

5-8-2013

Whigs in Court: Historiographical Problems with Expert Evidence

Gary Edmond

Follow this and additional works at: <http://digitalcommons.law.yale.edu/yjlh>



Part of the [History Commons](#), and the [Law Commons](#)

Recommended Citation

Edmond, Gary (2002) "Whigs in Court: Historiographical Problems with Expert Evidence," *Yale Journal of Law & the Humanities*: Vol. 14: Iss. 1, Article 3.

Available at: <http://digitalcommons.law.yale.edu/yjlh/vol14/iss1/3>

This Article is brought to you for free and open access by Yale Law School Legal Scholarship Repository. It has been accepted for inclusion in Yale Journal of Law & the Humanities by an authorized administrator of Yale Law School Legal Scholarship Repository. For more information, please contact julian.aiken@yale.edu.

Whigs in Court: Historiographical Problems with Expert Evidence

Gary Edmond*

I. INTRODUCTION

The trouble with all this is that it is setting the jury to decide, where doctors disagree. . . . But how can the jury judge between two statements each founded upon an experience confessedly foreign in kind to their own?

—Learned Hand (1901)¹

At the beginning of the twentieth century, Learned Hand expressed concern at the assessment of expert disagreement by the lay jury. While the debate over jury competence has continued, Hand's disquiet would appear to apply equally to historians, lawyers, and judges commenting on litigation involving protracted disputes between experts.² Hand's comment may actually raise the methodological question: how should historians and legal commentators approach and explain disagreements among experts and scientists during trials and appeals? This Article endeavors to sketch some tentative answers to that question, primarily through the review of several cases exemplifying the recent historiographical treatment of expert evidence. Recent approaches are conspicuous because, where

* B.A. (Hons), University of Wollongong, L.L.B. (Hons), University of Sydney, Ph.D., Cambridge University. Faculty of Law, University of New South Wales, Sydney 2052, Australia, and Visitor, Law Program, Research School of Social Sciences, The Australian National University, Canberra 0200 Australia. The author would like to thank David Mercer, Wilfred Prest, Rob Nelson, Ivan Crozier, Penny Pether, Paul Roberts, Mike Redmayne, Leighton McDonald, John Gava, John Williams and members of the Research Discussion Group, Adelaide University.

1. Learned Hand, *Historical and Practical Considerations Regarding Expert Testimony*, 15 HARV. L. REV. 40, 54 (1901).

2. *Id.* For a more general historical introduction see CAROL JONES, *EXPERT WITNESSES: SCIENCE, MEDICINE AND THE PRACTICE OF LAW* (1994) and Tal Golan, *The History of Scientific Expert Testimony in the English Courtroom*, 12 SCI. CONTEXT 7 (1999).

they focus on the evidentiary contests, they invariably draw upon idealized images of scientific knowledge and practice and usually attribute putatively proper values to the evidence—values frequently based on popular or official rationalizations of the litigation—and use them to interpret the entire litigation presupposing the continued existence and availability of those purportedly stable values. This later tendency might be described as *whiggish*, for it involves the decontextualized comparison of evidence from earlier trials and appeals with an allegedly proper value subsequently attributed at the (apparent) conclusion of the litigation. These approaches might also be characterized as *sociologies of error*, for they exhibit a tendency to explain purportedly *unreliable* evidence in sociological terms and purportedly *reliable* evidence on the basis of its intrinsic epistemic worth. The tendency to accept the value accorded to the evidence at the close of proceedings, trivializing the contingent and strategic processes involved in its production, articulation, and assessment, tends to invest the entire proceedings with a particular moral configuration that often facilitates recrimination and proposals for procedural reform.

In contrast, this Article aims to illustrate how such approaches to expert evidence achieve their coherence and plausibility on the basis of simplistic and artificial images of the sciences and at the cost of detailed attention to the legally influenced diachronic contest surrounding the resolution of expert disagreement. The celebration of putatively correct outcomes prevents analysts from asking more fundamental questions about why some types of knowledge are considered (and represented as) reliable and others unreliable, how that status is achieved through the course of adversarial litigation and whether legal institutions can or should be predominantly concerned with truth, or whether commitment to other, perhaps equally elusive, values such as defendants' rights, procedural fairness, and justice transform legal settings into quite peculiar epistemic domains.³

