"Science,""common sense," and DNA evidence: a legal controversy about the public understanding of science

Michael Lynch and Ruth McNally

Public Understanding of Science 2003; 12; 83
DOI: 10.1177/0963662503012001246

Public Understand. Sci. **12** (2003) 83–103

"Science," "common sense," and DNA evidence: a legal controversy about the public understanding of science

Michael Lynch and Ruth McNally

Law courts are important institutional settings in which public understanding of science is problematic. Courts have struggled with the question of how to handle scientific evidence in a system of justice in which lay jurors are responsible for deciding the facts of the case. Judicial conceptions of science and of jurors' capacities to understand scientific evidence inform decisions in particular cases. Such decisions, in turn, act as precedents that, for better or worse, embed judicial conceptions of public understanding of science into the workings of the legal system. This paper examines an English case in which the difference between "scientific" and "common sense" evidence was explicitly at stake. In this case *Regina v. Adams*, DNA evidence was used to convict the defendant of rape. The principal item of prosecutorial evidence was a match between DNA profiles developed from the defendant's blood and from semen recovered at the crime scene. Prosecution experts expressed the evidence in probabilistic terms, estimating a probability of one in 200 million that a DNA profile from a randomly chosen unrelated man in the relevant population would match the profile developed from the crime stain. Other so-called "common sense" evidence supported the defendant's not guilty plea. The case was appealed twice in the mid-1990s as the Court of Appeal deliberated over an innovative effort by the defense to counteract the prosecution evidence by converting the non-DNA "common sense" evidence into probability estimates. In its decisions, the Court reinstated a boundary between the "scientific" and "common sense" evidence, arguing that this boundary was necessary to preserve the jury's role as trier of fact. The paper's discussion of the court's boundary work addresses unresolved problems with the place of probability estimates in jury trials.

1. Introduction

On 6 April 1991, a young woman identified in court documents as Miss M was walking home in the early morning hours in a town north of London.[1] A stranger approached her and asked for the time. According to Miss M's testimony, when she glanced at her watch, the stranger grabbed her from behind, overpowered her, and raped her. She saw the man's face

 [0963-6625(200301)12:1;83–103;033246]

for only a few seconds.[2] Miss M reported the attack to the police and described her assailant as a white, clean-shaven man with a local accent, who was between 20 and 25 years old. Miss M underwent a physical examination, during which the police took a vaginal swab. Forensic analysis revealed traces of semen on the swab, and a DNA profile was developed. At the time, the Metropolitan Police had compiled a database consisting of a few thousand DNA profiles from unsolved crimes and convicted criminals. The profile developed from the semen recovered during Miss M's examination did not match any of the profiles on the database and was stored for future reference.

Two years later, Denis John Adams was arrested in connection with another sexual offence. The police took a blood sample from him, and forensic scientists developed a DNA profile from the sample. Forensic investigators ran Adams's DNA profile against the unsolved crime samples on the police database. According to the investigators, his profile matched the evidence recovered during the investigation of Miss M's rape, and so he was charged with that crime. According to the judicial summary of the case, the matching DNA profiles provided the only substantial evidence against Adams, and all of the other evidence supported his defense. The DNA evidence proved persuasive to the jury, and Adams was convicted in January 1995. In 1996, he successfully appealed the conviction and was retried. Again he was convicted, and again he appealed. The second appeal was heard in October 1997. This time the court rejected the appeal.

This paper treats *R. v. Adams* as a case study in the public understanding of science. The materials discussed in the paper are drawn from judicial summaries of the *Adams* trials and appeals, supplemented by interviews with a key expert witness for the defense and a prominent forensic analyst who worked for the prosecution. Our analysis of the particular case makes use of a larger body of documents, interviews, and observations from a long-term study of the forensic procedures, legal processes, and expert authority in cases involving DNA profile evidence. As readers who are trained in legal scholarship will recognize, this paper is not an account of how evidence law, admissibility standards, grounds of appeal, and procedural constraints apply to the case. Further, the materials used in the case are restricted to retrospective summaries and interviews, as well as observations made during the second appeal hearing.[3] We are interested in the case because of the way it explicitly addressed, and provisionally resolved, a distinction that has broad interest in social theory and social studies of science.[4] This is the distinction between science and common sense.

2. Background to the Trials

The *Adams* trials and appeals were concerned, in part, with the presentation of "scientific" and "common sense" evidence to a jury. We place quotation marks around "science" and "common sense" for two reasons: first, because the distinction marked by those terms is problematic, and second, because the terms explicitly appear in the Appeal Court summary of the case and are significant for the way the Court resolves the case. Consequently, while we would be hard pressed to defend a conceptual distinction between science and common sense, we shall treat the Court's *performative* use of the distinction as a "datum" for this study.[5] In the present case, the Appeal Court does more than *discuss* the relationship between "scientific" and "common sense" evidence, it stipulates how that relationship should be organized in the case at hand and other cases like it. What the judges say, therefore, serves to *effectuate*, within specific institutional circumstances, a particular set of

procedures related to the general theme of "public understanding of science." A further aspect of the judicial summary that is important to keep in mind is that it was written in the face of contested claims about the relationship between scientific evidence and public (juror) understanding of such evidence. Within the circumstances of the case, the judicial performance effectuated a resolution of a controversy about how the public can, and should, understand and use expert evidence.

What interests us in this paper is the way the Court deploys the distinction to mark out domains of competence and legitimate authority in a significant arena of public discourse. In other words, we are interested in how the Court marked, and practically managed, a contested boundary in the public understanding of science. The salience of the distinction for our discussion of the Court's "boundary work" is independent of the question of whether the distinction accurately "maps" the kinds of evidence deployed in the case.[6]

The *Adams* case involved an intersection of common law, genetic science, and probability theory. It is one of a series of cases in which courts in the United States, United Kingdom, and other justice systems struggled to incorporate DNA evidence into a system of justice that stresses lay participation and public accountability. The basic problem is that jurors have limited acquaintance with, and very little opportunity to learn about, the technical matters that experts are asked to present in court. Consequently, jurors (as well as many judges, prosecutors, and defense attorneys) can be at the mercy of experts who make uncontested assertions about what the evidence shows.[7] Moreover, on occasions when the experts disagree, the lay participants have no technical basis for deciding which claims to believe. Legal scholars and social scientists disagree about the extent to which juries are capable of understanding expert evidence. Many argue that judges and juries are ill prepared to distinguish genuine expertise from "junk science." However, others argue that the difficulties juries experience with expert evidence can be relieved by more effective presentation of such evidence in court. According to this view, juries are capable (or, at least, no less capable than judges are) of grasping what they need to know to decide a case.[8]

In addition to empirical questions about what jurors understand, or can be brought to understand, is a broader set of legal and political questions about the jury's role in the trial court. Although jury trials occur in a small, and decreasing, proportion of criminal and civil cases in the Anglo-American courts, the jury continues to stand proxy for the common citizen's place in the justice system.[9] Moreover, even in trials without a jury, the courts assign to the judge the role of "trier of fact" and circumscribe the role of expert witnesses. Legal precedents and procedures reserve the "ultimate issue"—the judgment about guilt or liability—to the trier of fact and restrict the role of experts to "informing" the court. Procedures of cross-examination and jury deliberation build in social-interactional and documentary mechanisms for "testing" evidentiary claims, exposing and dramatizing inconsistencies, and airing opposing views. In traditional jurisprudence, such procedures for openly presenting and testing evidence before a lay audience are likened to powerful machineries for exposing "the truth."[10]

The section that follows presents a brief summary of the techniques used to analyze forensic evidence in the *Adams* case. Then, after going into detail about the evidence and arguments used in the trials and appeals, the paper focuses on the Appeal Court's judgment about the proper relationship between "scientific" and "common sense" evidence. The paper concludes with a discussion of how the case exemplifies a problematic relationship between technical, probabilistic evidence and "common sense" understandings of the totality of evidence bearing on the case at hand.

The technique: single-locus probes

In the United States, United Kingdom, and many other nations, the laboratory routines and invisible entities of molecular biology have become routine tools for criminal investigation. DNA profiling (also called DNA testing, DNA typing, and DNA fingerprinting) involves a series of molecular biological techniques that have been used in criminal investigations starting in the mid-1980s. Technical terms like multi-locus probe (MLP), single-locus probe (SLP) and short-tandem repeat (STR) sequence analysis distinguish some of the different DNA profiling techniques that have been developed since 1985. Like other forensic applications of molecular biology, these techniques are comprised of a series of established laboratory routines for extracting DNA and visualizing DNA patterns. Although many of these routines are common to other fields of practice, they have been adapted specifically for purposes of criminal identification.

