did not participate in the murder, and thus there is a 0.01 probability that Eschaton did not participate and a 0.99 probability that he did. Further, as the criminal standard of proof 'beyond any reasonable doubt' is quantifiable as requiring to prove guilt with a high probability, it would seem that the standard is met, or at least so the prosecutor wishes to argue. But despite the high probability of guilt, many of us feel that we are not in a position to convict Eschaton.

The prosecutor offered a very minimalist theory of the crime. He refrained from establishing who participated in the murder and who did not; he only established that 99 prisoners participated and one did not. Such theory allows for many different possibilities without any commitment to one in particular. One possibility is that all prisoners except, say, prisoner 1 participated; another possibility is that the exception was prisoner 2; or prisoner 3; or prisoner 4; and so on. The prosecutor's theory allows for many possibilities to have occurred but not the one in which

[13] Data about the causes of wrongful conviction are partly available in the United States through the Innocence Project; see *www.innocenceproject.org*. To my knowledge, there are no official statistics or empirical data detailing how often statistical evidence is currently used in criminal trials.

[14] See Nesson (1979).

[15] We can suppose that the video recording is extremely detailed as to the ways in which the 99 prisoners killed the guard; the only unanswered question concerns the identify of the 99 guilty prisoners.

[16] To make this assumption plausible, suppose that the 99 prisoners appeared in the video while giving their shoulders to the camera, so that their faces were unrecognizable.

Eschaton is innocent. By leaving open so many possibilities, the prosecutor managed to establish his theory with a high degree of probability; his theory is, after all, 0.99 probable, conditional on the statistical evidence.[17]

Since the prosecutor's theory is compatible with 99 possibilities, Eschaton will have to disprove each single possibility. He will have to establish that the innocent prisoner was him, or almost equivalently, that no one of the other 99 prisoners was innocent. This would place a very high burden of proof on Eschaton. The burden would be less demanding if the prosecutor had left open less possibilities, or if he had committed himself to one possibility only.

Appendix 2: DNA evidence. DNA, fingerprint and eyewitness evidence all obey a similar pattern. Suppose that the perpetrator possesses certain facial, genetic, or fingerprint features, and that during the crime he left traces, either physically or in the witness memory. When the investigators single out a suspect and want to know whether he is the perpetrator or not, DNA, fingerprint, and eyewitness evidence can help them by means of what we may call *matching*: the facial features which the witness remembers to belong to the perpetrator might match or not those of the suspect; the fingerprint traces left on the crime scene might match or not those of the suspect; finally, the DNA traces on the crime scene might match or not match with the DNA profile of the suspect. If a match is declared, the investigators will have powerful incriminating evidence.

Further, DNA evidence is a form of trace evidence similar to fingerprints. Instead of depending on finger traces, it relies on traces such as hair, saliva, semen, etc. from which a DNA profile can be extracted, in a way analogous to how a fingerprint profile can be extracted from fingerprint traces. As long as fingerprint evidence is individualized, so is DNA evidence. It is also similar to eyewitness evidence: the latter relies on the cognitive traces in the eyewitness memory left by the sight of the perpetrator, while the former relies on physical traces.[18]

If the above is correct, DNA evidence must count as individualized as much as fingerprint and eyewitness evidence do if not more. But why then so much emphasis on its statistical nature? DNA evidence has a very striking statistical component, i.e. the frequency of the genetic features or DNA profile used to declare the match. DNA evidence is the only evidence for which those frequencies have been established in a dependable

[17] Given the 100 possibilities of how the collective murder could have happened, the prosecutor has left open 99 of them, so his theory is 0.99 probable.

[18] The difference between trace evidence (fingerprint or DNA) and eyewitness evidence is that, for the former, a match is declared by a forensic or laboratory analyst. This process is in principle publicly accessible to the fact-finders. Instead, when an eyewitness identifies a suspect as the perpetrator, the fact-finders are only left with the testimonial statement "it's him" or "it's not him." They have no access to the cognitive process through which the witness decided that the suspect was the perpetrator or not, yet this should not obscure the fact that some process of matching must go on implicitly inside the witness mind.

way, whereas the frequency of fingerprint or of facial features are virtually unknown.[19] Genetic frequencies have also been the center of wide controversies, and this might have led many to forget that DNA evidence is a form of trace evidence, just like fingerprints.