3. This usage is not intended to reify "justice" or "procedural fairness." Just as with the concept of "truth," in this Article it is suggested that their meanings be interpreted in context, as part of the flexible and contested matrix stimulated by adversarial litigation. The content or meaning of "truth" (or reliable knowledge/evidence), "rules," "justice," and "procedure" are all potentially "up for grabs" in the course of adversarial litigation. The primacy of "justice" and the role accorded to "truth" seem to be a point of tension in recent discussions of evidence and procedure. See, e.g., DANIEL FARBER, *BEYOND ALL REASON: THE RADICAL ASSAULT ON TRUTH IN AMERICAN LAW* (1997); PHANTOM RISK: *SCIENTIFIC INFERENCE AND THE LAW* (Kenneth Foster et al. eds., 1993); see generally PAUL GROSS & NORMAN LEVITT, *HIGHER SUPERSTITION: THE ACADEMIC LEFT AND ITS QUARRELS WITH SCIENCE* (1994); ALAN SOKAL & JEAN BRICMONT, *INTELLECTUAL IMPOSTURES* (1998).

A. The Temporal, Epistemic, and Social Context: Whig Historiography, Methodological Symmetry, and Idealized Images of Science

The foundations for the following discussion are predicated upon a combination of three theoretical approaches drawn from the history of science and the sociology of scientific knowledge. The first is derived from Herbert Butterfield's classic monograph, *The Whig Interpretation of History*,⁴ and the second is David Bloor's symmetrical postulate described in *Knowledge and Social Imagery*.⁵ The third seeks to incorporate a series of historical and empirical assaults on positivist, ahistorical, and idealized images of scientific knowledge and practice drawn from post-Kuhnian empirical and ethnographic studies of the sciences and historiography.⁶

According to Butterfield it was "part and parcel of the whig interpretation of history that it studies the past with reference to the present."⁷ The whig interpretation provides a simplified framework through which to interpret the complexities of history, enabling the historian fallaciously to "cut through this complexity."⁸ In effect, it represents a "short cut," a strategically abridged version of history rationalized on the basis of adherence to a retrospectively informed (set of) organising principle(s).⁹ This enables the historian to classify historical actors into those who "furthered progress and the men who tried to hinder it; so that a handy rule of thumb exists by which

4. HERBERT BUTTERFIELD, *THE WHIG INTERPRETATION OF HISTORY* (W.W. Norton & Co., 1965) (1931).

5. DAVID BLOOR, *KNOWLEDGE AND SOCIAL IMAGERY* (1991). The approach developed by Bloor is commonly described as the Strong Programme, and in combination with work by David Edge, Barry Barnes, and Donald MacKenzie, also characterized as the Edinburgh School.

6. For an early and influential statement of post-Kuhnian historiography of the sciences, consider Steven Shapin, *History of Science and Its Sociological Reconstructions*, 20 *HIST. SCI.* 157 (1982). For more particular historical and sociological concern with expert evidence and litigation, see *EXPERT EVIDENCE: INTERPRETING SCIENCE IN THE LAW* (Roger Smith & Brian Wynne eds., 1989); SHEILA JASANOFF, *SCIENCE AT THE BAR: LAW, SCIENCE AND TECHNOLOGY IN AMERICA* (1995); JONES, *supra* note 2; BRIAN WYNNE, *RATIONALITY AND RITUAL: THE WINDSCALE INQUIRY AND NUCLEAR DECISIONS IN BRITAIN* (1982); *Contested Identities: Science, Law and Forensic Practice*, 28 *SOC. STUD. SCI.* 675 (Mike Lynch & Sheila Jasanoff eds., 1998); *Science and Law*, 12 *SCI. CONTEXT* 3 (Tal Golan & Snait Gissis eds., 1999).

7. BUTTERFIELD, *supra* note 4, at 11. Another frequently cited description of whiggism is "The study of the past with one eye, so to speak, upon the present is the source of all sins and sophistries in history. . . . It is the essence of what we mean by the word 'unhistorical.'" *Id.* at 31-32. For further discussion consider A. Rupert Hall, *Can the History of Science Be History?* 4 *BRIT. J. HIST. SCI.* 207, 217 (1969); Mark Tushnet, *Interdisciplinary Legal Scholarship: The Case of History-in-Law*, 71 *CHI.-KENT L. REV.* 909, 914-17 (1996).