The SLP technique was used by the Metropolitan Police to analyze the samples in the *Adams* case.[11] The SLP technique produces a visual document—an autoradiograph—by using radioactive "probes" to mark the relative length (or molecular weight) of specific, highly-variable regions of DNA isolated from a sample. SLP does not provide a complete, or even partial, record of a person's DNA code. Instead of reading "the book of life," the SLP technique makes visible genetic variations (alleles) at specific chromosomal sites. Allelic variation at these sites is in terms of length (the length of the DNA sequence). Each probe is designed to visualize the different alleles that can occupy the same chromosomal site. The size of the alleles at that site is indicated by the position of one or two bands on an autoradiograph (one band if both alleles from each of the chromosome pair are the same size as each other (homozygous), two if they differ (heterozygous)). Together, a series of four, six, or more probes visualizes a pattern of bands that makes up the DNA profile for a given sample.

Forensic examiners compare profiles developed from crime scene samples, suspect samples, and control samples.[12] Six SLP probes will produce an SLP profile comprising between six and twelve bands, the position of each band being an indication of a different allele. When, for example, the pattern of bands developed from a suspect's blood sample does not align with the bands from a semen sample recovered during the investigation of a rape, the suspect is supposed to be excluded from further criminal consideration. When the two profiles match, i.e., when the bands all align with each other, it is evidence, though not certain proof, of identity. What remains to be determined is the probability that a given SLP profile occurs in the relevant population group or subgroup. This is a key difference between DNA "fingerprinting" and the established technique of friction ridge analysis (fingerprinting). The doctrine that each individual has unique fingerprints has been established for nearly a century. Accordingly, fingerprint examiners are not required to give probability estimates, but only to declare whether a latent print (a fingerprint recovered from a crime scene) and a suspect's print match do or do not match (an examiner can also testify that the evidence is inconclusive).[13]

Unlike fingerprint examiners, expert witnesses who present DNA matches are required to estimate the probability that two randomly selected, unrelated people from the same population group or "racial" subgroup would share the same set of alleles and thus have the same DNA profile. Forensic organizations, such as the United States Federal Bureau of Investigation, and the Forensic Science Service of England, Wales, and Northern Ireland, have developed estimates of how frequently each of the alleles in a given SLP profile occurs in a national population and/or its "racial" subgroups. Some alleles are extremely rare, occurring in less than 1 percent of the human population, while others are more common. By

consulting reference figures for the frequency of each allele in a given profile, and multiplying together the frequencies of all of the alleles in the profile, forensic analysts can come up with extremely low probability figures for a given profile, figures that indicate astronomical odds against finding random matches between profiles from unrelated individuals. Because DNA patterns are heritable, identical twins are expected to have identical profiles, and the probability of matching profiles should be higher for related individuals than for unrelated individuals.

The evidence

The DNA profile of the semen sample recovered from the rape victim, Miss M, was stored on a computer database. The database was a relatively small one that was used in the early 1990s.[14] The early database was held at the Metropolitan Police Forensic Science Laboratory in Lambeth, South London. It was a computerized index of SLP profiles developed from a few thousand persons convicted of crimes, and it also included a smaller number of body fluid stains associated with unsolved crimes. The vast majority of these crime stains were semen samples associated with sexual assaults.[15] Forensic scientists were able to check profiles from new suspects against the convict profiles and crime stain profiles in the database. This was how forensic employees at the Metropolitan Police laboratory were able to find a match between the DNA profile from Denis John Adams's blood sample and the profile from the semen stain collected during the investigation of Miss M's rape.

The case summary described Denis John Adams as a Caucasian man, who was 37 years old at the time of his arrest. As noted earlier, Miss M described her assailant as a Caucasian man between 20 and 25 years old. After Adams was arrested, he was placed in an identification parade (a police lineup), but Miss M. failed to identify him or anyone else in the lineup as her attacker. When Miss M saw Adams at the committal proceedings, she said he did not look like her attacker. According to her, Adams appeared to be much older than the attacker. The prosecution proceeded with the case in spite of the lack of other evidence linking Adams to the crime. At the trial, Adams gave alibi evidence, claiming he spent the entire night with his girlfriend on the date of the attack. His girlfriend supported his alibi.

The *Adams* case was interesting because virtually all of the evidence other than the DNA profile match supported the defendant's innocence. According to the court's summary, the "scientific" evidence supported the prosecution, whereas all of the "common sense" evidence supported the defense. The court may have overstated the discrepancy, because some of the circumstantial evidence supported (or, at least, did not conflict with) the prosecution's case. Adams lived in the local area, and he was Caucasian. And while he differed in some respects from the victim's description, it is possible to discount the discrepancies and the failure of the victim to identify him. Eyewitness testimony is notoriously fallible, and in this case the victim was asked to identify an assailant she had seen only briefly during a traumatic encounter two years earlier. Had Adams been, say, an eighty-year-old man of Asian extraction, the discrepancies would have been far more difficult for the prosecution's case. It also is not too difficult to discount Adams's alibi and the girlfriend's corroboration of it. Nevertheless, the *Adams* trial was construed by the court as a kind of test case in which a single powerful item of "scientific" evidence was weighed against an array of items of "common sense" evidence.

For our purposes, the case exhibited a special kind of "reality disjuncture": a dramatic confrontation between incommensurable accounts of "the same" event. In many rape cases, there is a disjuncture between the accused man's and the victim's accounts of the alleged event and the circumstances leading to it.[16] For example, when pursuing a "consent"

defense, the accused rapist argues that the victim explicitly or implicitly consented to the sexual act. In the absence of corroborating evidence, the jury is faced with choosing between disjunctive narratives. In the *Adams* case, according to the Appeal Court's summaries, the evidence supporting the accused rapist's guilt not only conflicted with the evidence supporting his innocence, it was a systematically different *kind* of evidence. The Court's summaries propose that the prosecution relied on expert evidence, whereas the defense relied on ordinary forms of identification and description.[17] This is only one case, but it is interesting for analytical purposes because of the unusually clear way it juxtaposed the credibility and weight of expert evidence, and specifically of DNA profile evidence, with that of other, more ordinary, forms of evidence. The Court summary aligned the two sides of the case with a set of epistemic distinctions.[18]

Prosecution	Defense
Expert judgments	Common sense judgments
Probabilistic evidence	Non-probabilistic evidence

3. The First Trial

During the first trial, Adams was represented by Mr. Ronald Thwaites, QC, a prominent barrister with strong civil libertarian leanings.[19] During this trial, and throughout the subsequent trials and appeals, the defense used an unusual two-pronged attack. As a first prong of attack, Thwaites and his main expert witness challenged the impressive, seemingly unassailable, probability figures the prosecution assigned to the DNA evidence. This form of challenge was not unusual, as the basis for such probability estimates had been challenged in many prior cases. The second prong involved an unusual, if not unique, attempt to employ a Bayesian approach to quantifying the weight of the "common sense" as well as the "expert" evidence. The two prongs of attack reinforced each other.

Assigning probabilities to the match

The match between the DNA profiles from the vaginal swab and from Adams's blood sample was translated into probabilistic language by the forensic scientists who summarized the evidence for the court. The scientists used different types of probabilistic statement at different times, when the evidence was analyzed, challenged, and reanalyzed during the trials and appeals. Some evidence statements used non-numerical terms to describe the DNA match probability.[20] In a statement dated 10 December 1993, Dr. Harris, a forensic scientist who analyzed the DNA evidence for the prosecution, stated "There is very strong evidence that Denis John Adams (the appellant) has shed this semen." Another forensic scientist, Ms. Lygo, in a statement on 17 November 1994, made a similar, though more formal, statement: "The scientific evidence provides strong support for the allegation that the semen staining on the high vaginal swab (ZH/11) originated from Denis John Adams." A third forensic scientist, Mr. Lambert, on 21 December 1994, also made a non-quantified assessment, while referring to the results of two different methods of analysis: ". . . the DNA SLP, and STR profiling results, when considered together, provide very strong support for the [prosecution's] view. . .."[21]

The defense did not take issue with these non-quantified statements about the probative value of the DNA evidence. The main points of contention concerned the use of quantitative probability estimates to formulate the "very strong support" for the prosecution's case. Perhaps because of the way the Crown's case against Adams concentrated on the matching

DNA profiles, the forensic scientists who analyzed the evidence sought to maximize the weight of that evidence by combining a series of probes and quantitative estimates. According to the testimony of Dr. Harris, an SLP technique with four probes was used and seven matching bands were visible in a comparison between the DNA profile of Adams's blood sample and the profile from the vaginal swab.[22] In his testimony, Harris acknowledged that when he examined the autoradiograph, one of the matching bands was so faint that he highlighted it with a marking pen in order to make it more clearly visible. He included the band in his estimate of the probability of the profile's combination of alleles in the white European population, and he came up with a figure of one in two million. Dr. Harris also said that he later added a fifth probe, which produced two additional matching bands. The revised probability estimate based on the nine matching bands became one in 200 million. This figure was rounded off from a more precise figure of one in 297 million, which Dr. Harris estimated was the probability that a randomly chosen unrelated man in the white European population would have a DNA profile that matched the DNA profile developed from the crime stain. He justified the rounded number in the interest of "conservatism": to avoid exaggerating the weight of the evidence against the defendant and to make the figures easier for the jury to understand.