That DNA evidence counts as individualized is apparent by how a defendant can rebut to it. A prosecutor's case based on DNA evidence will be sufficiently detailed. For it will contain information about when the perpetrator left the DNA trace at the crime scene; what actions he was performing which caused the DNA trace to be deposited; etc. The prosecutor, in other words, will have at least to establish that the presence of such DNA trace excludes any innocent behavior by whoever left the trace. Confronted with a relatively detailed theory of the crime, defendants' avenues for a defense will be analogous to those afforded to them when they are incriminated by means of fingerprint or eyewitness testimony.

[19]The frequencies of fingerprints and facial features are not discussed at all. This statistical aspect of fingerprint and eyewitness testimony is ignored completely. This is unfortunate. Two individuals might have the same fingerprints or might look exactly alike. No features are unique, or at least, science and statistics are unable to prove uniqueness. This means that the shortcomings of DNA evidence apply to fingerprint and witness evidence, with the difference that for DNA evidence we have reliable statistics concerning the frequency of genetic features, while we do not have any such statistics concerning facial-features and fingerprint patterns.

REFERENCES

Ronald J. Allen and Michael S. Pardo. The problematic value of mathematical models of evidence. *Journal of Legal Studies*, 36(1):107–140, 2007.

Mark Colyvan, Helen M. Regan, and Scott Ferson. Is it a crime to belong to a reference class? *Journal of Political Philosophy*, 9(2):168–181, 2001.

Itiel E. Dror, David Charlton, and Ailsa E. Peron. Contextual information renders experts vulnerable to making erroneous identifications. *Forensic Science International*, 156:74–78, 2006.

Hock Lai Ho. *Philosophy of Evidence Law*. Oxford University Press, 2008.

Jonathan J. Koehler and Daniel N. Shaviro. Veridical verdicts: Increasing verdict accuracy through the use of overtly probabilistic evidence and methods. *Cornell Law Review*, 75:247–279, 1990.

Elizabeth F. Loftus. *Eyewitness Testimony (revised edition)*. Harvard University Press, 1996.

Charles R. Nesson. Reasonable doubt and permissive inferences: The value of complexity. *Harvard Law Review*, 92(6):1187–1225, 1979.

NRC. *Strengthening Forensic Science in the United States: A Path Forward*. The National Academies Press, 2009.

Amit Pundik. *Statistical Evidence: In search of a Principle*. PhD thesis, University of Oxford, Faculty of Law, 2009.

Boaz Sangero and Mordechai Halpert. Why a conviction should not be based on a single piece of evidence: A proposal for reform. *Jurimetrics Journal*, 48, 2007.

Daniel J. Simons and CHristopher F. Chabris. Gorillas in our minds: Sustained inattention blidness for dynamic events. *Perception*, 28:1059–1074, 1999.

Alex Stein. *Foundations of Evidence Law*. Oxford University Press, 2005.

Sandra Guerra Thompson. Beyond a reasonable doubt? reconsidering uncorroborated eyewitness identification testimony. *UC Davis Law Review*, 41:1487–1545, 2008.

William C. Thompson and Edward L. Shumann. Interpretation of statistical evidence in criminal trials: The prosecutor's fallacy and the defense attorney's fallacy. *Law and Human Behaviour*, 11:167–187, 1987.

Laurence H. Tribe. Trial by mathematics: Precision and ritual in the legal process. *Harvard Law Review*, 84:1329–1393, 1971.

David T. Wasserman. The morality of statistical proof and the risk of mistaken liability. *Cardozo Law Review*, 13:935–976, 1991.

Gary L. Wells, Amina Memon, and Steven D. Penrod. Eyewitness evidence: Improving its probative value. *Psychological Science in the Public Interest*, 7(2):45–75, 2006.

Sandy L. Zabell. Fingerprint evidence. *Journal of Law and Policy*, 13:143–179, 2005.