8. BUTTERFIELD, *supra* note 4, at 14, 21, 22.

9. *Id.* at 22-23, 40, 43, 96-103. See also Mary Hesse, *Reasons and Evaluation in the History of Science*, in *CHANGING PERSPECTIVES IN THE HISTORY OF SCIENCE* 127 (M. Teich & R. Young eds., 1973).

the historian can select and reject and can make points of emphasis.”¹⁰

On occasion Butterfield suggested that some of the pitfalls associated with whig history might be remedied through careful attention to historical detail. Notoriously, he remarked that “if at any point we need further elucidation all that he [the historian] can do is to take us into greater detail, and make us see in still more definite concreteness what really did take place.”¹¹ That suggestion, drawing an analogy between historical practice and optical magnification, has generated criticism on the ground that it proposes a “specious inductivism.” The available historical record cannot be legitimately equated with the richness of the past, nor can ever-decreasing minutiae of facts “assemble themselves atheoretically into explanations.”¹² While the following account emphasizes the importance of richly textured contextual historiography, it makes no attempt to suggest that context alone is sufficient to account for all social activity.¹³ Perhaps more importantly, it serves to demonstrate that even accounts of expert evidence that provide tremendous contextual detail nevertheless risk serious historiographical limitations if they embrace historically and empirically implausible models of scientific knowledge and practice.¹⁴

10. BUTTERFIELD, *supra* note 4, at 11. See also Paschal Phillips, *The “Whig Interpretation” of Monastic History*, 31 ABR 201 (1980).

11. BUTTERFIELD, *supra* note 4, at 72; See also Adrian Wilson & T.G. Ashplant, *Whig History and Present-Centred History*, 31 HIST. J. 1, 7 (1988).

12. A. Rupert Hall, *On Whiggism*, 21 HIST. SCI. 45, 51-52 (1983). Butterfield and his student, Hall, seem to have accepted the proposition that greater attention to detail would prevent some whiggish tendencies. However, in adopting that proposition they appear to have underestimated the profound influence of assumptions about the sciences and asymmetrical explanations in shaping analysis in the history of the sciences. Consequently, more recent critics have suggested that one potential means of improvement is to abandon “present-centred concepts and to reconstruct the activities of agents in their original context” because flawed categories, such as idealized models of science, may well survive and subordinate the experience and findings of historical research. T.G. Ashplant & Adrian Wilson, *Present-Centred History and the Problem of Historical Knowledge*, 31 HIST. J. 253, 260-61 (1988).

13. See Quentin Skinner, *Meaning and Understanding in the History of Ideas*, 8 HIST. & THEORY 3 (1969). Skinner provided influential criticism of the existence of perennial problems in the history of ideas, and the inability of context to fully account for ideas. See also John Dunn, *The Identity of the History of Ideas*, 43 PHIL. 85 (1968). Similar objections have been raised in response to sociological attempts to explain behaviour and the success of some ideas on the basis of attributed interests. See BARRY BARNES, *INTERESTS AND THE GROWTH OF KNOWLEDGE* (1977); Michel Callon & Bruno Latour, *Don’t Throw the Baby out with the Bath School! A Reply to Collins and Yearly*, in *SCIENCE AS PRACTICE AND CULTURE* 343 (Andrew Pickering ed., 1992); Harry Collins & Steven Yearly, *Epistemological Chicken*, in *SCIENCE AS PRACTICE AND CULTURE*, *supra*, at 301; Donald MacKenzie, *Interests, Positivism and History*, 11 SOC. STUD. SCI. 498 (1981); Steve Woolgar, *Interests and Explanation in the Social Study of Science*, 11 SOC. STUD. SCI. 365 (1981).

14. See Ashplant & Wilson, *supra* note 12, at 6; Hall, *supra* note 12, at 51-52. But see John Diggins, *The Oyster and the Pearl: The Problem of Contextualism in Intellectual History*, 23 HIST. & THEORY 151 (1984). Significantly, Diggins’s analysis makes what amounts to routine

Accepting that Butterfield's text has been subjected to considerable criticism, primarily for lack of clarity and its inability to suggest a solution to the fact that contemporary assumptions and values influence all historical practice,¹⁵ including his own influential history of science,¹⁶ in this Article, the term "whig" will be restricted to denote the practice of employing (more) recent perspectives to interpret, often critically or ironically, earlier values, commitments, and actions.¹⁷ In relation to studies of litigation, the whig historian (or legal whig) will be shown to demonstrate a tendency to accept and privilege the standing of the evidence or findings at the outcome of legal proceedings, and to assess earlier performances and interpretations of the evidence against them. That is, legal whiggism is rendered conspicuous through a tendency to be present-centred in its evaluation of the past.¹⁸ Unlike its political counterpart, the appellation "legal whiggism" does not necessarily entail a glorification of the progression of certain cherished principles. Instead, it often results in a critical account of the treatment of a particular case in a manner that leads to proposals for legal reform.