Mr. Thwaites attacked the probability estimate by suggesting that Dr. Harris had engaged in an effort at "improving the numbers" to the point that "the probes became weaker and the technology was exhausted."[23] In other words, Thwaites accused the prosecution of stretching the technique beyond its limit of resolution in order to generate probability estimates that would persuade the jury to convict Adams on the DNA evidence alone.[24] Thwaites and his expert witnesses argued that analysis with the fifth probe was not performed under the same conditions as the earlier analysis, thus invalidating the enhancement of the probability figure from one in two million to one in 200 million. They also challenged Harris's procedure of manually enhancing the evidence. They argued that Dr. Harris's manual mark was illegitimate, and that the particular "match" should be excluded from the probability estimate. According to the multiplication rule, this would alter the estimate by a factor of ten, changing the estimate from one in two million to one in 200,000.

In addition to using different versions of the SLP technique, the forensic scientists also used STR, a more recently developed technique. This is a DNA profiling method that employs the polymerase chain reaction (PCR) together with other laboratory techniques and instruments to analyze "short tandem repeat" (STR) sequences of DNA. The STR technique was adopted by forensic services in the UK and elsewhere, and became the basis for the national DNA profile database, which was developed by the UK Forensic Science Service (the first profiles were placed on the STR database in 1995). In the first *Adams* trial, the judge decided to exclude the STR evidence. Had the evidence been presented to the jury, and had the resulting estimate been numerically combined with the SLP estimate (a procedure that the defense was prepared to criticize), the resulting probability figures would have been even more impressive. The decision not to present the STR evidence to the jury had to do with avoiding potential confusion arising from the combination of different techniques.

To complicate matters further, the defense contended ("somewhat faintly" according to a judicial summary) that the analysis should take account of the fact that Adams had a brother.[25] The DNA experts acknowledged that the brother "was a complicating factor whose existence reduced the probability of a match to 1 in 220."[26] The court chose to disregard this possible complication, because the brother's blood was not profiled and the defense did not suggest that he might have been involved in the crime.[27]

Without going more deeply into the matter, it should be clear that the question of just how to formulate the match probability was contentious. Depending upon the number of probes used, and upon which visual results were counted as adequate, the probability estimates varied from one in 200 million to one in 200,000. If the brother was included in the narrative, the number became one in 220. Given these discrepancies, the rounding off of 297 million so that it becomes 200 million would seem a minor adjustment (although this too was contested by the defense).

In spite of the variety of numerical and non-numerical formulations that were given for the estimate, the figure of one in 200 million was used throughout the *Adams* trials and appeals. The defense continued to contest this figure but also used it for convenience.

The Bayesian approach.

A distinctive argumentative strategy was used during the *Adams* trials and appeals. This strategy involved an effort to enable the jury to use a Bayesian procedure for translating all of the evidence into comparable probability estimates. The defense appealed to "elementary fairness" when arguing that all of the evidence should be given a symmetrical treatment. According to the argument, the "common sense" evidence presented by the defense should be weighed on the same scale as the "scientific" evidence presented by the prosecution, and both should be analyzed with the Bayesian formula: posterior odds = prior odds $\times$ likelihood ratio. The defense argued:

> Unless the defence-oriented evidence could be successfully translated into numerical form and combined by the jury with the prosecution statistical estimate, the defence were prejudiced because of the tendency of the single large statistic to overwhelm the other evidence. The jury were in effect given only one half of the equation: the prosecution reasoning being how unlikely he is to be innocent, given the finding in relation to the DNA. But the defence were attempting to say how unlikely he is to be guilty, given the rest of the evidence, if it could be put to the same form to enable to jury to complete the equation.[28]

The defense emphasized that Bayes's Theorem already was incorporated into the "practices and procedures used by the FSS in their out-of-court preparation for the presentation of DNA evidence in court."[29] According to this argument, the prosecution had already opened the door to the Bayesian approach. The only question was whether the door should also be kept open for the defense evidence.

It might seem at this point that there is a clear distinction between the quantitative DNA evidence presented by the prosecution and the qualitative "common sense" evidence presented by the defense, but this is exactly the distinction that the defense challenged. Thwaites and his expert witnesses argued that forensic scientists use empirical data from studies of reference populations when making estimates of the probability of a random match, but that the "single large statistic" they develop for a specific case depends on what the scientists deem relevant to the case at hand. As we have seen, in the *Adams* case, the one in 200 million estimate developed by the forensic analysts was based on reference figures for the European Caucasian population, and it did not take account of the possibility that Adams's brother might be included in the suspect pool. The figure also incorporated contested judgments about which bands on an autoradiograph were clear enough to declare as evidence, and how much rounding-down was appropriate. The derivation and preservation of the "single large statistic" of one in 200 million thus depended on a series of decisions about relevant evidence, adequate procedure, and reasonable estimation. In brief,

the defense argued that the prosecution's quantitative evidence was based upon a series of judgments about the case at hand; judgments that, within the specialized domain of forensic analysis, were not essentially different from the "qualitative common sense" judgments that entered into the weighing of the defense evidence. Consequently, the defense proposed to extend quantitative methods of estimation to cover the relevance and weight of the alibi, the misidentification, and other evidence in the case.

An expert witness for the defense who argued in favor of the Bayesian approach was Professor Peter Donnelly, of Queen Mary and Westfield College, University of London. Donnelly, who later became Head of Statistical Science in the Department of Statistics at Oxford, argued that Bayes's Theorem was "the only logical and consistent way" for the jury to consider the DNA evidence together with the other evidence in the case. Donnelly's intervention was more than a matter of registering an expert opinion. He also developed a questionnaire to be used by the members of the jury. The prosecution did not object to the procedure, and a prominent forensic scientist working with the prosecution, Dr. Ian Evett, reviewed the questionnaire before it was introduced into the trial. The judge ruled that it would be acceptable for individual jurors to use on a voluntary basis.

The questionnaire

The questionnaire consisted of a series of seven questions, each of which asked an individual juror to make probability estimates concerning four items of evidence: (1) that a local man would have committed the offense, (2) that the victim would not have identified the defendant, (3) that Adams would have given the evidence he did in favor of his innocence, and (4) that Adams would have been able to call alibi evidence on his behalf. With the exception of the first item, which Donnelly said was a "starting point," there were two questions for each item. Donnelly did not tell the jurors which figures to use for their estimates, but he did supply some illustrative figures when he demonstrated the procedure. Here is how he worked through the series of items:

1. What is the chance, assuming nothing else about the case, that the rapist came from the local area? Donnelly cited local council data that 153,000 men between ages of 18 and 60 lived within a 15-kilometer radius of the crime scene. For purposes of calculation, he rounded off this figure to 150,000. He then gave an illustrative estimate of a 75 percent chance that the attacker was a local man. Donnelly then divided 150,000 by 0.75, resulting in a figure of 200,000. According to the undisputed facts of the case, Adams was a local man so that *in the absence of any other evidence* the odds of his being the man who "shed" the semen sample recovered during the victim's examination were 200,000 to one. Donnelly explained this procedure as follows:

> The . . . rationale is that if we were certain it was a local man and there are 153,000 relevant local men, the fact that it is a particular one, Mr. Adams, we would assess as 1 over 150,000. If we are not certain it is a local man, our assessment for a particular local man will be decreased a bit. The way that changes is that you take your 150,000 which are the odds, 150,000 to 1—are the odds on innocence—and you increase that by 100 divided by the percentage, 75 percent. . . . We have now changed to a situation where our odds on a particular local man, Mr. Adams, is the one of interest, but a particular local man being the true rapist is 200,000 to one now.[30]

2. What is the chance that the victim would fail to identify Adams? The questionnaire contained two questions relevant to this item. One question asked the jurors to state the

probability that, if Adams were innocent, he would not match the victim's description. The second question asked jurors to state the probability that, if Adams were guilty, he would not match the victim's description.