The second analytical tool is methodological symmetry.¹⁹ In an attempt to extend the efforts of earlier sociologists of knowledge such as Mannheim and Durkheim,²⁰ the philosopher David Bloor proposed a systematic extension to the study of the sciences designed to incorporate an examination of the production and content of

reference to a broader context to support his various arguments that attempt to displace intellectual history from a specific contextual milieu.

15. See, e.g., WILLIAM H. DRAY, ON HISTORY AND PHILOSOPHERS OF HISTORY 164-90 (1989); Ashplant & Wilson, *supra* note 12; William H. Dray, *J.H. Hexter, Neo-Whiggism and Early Stuart Historiography*, 26 HIST. & THEORY 133 (1987); G.R. Elton, *Herbert Butterfield and the Study of History*, 27 HIST. J. 729 (1984); Hall, *supra* note 12; Wilson & Ashplant, *supra* note 11.

16. HERBERT BUTTERFIELD, THE ORIGINS OF MODERN SCIENCE 1300-1800 (1950). For criticism see Andrew Cunningham & Perry Williams, *De-centring the "Big Picture": The Origins of Modern Science and the Modern Origins of Science*, 26 BRIT. J. HIST. SCI. 407 (1993). Also consider the qualified defense by Regis Cabral, *Herbert Butterfield (1900-79) as a Christian Historian of Science*, 27 STUD. HIST. & PHIL. SCI. 547 (1996).

17. This is consistent with the framing of Butterfield's major concerns espoused by Marshall Poe, *Butterfield's Sociology of Whig History: A Contribution to the Study of Anachronism in Modern Historical Thought*, 25 CLIO 345 (1996), and the notion of "present-centredness" articulated by Wilson & Ashplant, *supra* note 11, at 11.

18. "Present-centredness" is Ashplant and Wilson's refinement of Butterfield's concept of whiggism. In their analysis, whiggism is merely one aspect of present-centredness—typified by a particular concern with value judgements and origins. Ashplant & Wilson, *supra* note 12, at 261.

19. BLOOR, *supra* note 5, at 1-23.

20. See, e.g., EMILE DURKHEIM, THE ELEMENTARY FORMS OF THE RELIGIOUS LIFE (1915); EMILE DURKHEIM, THE RULES OF SOCIOLOGICAL METHOD (1938); KARL MANNHEIM, ESSAYS ON THE SOCIOLOGY OF KNOWLEDGE (1952); KARL MANNHEIM, IDEOLOGY AND UTOPIA (1936). For a discussion see MICHAEL MULKAY, SCIENCE AND THE SOCIOLOGY OF KNOWLEDGE 1-26 (1979).

scientific knowledge. Conceptually, this approach extended well beyond the circumstances surrounding the direction and degree of scientific investigation, such as funding, personal and professional recognition, and reward, that had characterized an externally oriented sociology of science exemplified by the work of Robert Merton and other primarily North American scholars.²¹ Bloor's approach was a response to an apparent reluctance to describe the social conditions attending the production as well as the actual content of ostensibly reliable knowledge in the sociology, philosophy, and, to a lesser extent, the history of the sciences.²² Bloor was concerned that previous attempts to explain the production and content of scientific knowledge were routinely divided, asymmetrically, into those representing true knowledge and those representing false knowledge. Following from this division, most commentators proceeded as though only the activities of scientists who had produced false (or flawed) knowledge could (or needed to) be explained sociologically; on the basis of the entry of egregious social factors such as ideology or other inappropriate attitudes and assumptions. These accounts are sometimes referred to as "sociologies of error," because they provide an explanation for why a research program failed.²³ For

21. See, e.g., BERNARD BARBER, *SCIENCE AND SOCIAL ORDER* (1952); JOSHUA BEN-DAVID, *THE SCIENTIST'S ROLE IN SOCIETY: A COMPARATIVE STUDY* (1971); WARREN HAGSTROM, *THE SCIENTIFIC COMMUNITY* (1965); ROBERT MERTON, *SCIENCE, TECHNOLOGY & SOCIETY IN SEVENTEENTH CENTURY ENGLAND* (Harper & Row 1970) (1938) [hereinafter MERTON, *SCIENCE, TECHNOLOGY, & SOCIETY*]; ROBERT MERTON, *THE SOCIOLOGY OF SCIENCE: THEORETICAL AND EMPIRICAL INVESTIGATIONS* (1973); *THE SOCIOLOGY OF SCIENCE* (Bernard Barber & Walter Hirsch eds., 1962). For an overview and discussion, consider DAVID HESS, *SCIENCE STUDIES: AN ADVANCED INTRODUCTION* 52-80 (1997), and BARRY BARNES, *SCIENTIFIC KNOWLEDGE AND SOCIOLOGICAL THEORY* 99-124 (1974).

22. In the historiography of the sciences, there have been several pioneering historians who examined the role of purportedly "non-scientific" influences in the work of some of the most distinguished early modern natural philosophers. For example, alchemical and natural magical ideas of action at a distance may have been influential on the gravitational work of Newton and its cautious reception by natural philosophers of the mechanical persuasion. See, e.g., ALAN DEBUS, *THE ENGLISH PARACELSAINS* (1966); BETTY DOBBS, *THE FOUNDATIONS OF NEWTON'S ALCHEMY, OR "THE HUNTING OF THE GREEN LYON"* (1975); BRIAN EASLEA, *WITCH-HUNTING, MAGIC AND THE NEW PHILOSOPHY: AN INTRODUCTION TO THE DEBATES OF THE SCIENTIFIC REVOLUTION 1450-1750* (1980); FRANCES YATES, GIORDANA BRUNO AND THE HERMETIC TRADITION (1964); Keith Hutchinson, *What Happened to the Occult Qualities in the Scientific Revolution?*, 73 *ISIS* 233 (1982); J. McGuire & P. Rattansi, *Newton and the "Pipes of Pan,"* 21 *NOTES & REC. ROYAL SOC'Y LONDON* 108 (1966). For an indication of the influence of the practical arts, the printing press, and broader social circumstances, consider ELIZABETH EISENSTEIN, *THE PRINTING PRESS AS AN AGENT OF CHANGE* (1979); MERTON, *SCIENCE, TECHNOLOGY, & SOCIETY*, *supra* note 21; PAOLO ROSSI, *PHILOSOPHY, TECHNOLOGY, AND THE ARTS IN THE EARLY MODERN ERA* (S. Attanasio trans., 1970); Boris Hessen, *The Social and Economic Roots of Newton's Principia*, in *SCIENCE AT THE CROSSROADS* (N. Bukharin ed., 1931).

23. See Ashplant & Wilson, *supra* note 12, at 257-59; R.W.K. Hinton, *Five Points About Whig History*, 9 *HIST. TODAY* 720, 722 (1959); John Pickstone, *Past and Present: Knowledge in the Practice of the History of Science*, 33 *HIS. SCI.* 203, 210 (1995). All of these works refer

historians, this type of explanation invariably involves retrospective assessment and implicitly access to the historically vindicated (or tested) correct outcome. As we shall see, this retrospective or whiggish dimension has become an important impediment to more symmetrically oriented approaches to expert evidence. Alternatively, accounts of scientific developments that are characterized as fundamentally correct require little, if any, social explanation. They purport to be fully explained, or more accurately determined, by the evidence or *Nature*.

In response, Bloor proposed to erode the division and require that both true and false approaches and conclusions should be explained utilizing the same type of accounting resources; that is, subject to the same type of causal explanation. This then, is the meaning of methodological symmetry. Bloor endeavors to explain the production and deployment of all knowledge using similar types of explanation. In this way the character of *Nature* cannot be used to arbitrate disagreement about its character.²⁴ Appeals to *Nature*, proper method, community acceptance, logic, or the evidence should not be invoked (or accepted by the analyst) to make some types of knowledge claim appear self-evident or obvious. The adoption of methodological symmetry directs attention to the reasons—frequently social reasons as knowledge and methods form part of technical and public disputes—surrounding the (often temporary) ascendancy of particular knowledge claims. For the purposes of this Article, the attribution of intrinsic value to some particular configuration of the evidence should not, without more, be understood as sufficient to explain the outcomes of trials and appeals. Indeed, this is so even before peculiarly legal and procedural factors are introduced into the matrix.²⁵

Use of the symmetry postulate, like contextual sensitivities derived from theoretical concerns about whiggism, requires specific attention to the situated production and reception of knowledge claims.²⁶ It

to asymmetrical tendencies in their discussions of whig-related historiography.