> What matters is the ratio of those two figures. . . how much more likely one thinks the evidence is if Mr. Adams is guilty than if he is innocent, or how much less likely?[31]

For illustrative purposes, Donnelly estimated that if Adams were innocent there is a 90 percent chance that he would not match the victim's description, but if he were guilty there is a 10 percent chance that he would not match the description. The ratio between the two hypothetical estimates would thus be 1:9. According to the Bayesian logic of the questionnaire, this ratio can be multiplied by the ratio given for the first item: 1:200,000 multiplied by 1:9. This results in a probability of one in 1.8 million. When asked by the judge, ". . . does that represent the chance that Adams is guilty?" Donnelly answered "Yes." Note that this answer pertained only to the first two items of evidence, without taking into account any other evidence in the case. This procedure was then carried through the remaining two evidence items.

3. What is the chance of the defendant's evidence? In his illustrative calculation, Donnelly estimated that there was an equal chance that a guilty or innocent defendant would give the same evidence that Adams provided the court on his own behalf. In Donnelly's illustration, this item was "neutral," represented by a 1:1 ratio, with no effect on the calculations from the first two items.

4. What is the chance of the alibi evidence? In answer to the first question on this item, Donnelly estimated a 25 percent chance that a guilty defendant would be able to produce the alibi evidence in this case, and in answer to the second question, he estimated a 50 percent chance that an innocent defendant would produce the same evidence. Consequently, according to this illustration, it would be twice as likely that Adams would have the alibi evidence if he were innocent than if he were guilty. Multiplying the ratio 1:1.8 million by 1:2 yields a ratio of 1:3.6 million. "So on this entirely hypothetical basis we have got the figure of the probability of Adams being guilty from the non-DNA evidence as 1 in 3.6 Million."[32]

Having developed this estimate of the non-DNA evidence, Donnelly then combined it with the prosecution's DNA evidence figures. Recall that the prosecution gave a figure of 1:200 million for the chance that a matching profile would be found between the crime stain and a randomly chosen unrelated man from the European Caucasian population. Again, Donnelly stated the evidence in terms of a ratio between two hypothetical estimates. According to his calculation, if Adams was the actual rapist, the chance of a match between his blood sample and the crime scene evidence is 1:1 (100%), because both samples derive from the same source.[33] If, however, Adams was not the rapist, the chance of his DNA profile matching the crime stain would be 1:200 million (for the sake of the illustration, this figure accepts the prosecution's estimate, though as noted earlier the defense challenged that figure).

Recall Donnelly's illustrative figure from the exercise with the non-DNA evidence that there is one chance in 3.6 million that Adams is guilty. When this probability figure is weighed against the prosecution's figure of one chance in 200 million that Adams is innocent, the resulting estimate is 1:200/3.6 or one chance in approximately 55 that Adams is innocent. In other words, according to Donnelly's illustrative exercise, taking all of the evidence into account, it is 55 times more likely that Adams is guilty rather than innocent.[34]

Note that a successful challenge to the prosecution's estimate of the probability of the DNA profile match also can affect the result. As mentioned earlier, Donnelly challenged the visual match between Adams's sample and the crime scene evidence by stating that the indistinct autoradiograph band that Dr. Harris enhanced with his marking pen should not be counted in the multiplication of allele frequencies in the profile. If this band is excluded, then the DNA profile match estimate is reduced by a factor of ten, becoming one in 20 million (this is because in a single locus probe technique, the frequency estimates for each of the matching allele bands are multiplied together). Consequently, if we divide one in 20 million by one in 3.6 million, we get an estimate of one in 5.5. Accordingly, it is 5.5 times more likely that Adams is guilty rather than innocent. If the defense had managed to knock down the DNA estimate by another factor of ten, to one in two million, the result would drop accordingly: one in two million divided by one in 3.6 million, or roughly one chance in two that Adams is guilty.[35] And, finally, if Adams's brother had been included in the suspect pool, according to the Bayesian procedure, the balance of evidence favors his innocence, as there is now a chance of only one in eighteen that he is guilty. In a criminal trial in the UK (and in many other systems), jurors are instructed to register a not-guilty verdict if they maintain "reasonable doubt" about the totality of evidence against the accused person. There is no single point in a continuum of probabilities at which "reasonable doubt" begins, but it should seem clear that, while a probability of innocence of one in 290,000,000 (a number several times the size of the UK population) might seem unassailable, probabilities like one in 55, and surely one in 5.5, can much more easily provide a basis for doubt.

4. Appeals

The first appeal

In spite of the defense's efforts, on 24 January 1995, Adams was convicted for the rape of Miss M. The verdict indicated that the jury believed that the DNA evidence was sufficiently strong to prove Adams's guilt beyond a reasonable doubt. Adams appealed the conviction, and the Court of Appeal heard the case in April 1996. The grounds of appeal were, first, that the DNA evidence presented by the prosecution was insufficient by itself to establish guilt, and, second, that the judge had inadequately instructed the jury to apply Bayes's Theorem to the facts of the case. On 9 May 1996, the Court of Appeal announced that it accepted the defense's position that the trial court had not prepared the jury properly, and it referred the case for retrial. However, the Court rejected the other ground of appeal, and stated that it is up to the jury to decide if the DNA evidence is sufficient for conviction in the absence of corroborating evidence.[36]

According to the summaries provided by the Court of Appeal, in the two *Adams* trials, the DNA evidence stood alone as the primary, if not the sole, item of prosecutorial evidence. The Court (especially during the second appeal hearing) construed the *Adams* case as a test of the relative credibility and weight of DNA evidence, and in both appeals the Court raised vexed questions about the way the expert evidence may have dominated the jury's verdict. Importantly, for our purposes, the Court of Appeal also commented on the appropriateness of using Bayes's theorem to summarize the non-DNA as well as the DNA evidence.

> . . . we have very grave doubt as to whether that evidence [using Bayes theorem] was properly admissible, because it trespasses on an area peculiarly and exclusively within the province of the jury, namely the way in which they evaluate the relationship between one piece of evidence and another.

The "grave doubt" in this case had to do with the principle of "jury usurpation": an encroachment upon the jury's role as trier of fact in criminal cases. In the English common law tradition, expert witnesses are recognized as persons with specialized knowledge or experience that lay members of the court are unlikely to share. Although expert testimony about relevant facts can be authoritative (especially when such testimony is not challenged by other expert witnesses), it is also restricted. In principle, experts inform and advise the court about particular matters relevant to the case; they are not supposed to pronounce upon the ultimate guilt or innocence of the defendant.[37] The court went on to say that:

> The Bayes theorem may be an appropriate and useful tool for statisticians and other experts seeking to establish a mathematical assessment of probability. Even then, however, as the extracts from Professor Donnelly's evidence cited above demonstrate, the theorem can only operate by giving to each separate piece of evidence a numerical percentage representing the ratio between probability of circumstance A and the probability of B granted the existence of that evidence. The percentages chosen are matters of judgment: that is inevitable. But the apparently objective numerical figures used in the theorem may conceal the element of judgment on which it entirely depends.[38]

The Court's criticism was both legal and methodological. Much in the fashion of skeptical philosophers of science, the court reasoned that the quantification procedure would create only the *appearance* of objective judgment. More seriously, the Court suggested that the procedure would substitute a limited calculus for the jury's legal mandate to examine all of the evidence in the singular case. The court argued that

> . . . the theorem's methodology requires. . . that items of evidence be assessed separately according to their bearing on the accused's guilt, before being combined in the overall formula. That in our view is far too rigid an approach to evidence of the type that a jury characteristically has to assess, where the cogency of (for instance) identification evidence may have to be assessed, at least in part, in the light of the strength of the chain of evidence in which it forms part.[39]

The Court's rejection of the Bayesian procedure was based on a strict demarcation between a formulaic mode of individual reasoning and a collective form of common sense deliberation:

> More fundamentally, however, the mathematical formula, applied to each separate piece of evidence, is simply inappropriate to the jury's task. Jurors evaluate evidence and reach a conclusion not only by means of a formula, mathematical or otherwise, but by the joint application of their individual common sense and knowledge of the world to the evidence before them.[40]

The Court's judgment clearly rejected the rationale Donnelly had earlier given for introducing the Bayesian procedure. As he put it, Bayes's theorem was "the only logical way to assess all the evidence." He did not distinguish the procedure from common sense, nor did he associate it with "objective" outcomes. Instead, he recommended it as a tool that would enable each juror to represent different evidential judgments, and to combine their weights, on a coherent quantitative scale. He did not tell the jurors which probabilities to associate with specific items of evidence, such as the alibi evidence and the victim's failure to identify Adams. Far from being a source of "objective" judgments, for Donnelly and Thwaites the procedure was a way to counteract the *apparently objective* numerical figures representing the weight of the prosecution's DNA evidence; figures that can be shown to

depend upon a chain of fallible practices and judgments. The defense thus justified the use of the Bayesian method as a democratic procedure in the numero-politics of a criminal trial; a procedure that enabled "common sense" evidence to be weighed on the same scale as "scientific" evidence.