24. HARRY COLLINS, *CHANGING ORDER: REPLICATION AND INDUCTION IN SCIENTIFIC PRACTICE* (1985); BRUNO LATOUR, *SCIENCE IN ACTION: HOW TO FOLLOW SCIENTISTS AND ENGINEERS THROUGH SOCIETY* (1987).

25. Adversarial jurisdictions provide fertile conditions for symmetrical analysis because the various parties tend to develop antagonistic cases and evidence, anticipating and responding to the opposing case. In analyzing the competing knowledge claims, commentators should not simply accept one version without reflecting upon the influence of rules and procedures as well as the traditions of application and other, purportedly, extra-legal influences which are all constitutive of what is relevant and admissible evidence.

26. See generally MICHEL FOUCAULT, *THE ARCHAEOLOGY OF KNOWLEDGE* (Alan Sheridan Smith trans., 1972); MICHEL FOUCAULT, *THE ORDER OF THINGS: AN ARCHAEOLOGY OF THE HUMAN SCIENCES* (1973); MICHEL FOUCAULT, *POWER/KNOWLEDGE* 83-85 (Colin Gordon et al. trans., 1980); Donna Haraway, *Situated Knowledges*, in SIMIANS,

should encourage the analyst to embark on a comprehensive assessment of all the expert knowledges, involving an attempt to explain variations in the status of knowledge claims (the relevance and credibility of experts) as well as differing representations of their production, articulation, and recognition at particular times and places.²⁷

At this juncture it is appropriate to complete the theoretical introduction by problematizing the prevalent use of simplistic and abstract images of scientific knowledge and practice.²⁸ Adopting this course should serve to prevent analysts from rigidly adhering to controvertible *a priori* images of the sciences to inform (and reinforce) their asymmetrical analyses of true and false beliefs. Historians of science will presumably be conversant with a range of post-Kuhnian approaches to the sciences. Those approaches, however, may be less familiar to legal historians, legal commentators, and judges.²⁹ Though the following overview is not definitive, it does evince some sense of the dynamics and issues of particular concern to a large number of specialist historians and sociologists of science. To some extent it may operate as a historiographical minimalist position, regardless of

CYBORGS AND WOMEN: THE REINVENTION OF NATURE 183 (1991).

27. See, e.g., SOCIAL CONSTRUCTIONISM, DISCOURSE AND REALISM (Ian Parker ed., 1998); Callon & Latour, *supra* note 13; Collins & Yearly, *supra* note 13; Derek Edwards et al., *Death and Furniture: The Rhetoric, Politics, and Theology of Bottom Line Arguments Against Relativism*, 8 HIST. HUM. SCI. 25 (1995). Accepting that symmetry tends to encourage an epistemologically relativist stance should not prevent non-relativists from adopting it as an attitude or methodological stricture rather than as an epistemology. The analyst is free to prefer a particular position or accept a hierarchy, but in doing so should explain the reasons for that preference, without dismissing opposing claims as irrational. Relativism does not require that all knowledge be treated and conceived as equal in value, and this is where the contextual plausibility of accounts, shared assumptions among groups, and the ability to be able to *accomplish* things all come into play. For debates around images of technical efficacy, consider WIEBE BIJKER, OF BICYCLES, BAKELITE AND BULBS: TOWARD A THEORY OF SOCIOTECHNICAL CHANGE (1995); THE SOCIAL SHAPING OF TECHNOLOGY: HOW THE REFRIGERATOR GOT ITS HUM (Donald MacKenzie & Judy Wajcman eds., 1985).