Because none of the parties to the appeal had argued against the admissibility of the Bayesian procedure, at the first appeal the Court of Appeal did not give an authoritative ruling on the matter, but it did express the opinion that Bayes's Theorem was not appropriate for aiding jurors to weigh all of the evidence in the case.

Retrial and second appeal

In the retrial, the defense again invited the jury to use the Bayesian procedure to calculate probabilities for all of the evidence, and Thwaites and Donnelly provided the jury with the questionnaire. The judge expressed reservations about the procedure, but permitted individual jurors to use the questionnaire on a voluntary basis. The judge instructed the jury, "If you feel able to use the questionnaire to operate Bayes's Theorem and you find it almost as easy as kiss your hand (sic.) to give answers, then you have the opportunity to do it, having not only your own copies but you will have when you go out an extra blank one to fill in your collective view if you want to." The judge added, however, that the jury's duty was to "consider your verdict amongst yourselves, all of you together and not with one huddled in a corner with his calculator."[41]

The arguments by the prosecution and defense were much the same as in the first trial, and once again the jury convicted Adams and the defense pursued an appeal. The terms of the second appeal were similar to the first. In a written summary, Mr. Thwaites argued that, in a case in which the prosecution was allowed to present DNA evidence in terms of probability figures, it was only fair that the defense should be allowed to use Bayes's Theorem to represent the non-DNA evidence in an equivalent way.[42] He added that the Bayesian approach is sound, logical, and approved by expert opinion. Thwaites also submitted that the trial judge's instructions to the jury gave insufficient (and even "flippant") instructions about using the Bayesian approach. Thwaites added that when the trial judge contrasted the Bayesian approach to the jurors' "common sense" reasoning, he misled the jury by ignoring the way Donnelly's questionnaire provided a means for making "scientific" and "common sense" evidence explicit and comparable.

The Court of Appeal considered the second appeal on 16 October 1997, and rejected each of the arguments in favor of the appeal. The Court also used the opportunity to give further authority to its earlier misgivings about the Bayesian procedure. Once again, the Court drew a firm distinction between the Bayesian approach and "a more conventional application of judgment." The Court stated that quantification was appropriate for the prosecution's DNA evidence, in so far as that evidence was "based... on empirical statistical data, the data and the deductions drawn from it being available for the defence to criticise and challenge."[43] The Court characterized the defense's evidence as "non-scientific, non-DNA evidence" and rejected the argument that such evidence should be represented in statistical form. To support this decision, the Court cited the judgment on the first *Adams* appeal, which as noted earlier, stated misgivings about whether the Bayesian procedure should be allowed. The Court also cited a more recent appeal court ruling in the case *R v. Doheny and [Gary Andrew] Adams*, which endorsed the first (Denis John) *Adams* appeal court's objection that the Bayesian procedure "plunges the jury into inappropriate and unnecessary realms of theory and complexity deflecting them from their proper task."[44] This chain of citations performed a kind of a bootstrapping operation on the "misgivings" initially

stated by the Court of Appeal in the judgment on the first *Adams* appeal.[45] Now, a year later, the second *Adams* appeal court could cite the earlier judgment together with the *Doheny & Adams* judgment as "previous rulings" on the matter in question. The Court concurred that "we regard the reliance on evidence of this kind in such cases as a recipe for confusion, misunderstanding, and misjudgment, possibly even among counsel, but very probably among judges and, as we conclude, almost certainly among jurors."[46] The court affirmed the "conventional" manner in which juries assess the evidence.[47] They concluded by saying:

> We are very clearly of opinion that in cases such as this, lacking special features absent here, expert evidence should not be admitted to induce juries to attach mathematical values to probabilities arising from non-scientific evidence adduced at the trial.[48]

The Court dismissed the appeal, and at that point, the *Adams* case was closed. Adams was about to come up for parole, so the outcome of the appeal was no longer crucial to his fate. The ruling about the Bayesian procedure may have more lasting significance, however.

5. Conclusion

The ultimate issue

The *Adams* case raised interesting issues that are not limited to the particular case or even to criminal law. A number of scholars have observed that statistical, probabilistic reasoning has become increasingly prevalent in many areas of professional life and public decision making.[49] Where the word "probability" once had no association with numbers, and non-numerical evaluations were deemed sufficient for almost all decisions, the precise calculation of risk is now said to be an emblem of late-modern society.[50] Many of the scholars who characterize this global trend express concern about the intimidating sense of objectivity associated with statistical treatments of matters that were once consigned to "common sense," and they raise questions about the ability of the numerically challenged masses to participate in decisions that are increasingly consigned to experts.[51] This is exactly what troubled the Court of Appeal when it reviewed the *Adams* trials and the Bayesian procedure advocated by the defense. The Court openly worried about the possibility that a technical method of calculative decision making would usurp the jury's (or, in a trial in which there is no jury, the judge's) responsibility to decide the "ultimate issue" of guilt or innocence.

Both *Adams* appeal judgments affirmed the role of "common sense" for making a holistic assessment of the totality of evidence presented in the singular case. As the courts recognized, such assessments not only involve discrete questions of fact; they also take into account the demeanor of witnesses and the credibility of testimony. They involve elements of trust, which are fallible, difficult to justify, and impossible to quantify.[52] The Court also stated that the jurors' common sense rests upon knowledge of the world, which cannot be reduced to an expert system.[53] The Court emphasized the social, as opposed to the individual, basis of jurors' judgments and affirmed the necessity for jurors collectively to deliberate about the ultimate issue. In sum, the two *Adams* appeal court decisions characterized the Bayesian approach as an individualistic, reductive calculus that creates a misleading or potentially confusing *appearance of objectivity* when applied to "non-scientific" evidence.

However, the *Adams* appeals did not simply result in an across the board rejection of a probabilistic calculus, nor did they simply uphold common sense against a technocratic procedure that threatened to subjugate (or "usurp") it. The Court did not discount the appropriateness of the Bayesian method for assisting judgments about *scientific* evidence,

and the Court did not go along with the defense's argument for an equivalent quantitative framing of the DNA and non-DNA evidence.[54] Donnelly and Thwaites argued that the courts should give the same quantitative treatment to all facts in the case, regardless of whether they are associated with empirical science. The Court rejected this argument and continued to demarcate "scientific" DNA evidence from "common sense" non-DNA evidence, and it ruled that numerical estimation was appropriate for the former type of evidence but not the latter. The Court of Appeal affirmed the jury's jurisdiction over the "ultimate issue," but at the same time it affirmed the "scientific" legitimacy of probability estimates. The Court's use of the science/non-science distinction produced an interesting twist on the theme of boundary work.[55] The Court did not simply assign special authority to "expert" or "scientific" testimony. Instead, it stipulated limits to "scientific" testimony and ascribed global authority to the jury's collective "common sense" deliberations.

Like other forms of graphic and statistical evidence, DNA profiles and probability estimates can confer such impressive weight in a case that it may seem pointless to dispute them.[56] A concern that was voiced soon after DNA profiling was introduced into criminal prosecutions was that technically unprepared judges and jurors would accord undue weight to DNA matches, so that forensic reports of DNA matches would function, in effect, as verdicts.[57] Accordingly, the immense odds against a random match supplied by forensic experts threatened to overwhelm (and thus "usurp") the jurors' assessment of all the evidence. The illustration of the Bayesian approach that Donnelly gave during his testimony showed that the reported odds against a defendant, even when as steep as 200 million to one, do not preclude a judgment of reasonable doubt about the totality of evidence in the case.[58] The juries in the two *Adams* trials apparently were not swayed by the demonstration, but the defense argued that the jury was inadequately prepared by the judge to use the method. From the defense's point of view, the Bayesian procedure, far from being a reification of expert evidence at the expense of common sense, was designed to give common sense a fighting chance when faced with the scientific authority and impressive probability figures associated with DNA evidence.

The Thwaites–Donnelly equivalence strategy was a matter of fighting fire with fire, but there are other ways to give equivalent treatment to DNA and non-DNA evidence. One alternative would be to avoid probability estimates in reports of DNA matches. This is what population geneticist Richard Lewontin advocates. Lewontin, a prominent critic of the use of DNA profiling in the courts (and more generally, a prominent critic of genetic reductionism), argues that probability figures should be barred from jury trials. He gives two reasons for this proposal. One is that it is impossible within the time constraints of a trial to educate a jury sufficiently to critically evaluate probabilistic evidence (according to Lewontin, it is difficult enough for undergraduate students at an elite university to grasp the relevant aspects of probabilistic reasoning after several months of study).[59] Interestingly, during the second appeal, Thwaites also stated that most lay people are "hopeless" when it comes to statistics, and they need an expert to assist them.[60] However, unlike Lewontin, he also asserted that statistics represent the "logic" of common sense reasoning, and that lay persons need experts to elucidate the statistical logic that is both inherent in their thinking and beyond their comprehension.

Lewontin's other argument is that the reference populations used for calculating probability estimates in criminal cases are never precisely tailored to the circumstances of the case at hand. In the *Adams* trial, for example, the calculus changes remarkably when the pool of potential suspects includes close relatives (specifically, in this case, a brother). Forensic organizations estimate allele frequencies for national, regional, and "racial" populations and subgroups, but these estimates do not take into account the unique makeup

of the pool of suspects for a particular crime. It would be impossible to specify such a group in advance of a trial, since the outlines of the group depend upon emergent judgments about the crime and about who in the world might have committed it. In addition, except in cases in which specific objections are made (such as when the *Adams* defense objected to Dr. Harris's enhancement of a band on an autoradiograph), the forensic estimates typically do not take into account the possibility of error or fraud in the handling and analysis of the evidence.[61] Consequently, no "single impressive statistic" can cover variations in comparison groups, police and forensic practices, and other singular contingencies of a case. Moreover, not all of these practices and contingencies can be traced back to the actions, procedures, and controls associated with a science. According to Lewontin's criticisms, the probabilistic estimates presented in court arise from an incalculable background of police practices, technical and non-technical judgments, and singular relevancies associated with the case at hand.

The argument to bar statistical evidence from the tribunal of fact is the opposite of the Bayesian argument, but both arguments give consistent approaches to the "ultimate issue." Both arguments challenge the *Adams* Appeal Court's reliance on an essential distinction between "scientific" evidence, which is legitimately stated in probabilistic terms, and "non-scientific" evidence, which should be left to the conventional, "common sense" judgments and deliberations of the jury. Both Donnelly's probabilistic arguments and Lewontin's objections to probability figures attempt to undermine the distinction between scientific and common sense evidence in favor of an equivalent representation and weighing of the facts in a case. Although the *Adams* Appeal Court mentioned concerns with jury usurpation, its categorical distinction between scientific and non-scientific evidence created a preserve of expert authority beyond the reach of the jury's "common sense."

Acknowledgments

The research for this paper was supported by a grant from the British Economic and Social Research Council (ESRC). "Science in a Legal Context: DNA Profiling, Forensic Practice, and the Courts" (R000235853), July 1995 to March 1998. We would like to thank three anonymous reviewers for their astute comments and criticisms of an earlier draft.

References

1 A collection of summaries was used for reconstructing this case: 1st Appeal, *R. v. Adams*, ([1996] 2 Cr App Rep 467), Judgment by Lord Justice Rose (Panel: Rose, Hidden, Buxton), Royal Courts of London, 26 April 1996 (*The Times*, London, 9 May 1996: 36); 2nd Appeal, *R. v. Adams* (Denis), 1st Court of Appeal, Criminal Division, Judgment by Mr. Justice Potts (Panel: Lord Bingham, Chief Justice, Potts, Butterfield), Royal Courts of Justice, London, 16 October 1997; and Ronald Thwaites Q.C. and Marc Brittain, "Perfected Grounds of Appeal & Skeletal Argument," *Denis John Adams v. Regina*, in The Court of Appeal, Criminal Division, 25 September 1997. Ruth McNally attended the second appeal, and McNally and Lynch interviewed Peter Donnelly (24 November 1997), an expert witness for the defense, and Ian Evett (18 December 1998), an expert for the prosecution.

2 The judicial summary stated that the attack came from behind, and the victim only saw her assailant briefly. Donnelly (interview 24 November 1997, p. 12), provided a different version of witness's testimony, and stated that the day after the rape the victim told the police that she had a good view of her attacker and was confident that she would be able to recognize him again. In fact, according to Donnelly, she did claim to see him walking down the street a couple of years later, but the police were unable to locate the man at the time.

3 See note 1 for references to judicial summaries. The larger project is a continuing study of the history of legal, scientific, and public controversies about forensic DNA profiling. It examines other trials, appeal cases, and admissibility hearings in the United States and United Kingdom, which raised issues that are less prominent in the *Adams* case. The issues include practical and legal contingencies for the handling and formulation of DNA evidence, the continuity of evidence (or chain of custody), and the development of DNA databases in the U.S.

and U.K. The most recent phase of the project examines the implications of closure in the controversy about DNA evidence and its implications for other forms of evidence (particularly fingerprint examination). Participants in the study include Kathleen Jordan, Ruth McNally, Simon Cole, and Michael Lynch. For a series of articles based on the study, see Michael Lynch and Sheila Jasanoff (eds.), "Science, law, and forensic practice," *Social Studies of Science* 28, nos. 5/6 (1998). Also see, Simon Cole, *Suspect Identities: A History of Fingerprinting and Criminal Identification* (Cambridge, MA: Harvard University Press, 2001); Linda Derksen, "Towards a sociology of measurement: making subjectivity invisible and negotiating measurement in the case of DNA fingerprinting," *Social Studies of Science* 30, no. 6 (2000): 803–845; Jennifer Mnookin, "Fingerprint evidence in an age of DNA profiling," *Brooklyn Law Review* 67, no. 1 (2001): 13–70; Michael Lynch and Ruth McNally, "Science, common sense, and common law: courtroom inquiries, and the public understanding of science," *Social Epistemology* 13, no. 3 (1999): 183–196; and Kathleen Jordan, *Sociological Investigations into the Mainstreaming of the Polymerase Chain Reaction* (Ph.D. Dissertation. Department of Sociology, Boston University, 1997).

4 This focus on a particular case is inspired by Harold Garfinkel's idea of "perspicuous settings": specific occasions, cases, or practices that explicate fundamental concepts, themes, and distinctions that social theorists tend to address abstractly, often in endless debates with one another. The idea is to find a setting and embed a study in that setting, in order to explicate a relevant concept, theme, or distinction (in this case, the science/commonsense distinction). Our reliance on textual reconstruction, and our limited legal backgrounds, greatly limits the extent to which this study is "embedded," in the way that Garfinkel recommends, in the case and its legal circumstances. See Harold Garfinkel, *Ethnomethodology's Program: Working Out Durkheim's Aphorism* (Lanham, MD: Rowman & Littlefield, 2002), 181–184.

5 J. L. Austin devised the category of "performative utterance" to describe linguistic expressions that, when uttered in the appropriate circumstances *institute* or *accomplish* real-worldly events. J L. Austin, *How to do Things with Words* (Oxford: Clarendon, 1962). So, for example, under the appropriate ceremonial conditions, an official can christen a ship or pronounce a couple to be married. Such declarations do not describe preexistent conditions; instead, they *perform*, or contribute to the performance of, the state of affairs in question.

6 The term "boundary work" was coined by Thomas Gieryn as a way to respecify the science/non-science distinction as a social phenomenon. See Thomas Gieryn, "Boundary-work and the demarcation of science from non-science," *American Sociological Review* 48 (1983): 781–795; Thomas Gieryn, "Boundaries of science," in *Handbook of Science, Technology & Society*, ed. Sheila Jasanoff, Gerald Markle, James Petersen et al., (Beverly Hills, CA: Sage, 1994), 393–443; Thomas Gieryn, *Cultural Boundaries of Science: Credibility on the Line* (Chicago: University of Chicago Press, 1999). Unlike the philosophers who have attempted to formulate principled demarcation criteria, Gieryn recommends a social-historical orientation to efforts by scientific associations and courts to accord or deny scientific status to candidate fields such as phrenology, creation science, and various social sciences.

7 As an anonymous reviewer of an earlier draft of this paper noted, judges, attorneys, and expert witnesses are in a different situation from that of a jury. Unlike jurors, judges and attorneys often encounter repeated presentations of similar types of evidence and repeated performances by the same expert witnesses. Witnesses, attorneys, and judges thus have an opportunity to learn from prior cases and to adjust their actions accordingly. It is not necessarily the case, however, that judges have superior understanding of technical issues than jurors. See Gary Edmond and David Mercer, "Scientific literacy and the jury: reconsidering jury 'competence'," *Public Understanding of Science* 6 (1997): 329–357.