28. In adopting the terms simplistic, idealized and idealistic, traditional and conventional as adjectives for a particular approach to scientific knowledge and practice, I have endeavored to produce an ideal-type or generalized representation to facilitate analysis. This image is deliberately similar to the more popular and general images of science captured by Michael's concept of science-in-general. That is, an image of science that lacks the messiness of local contingency characteristic of science-in-particular. See Mike Michael, *Lay Discourses of Science: Science-in-General, Science-in-Particular, and Self*, 17 SCI. TECH. & HUM. VALUES 313 (1992).

29. See generally David Caudill & Richard Redding, *Junk Philosophy of Science?: The Paradox of Expertise and Interdisciplinarity in Federal Courts*, 57 WASH. & LEE L. REV. 685 (2000); Gary Edmond & David Mercer, *Representing the Sociology of Scientific Knowledge and Law*, 19 SCI. COMM. 307 (1998). Significantly, legal commentators and judges seem to be more comfortable and sophisticated in their discussion and analysis of law and legal procedure which, unlike scientific evidence, they are less likely to reify. This will be illustrated in the example of Nobles and Schiff's account of some of the scientific evidence in the case of the Birmingham Six (Part III.B, *infra*) where their analysis of judicial review appears more sophisticated than the accompanying discussion of scientific evidence.

the particular epistemologies adopted by historians or attributed to scientists.³⁰

So, as a general overview,³¹ it is probably fair to say that many contemporary historians and sociologists of science would dismiss, as implausible, the existence of an historically stable, universal, prescriptive, and efficacious scientific method doctrine.³² On the basis of work by Kuhn, Polanyi, and Collins, formal education and socialization into a research tradition or research institution seem to be more important than knowledge of idealized or philosophical formulations.³³ Similarly, historical and empirical studies have been unable to locate a simple set of institutional or personal commitments or norms consistently adhered to by scientists. Norms conventionally associated with the sciences, such as disinterestedness and skepticism, are better understood as a complex moral language, susceptible to strategic deployment, rather than as a guide to scientific practice.³⁴ On examination, peer review and publication do not appear to operate in the rigorously skeptical manner typically attributed to them. Scientific journals, especially the most prestigious, have many, sometimes competing, obligations among, for example, rapid dissemination, maintaining broad interest, and providing the most accurate information possible. Submissions are not always thoroughly reviewed and are only rarely replicated. Perhaps it is not surprising to find that in undertaking review, scientists tend to be more sympathetic to individuals whose ideas and theories are

30. Many of the following characteristics are accepted by historians and sociologists with realist sympathies. See, e.g., STEPHEN COLE, *MAKING SCIENCE* (1992); Pickstone, *supra* note 23.

31. For a more detailed account of post-Kuhnian approaches to expert evidence, see Gary Edmond, *Judicial Representations of Scientific Evidence*, 63 *MOD. L. REV.* 216 (2000).

32. See, e.g., COLLINS, *supra* note 24; TREVOR PINCH, *CONFRONTING NATURE: THE SOCIOLOGY OF SOLAR-NEUTRINO DETECTION* (1986); *THE POLITICS AND RHETORIC OF SCIENTIFIC METHOD* (John Schuster & Richard Yeo eds., 1986). Notwithstanding a range of empirical studies and philosophical and sociological critiques, many scientists (and non-scientists) remain genuinely committed to highly abstract and often artificial images of scientific practice and progress. It is perhaps not surprising to find that individuals and institutions routinely relying upon the authority and credibility of scientific knowledge are often reluctant to acknowledge or explore the more contingent, tacit, political, competitive, and increasingly commercial dimensions of scientific organization and practice, or their influence on the constitution of scientific knowledge. This is particularly true of those, like judges, embroiled in public or professional controversies or science pedagogy.

33. THOMAS KUHN, *THE STRUCTURE OF SCIENTIFIC REVOLUTIONS* (1962); MICHAEL POLANYI, *PERSONAL KNOWLEDGE* (1958); COLLINS, *supra* note 24; see also Michael Mulkey & Nigel Gilbert, *Putting Philosophy to Work: Karl Popper's Influence on Scientific Practice*, 11 *PHIL. SOC. SCI.* 389 (1981).