8 For discussions of the general problems associated with the role of expert evidence in legal and legislative decisions, see Roger Smith and Brian Wynne (eds.), *Expert Evidence: Interpreting Science in the Law* (London: Routledge, 1989); Sheila Jasanoff, *Science at the Bar: Law, Science and Technology in America* (Cambridge, MA: Harvard University Press, 1996); and Gary Edmond and David Mercer, "Litigation life: law-science knowledge construction in (Bendectin) mass toxic tort litigation," *Social Studies of Science* 30, no.2 (2000): 265–316. Jury competence is a subject of divergent views. Peter Huber, *Galileo's Revenge: Junk Science in the Courtroom* (New York: Basic Books, 1991), emphasizes the incompetence of jurors and other lay participants in trials to tell the difference between genuine and "junk" science. Lindsey and his colleagues suggest that jury understanding can depend upon how evidence (particularly probabilistic evidence) is presented. Samuel Lindsey, Ralph Hertwig, and Gerd Gigerenzer, "Communicating statistical evidence," *Jurimetrics* (2002), in press. For a similar suggestion made in an experimental study of probabilistic formulations of DNA evidence, see F. Taroni and C. Aitken, "Probabilistic reasoning in the law: part 1: assessment of probabilities and explanation of the value of DNA evidence," *Science & Justice* 38, no. 3 (1998): 165–177.

9 The *New York Times* (2 March 2001, A-1, 17) reported that in the past 40 years, jury trials ending with verdicts in the United States had dropped from 5.4 to 1.5 percent of civil cases and from 10 to less than 5 percent of criminal trials.

10 John Henry Wigmore famously stated that the presentation and testing of evidence in the trial court is "the greatest legal engine ever invented for the discovery of truth," in *Wigmore's Code of the Rules of Evidence in Trials of Law*, 3rd ed. (Boston: Little Brown, 1940), 29.

11 For a more elaborate description of DNA profiling techniques oriented to lay readers, see National Research Council, *The Evaluation of Forensic DNA Evidence* (Washington D.C.: National Academy Press, 1996). In the mid-1990s, the STR system began to replace the SLP technique, and the latter is rarely used anymore. However, because SLP was the key profiling technique used in *Adams*, it remains the focus of our discussion.

12 This capsule summary leaves out a great deal of technical detail. It also leaves out many of the contingent practices involved in collecting, transporting, and handling evidence. For more elaborate discussion of procedural descriptions of DNA profiling and their relations to practical contingencies, see Kathleen Jordan and Michael Lynch, "The dissemination, standardization, and routinization of a molecular biological technique," *Social Studies of Science* 28, nos.5/6 (1998): 773–800; and Saul Halfon, "Collecting, testing and convincing: DNA experts in the courts," *Social Studies of Science* 28, nos. 5/6 (1998): 801–828.

13 Recently, and partly as a result of the widespread recognition of DNA profiling as the new gold standard of criminal justice, the older gold standard—fingerprinting—has undergone more critical scrutiny than it had for decades. In U.S. Federal Court, and in an increasing number of state courts, starting with *US v. Mitchell*, No. 96–407 (E.D. Pa. 1999), fingerprint examination has been subject to admissibility hearings. Thus far, the courts have upheld the practice, but a decision by a prominent federal judge (Louis H. Pollak) in *US v. Plaza*, No. 98–362 (E.D. Pa. 2002) restricted the presentation of fingerprint "matches" in courts. This decision attracted a great deal of media attention, as well as strong protests from the forensic community. Judge Pollak was persuaded to reconsider the decision, and a few months later he retracted the part of the ruling that restricted fingerprint examiners from explicitly using the word "match" when presenting fingerprint evidence in court. Challenges to fingerprinting continued to be mounted, both in court and in the forensic literature. Ian Evett of the British Forensic Science Service (FSS) in a co-authored paper with Christophe Champod (also of the FSS) criticized the non-probabilistic way in which fingerprinting evidence is presented in court (Christophe Champod & Ian Evett, "A probabilistic approach to fingerprint evidence," *Fingerprint World* 27, no. 105 [July 2001]: 95–107). Evett also was involved in the *Adams* case, as an expert for the prosecution who consulted with Peter Donnelly (the expert for the defense) over the design of the questionnaire used in that case. The challenges to fingerprint examination do not contest the doctrine that each individual has unique fingerprints. Instead, they focus on the fallibility of fingerprint examination, insisting that the practice has unknown reliability and the method lacks a "scientific" and/or probabilistic basis. See Simon Cole, *Suspect Identities: A History of Fingerprinting and Criminal Identification* (Cambridge, MA: Harvard University Press, 2001, 287ff), and Jennifer Mnookin, "Fingerprint evidence in an age of DNA profiling." The relationship between DNA profiling and fingerprinting is the subject of current research by Michael Lynch and Simon Cole.

14 In 1994, the Home Office announced an initiative to construct a national DNA profile database. This national database uses the STR system. The first profiles were entered in April 1995.

15 For a description of the database, see J. E. Allard, "Series sexual crimes identified by a DNA computerised database," *Advances in Forensic Haemogenetics*, 4 (1992): 295–297.

16 Melvin Pollner introduced the idea of "reality disjunctures" to describe incommensurable accounts of (arguably) the same experience (Melvin Pollner, "The very coinage of your brain: the anatomy of reality disjunctures," *Philosophy of the Social Sciences* 5 [1975]: 411–430). Examples include patient reports of experiences that psychiatrists attribute to hallucinations and contrary testimony about the speed of a particular automobile by defendant and arresting officer in traffic court. Paul Drew examines testimony at a rape case in which the victim and defendant give different versions of the circumstances leading to the alleged crime. See Paul Drew, "Contested evidence in courtroom cross-examination: the case of a trial of rape," in *Talk at Work: Interaction in Institutional Settings*, eds. John Heritage and Paul Drew (Cambridge: Cambridge University Press, 1992), 470–520.

17 As a reviewer of an earlier draft noted, the defense also relied upon expert evidence (viz., the testimony of Peter Donnelly). The Appeal Court tended to draw a sharp distinction between expert and commonsense evidence, independently of Donnelly's role as an expert. Donnelly presented himself as an expert who assisted the defense with a critique of the prosecution's "expert" evidence and who offered to represent the "common sense" evidence with a Bayesian procedure.

18 Judgment, 1st appeal, p. 9 (page numbers are to a computer aided transcript by Smith Bernal, London).

19 We did not have direct access to the two trials, and so our reconstruction is limited by the partial and retrospective summaries, citations, and quoted portions of transcript provided by the Court of Appeal.

20 Quotations of witness statements by Dr. Harris, Ms. Lygo, and Mr. Lambert were taken from the document Thwaites and Brittain, (cit. no. 1), p. 2, sec. 3.1.

21 Ibid.
22 The reason that four probes would result in seven bands is that three of the probes identified heterozygous sites in the particular samples (two each for three probes) and one identified a homozygous site.
23 Thwaites and Brittain, (cit. no. 1), p. 3, sec. 3.5.
24 This is an instance in which the features attributed to "pathological science" are used as argumentative resources for discounting (in this case, reducing the weight of) particular evidence. The theme of pathological science was introduced by Irving Langmuir, *Pathological Science*. General Electric, R&D Center Report No. 68-C-035, (Schenectady, N.Y.), 1968. For a study of particularistic uses of arguments to support and discount evidence, see G. Nigel Gilbert and Michael Mulkay, *Opening Pandora's Box: An Analysis of Scientists' Discourse* (Cambridge, UK: Cambridge University Press, 1984).
25 Judgment, first appeal, 1996 (cit. no. 1), p. 3.
26 Judgment, second appeal, 1997 (cit. no. 1), p. 3. The statement is confusing, because a match probability of 1 in 220 is much *higher* (and not lower, as the statement implies) than a probability of, for example 1 in 200 million. Perhaps the statement meant to say that including the brother would reduce the odds against the possibility of a profile from someone other than Denis John Adams matching the crime stain.
27 For an illuminating discussion of how burden of proof for including particular persons in the suspect pool is shifted from prosecution to defense in DNA profiling cases, see Michael Redmayne, "Doubts and burdens: DNA evidence, probability and the courts," *Criminal Law Review* (1995); 464–482.
28 Thwaites and Brittain, (cit. no. 1), p. 4.
29 Ibid., p. 5.
30 Donnelly, quoted in Judgment, 1st Appeal, 1996.
31 Ibid.
32 Ibid.
33 The 100 percent figure ignores the chance of a false negative (for example, a non-match between Adams's blood profile and the crime stain due to an error in the collection or analysis of the samples).
34 This calculation must assume that the series of estimates covers all of the relevant evidence. They must be exhaustive. Otherwise, it would be inappropriate to make a guilty or not-guilty judgment on the basis of the series of estimates. It is obvious that the four items of defense evidence can be divided further. One could, for example, treat the witness's misidentification of Adams in the police line-up as a separate item from the mismatch between Miss M's description of the perpetrator's age and Adams's age. According to Donnelly (interview, cit. no. 1, p. 21), as long as all of the evidence is represented in the series of estimates, the overall calculation should, in principle, not be affected by the way it is parsed into separate items. Others, including a Mr. Lambert, who testified for the prosecution in the second trial, expressed misgivings on this point, arguing that the questionnaire that Donnelly prepared did not cover all of the relevant evidence, or all of the evidence that particular jurors might consider relevant (second appeal summary, cit. no. 1, p. 7).
35 Donnelly also criticized the forensic scientists for not adding a correction factor for sampling error. This had a relatively minor effect (1 percent) on the resulting estimate.
36 In a later case, *R v. Watters* (Court of Appeal [Criminal Division], October 19, 2000), the panel of judges on the Court held that DNA evidence without corroborating evidence was not sufficient to convict in that case. This judgment referred to the specific circumstances of the case: a burglary case in which an STR profile developed from a cigarette butt found at the crime scene matched the defendant's profile. The defendant had two brothers, and including them in the suspect pool greatly increased the probability that another person besides the defendant could have matched the crime scene evidence.
37 An anonymous reviewer of the draft of this article commented that we make too much of the "ultimate issue," because, in practice, it is often ignored and easily circumvented. We have no reason to doubt such an assessment, and if it is so that the "ultimate issue" is less than ultimate, then the Appeal Court's insistence upon its relevance should be viewed as a contingent way of settling the case at hand, rather than as an invocation of an omnirelevant and binding judicial principle for any comparable case.
38 First appeal, (cit. no. 1), p. 15.
39 Ibid.
40 Ibid.
41 Summing up of His Honour Judge Pownall, quoted in Second appeal, 1997, p. 8.
42 A summary of Thwaites's arguments was presented in the Lord Chief Justice's summary, 2nd appeal, (cit. no. 1), pp. 5–8.
43 Second appeal, (cit. no. 2), p. 10.
44 *R v Adams* (1996) 2 Court of Appeal, R 467, cited in *R v Doheny and Adams* (1997) 1 Court of Appeal R 369, 374 G, which in turn is cited in Ibid., p. 11. < AU: please add note number to #44 above >

45 The theme of judicial bootstrapping deserves more attention than we give it within the frame of this paper. See Gary Edmond and David Mercer, "Litigation life: law-science knowledge construction in (Bendectin) mass toxic tort litigation."
46 Second appeal, (cit. no. 1), p. 12.
47 Interestingly, the court exemplified the "conventional" approach by describing how a juror might work through the various elements of the *Adams* case. This description included a sequence of probabilistic judgments, much as a Bayesian might describe them, except that the judgments would not be numerical.
48 Second Appeal, (cit. no. 1), p. 13.
49 See Ian Hacking, *The Emergence of Probability: A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference* (Cambridge, UK: Cambridge University Press, 1975); Gerd Gigerenzer, et al., *Empire of Chance: How Probability Changed Science and Everyday Life* (Cambridge, UK: Cambridge University Press, 1989); Theodore Porter, *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life* (Princeton: Princeton University Press, 1995).
50 Ulrich Beck, *Risk Society: Towards a New Modernity* (London: Sage, 1992).
51 See Roger Smith and Brian Wynne (eds.), *Expert Evidence: Interpreting Science in the Law* (London: Routledge, 1989); and Alan Irwin and Brian Wynne (eds.), *Misunderstanding Science? The Public Reconstruction of Science and Technology* (Cambridge, UK: Cambridge University Press, 1996).
52 See Steven Shapin, "Cordelia's love: credibility and the social studies of science," *Perspectives on Science* **3** (1995): 255–275.
53 This distinction resonates with criticisms of artificial intelligence and expert systems. See Hubert Dreyfus, *What Computers Can't Do* (New York: Harper and Row, 1979); and H. M. Collins, *Artificial Experts: Social Knowledge and Intelligent Machines* (Cambridge, MA: MIT Press, 1990).
54 Readers familiar with the sociology of scientific knowledge may be reminded of the "equivalence postulate" advocated by Barry Barnes and David Bloor in their proposals for a Strong Programme in the sociology of scientific knowledge. Barry Barnes and David Bloor, "Relativism, rationalism and the sociology of knowledge," in *Rationality and Relativism*, ed. Martin Hollis and Stephen Lukes (Cambridge, MA: MIT Press, 1982), 21–47. Their equivalence postulate states that "all beliefs are on a par with one another with respect to the causes of their credibility" (p. 23). Accordingly, sociologists of knowledge should give the same form of explanation to all instances of knowledge or belief, regardless of whether those beliefs presently enjoy the status of scientific truth. Bloor's better-known postulates of "symmetry" and "impartiality" are similar. David Bloor, *Knowledge and Social Imagery* (London: Routledge and Kegan Paul, 1976). It is possible to identify parallel themes between the sociology of scientific knowledge (SSK) and the arguments in an adversary trial (see Michael Lynch, "The discursive production of uncertainty: the O. J. Simpson 'Dream Team' and the sociology of knowledge machine," *Social Studies of Science* 28, nos. 5/6 (1998): 829–868). However, it would be misleading to suppose that, in the *Adams* case, the participants in the trial are acting in accordance with the concepts and methodological strategies of SSK. It would be equally misleading to suppose that Thwaites and Donnelly lay out a position that is akin to a radically skeptical or "relativist" treatment of *all* knowledges as being "on a par with one another." Thwaites and Donnelly seek to trump one "scientific" card (forensic DNA evidence) with another "scientific" (or mathematical) card (Bayesian probability). Their questionnaire is designed to allow jurors to weigh "scientific" (legal) evidence and "common sense" (legal) evidence on the same scale, but this circumscribed strategy does not suggest or imply a general hypothesis about "the causes" of credibility. Instead, it is a calculative device for inscribing and weighing credibility judgments.
55 See Gieryn, "Boundary-work and the demarcation of science from non-science."
56 See Cole, *Suspect Identities*, for a discussion of appeal cases in the United States in the late 1920s and early '30s, in which the issue of jury usurpation pertained to the testimony of fingerprint examiners. In some cases, the appeal courts held that fingerprint examiners encroach upon the jury's province when they describe matching latent and rolled prints as matters of "fact" rather than expert "opinion." See *State v. Steffen* 230 N.W. 536 (Iowa, 1930).
57 For a sociological discussion of procedural and statistical problems associated with early versions of DNA profiling, see Derksen, "For a sociology of measurement." A key point in criticisms of court treatments of DNA profile evidence is that such evidence should be treated as "reports" by particular agents and agencies rather than simple scientific facts. Jonathan J. Kohler, "On conveying the probative value of DNA evidence: frequencies, likelihood ratios, and error rates,' *University of Colorado Law Review* 67 (1996): 859–886, at 868.
58 By the year 2000, probes were developed for as many as 13 DNA loci in the STR system. Estimates of more than one trillion to one were given against the odds against random matches between profiles from unrelated persons. With such impressive figures in hand, it has been argued that it will no longer be necessary even to consider the possibility that more than one person in the relevant suspect population could possibly match a

given crime stain profile. Exceptions would be made for identical twins, and (with low likelihood) other very closely related persons. See Bruce Budowle, Ranajit Chakraborty, George Carmody, et al., "Source attribution of a forensic DNA profile," *Forensic Science Communications* 2, no. 3 (July, 2000) (available at: http://www.fbi.gov/programs/lab/fsc/backissu/july2000/source.htm).

59 Richard Lewontin, Harvard University, interviewed by Kathleen Jordan, 7 April 1998. The position he developed in the interview is consistent with one he has held all along. For a non-technical account of his argument, see his *Biology as Ideology: The Doctrine of DNA* (New York: Harper Collins, 1992), 78–83.

60 Ruth McNally, notes taken while attending second appeal, 17 October, 1997.

61 Jonathan Kohler ("On conveying the probative value of DNA evidence," p. 868) argues that "the possibility of laboratory error is substantially larger than the possibility of a coincidental match." For an argument against the policy of reporting match probabilities without taking laboratory errors (or evidence collection, mislabeling, etc.) into account, see William C. Thompson, "DNA evidence in the O. J. Simpson trial," *University of Colorado Law Review* 67 (1996): 827–857.

Authors

Michael Lynch, Department of Science & Technology Studies, Cornell University, Ithaca, New York, 14853.

Ruth McNally, School of Business & Management, Brunel University.