34. Ian Mitroff, *Norms and Counter-Norms in a Select Group of the Apollo Moon Scientists: A Case Study in the Ambivalence of Scientists*, 39 *AM. SOC. REV.* 579 (1974); Michael Mulkey, *Interpretation and the Use of Rules: The Case of the Norms of Science*, in *SCIENCE AND SOCIAL STRUCTURE* 111 (T. Gieryn ed., 1980).

consonant with or reinforce their own.³⁵ In addition, there have been protracted difficulties surrounding attempts to locate and delimit identifiable communities of scientists—of particular relevance for admissibility decisions in legal settings—sharing assumptions and a range of commitments on all but the most fundamental issues.³⁶ The boundaries used to distinguish between the sciences, and demarcate scientific from non-scientific activity, seem more comprehensible when understood as flexible and strategically manipulated.³⁷ Perhaps the final insult to traditional models is the contention that historically it is extremely difficult to locate an homogeneous entity or pursuit that is easily characterized as *Science*.³⁸ This is not to suggest

35. DARYL CHUBIN & EDWARD HACKETT, *PEERLESS SCIENCE: PEER REVIEW AND U.S. SCIENCE POLICY* (1990); STEVE EPSTEIN, *IMPURE SCIENCE: AIDS, ACTIVISM AND THE POLITICS OF KNOWLEDGE* (1996); SHEILA JASANOFF, *THE FIFTH BRANCH: SCIENCE ADVISERS AS POLICYMAKERS* (1990); MARCEL LA FOLLETTE, *STEALING INTO PRINT: FRAUD, PLAGIARISM AND MISCONDUCT IN SCIENTIFIC PUBLISHING* (1992); Harry Collins, *Tantalus and the Aliens: Publications, Audiences and the Search for Gravitational Waves*, 29 *SOC. STUD. SCI.* 163 (1999); Gary Travis & Harry Collins, *New Light on Old Boys: Cognitive and Institutional Particularism in the Peer Review System*, 16 *SCI. TECH. & HUM. VALUES* 322 (1991). Contrast the legally oriented commentary by Elie Chan, *The "Brave New World" of Daubert: True Peer Review, Editorial Peer Review and Scientific Validity*, 70 *N.Y.U. L. REV.* 100 (1995).

36. Historically, notwithstanding the difficulty of ascertaining whether an assumption or technique was accepted by a particular *community* of scientists, its apparent general acceptance was frequently embraced by judges as a means of resolving admissibility decisions. The so-called general acceptance test, derived from the case of *U.S. v. Frye*, 293 F. 1013 (D.C. Cir. 1923), was criticized as being both excessively liberal and restrictive and was rhetorically compromised in U.S. federal courts by the decision in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993), which determined that the *Frye* test did not, on its own, survive the enactment of the *Federal Rules of Evidence* in 1975. See Gary Edmond, *Deflating Daubert: Kumho Tire Co. v. Carmichael and the Inevitability of General Acceptance (Frye)*, 23 *U.N.S.W. L.J.* 38 (2000); Paul Giannelli, *The Admissibility of Novel Scientific Evidence: Frye v. United States, a Half Century Later*, 80 *COL. L. REV.* 1197 (1980).

37. See *THE DISUNITY OF SCIENCE* (Peter Galison ed., 1996); THOMAS GIERYN, *CULTURAL BOUNDARIES OF SCIENCE: CREDIBILITY ON THE LINE* (1998); KAREN KNORR-CETINA, *EPISTEMIC CULTURES: HOW THE SCIENCES MAKE KNOWLEDGE* (1999); Karen Knorr-Cetina, *Scientific Communities of Transepistemic Arenas of Research? A Critique of Quasi-Economic Models of Science*, 12 *SOC. STUD. SCI.* 101 (1982).

38. A number of historians have emphasized the importance of adopting contextually appropriate terminology to study the history of the sciences and technology. Indeed among early modern historians there is a tendency to study natural philosophy, a more expansive set of practices and metaphysical commitments—including a theological sensitivity—influencing the production of knowledge about the natural world. See Cunningham & Williams, *supra* note 16; Andrew Cunningham, *Getting the Game Right: Some Plain Words on the Identity and Invention of Science*, 19 *STUD. HIST. & PHIL. SCI.* 365 (1988); Simon Schaffer, *Natural Philosophy, in THE FERMENT OF KNOWLEDGE: STUDIES IN THE HISTORIOGRAPHY OF EIGHTEENTH CENTURY SCIENCE* 55 (G. Rousseau & R. Porter eds., 1980); John Schuster & Graeme Watchirs, *Natural Philosophy, Experiment and Discourse: Beyond the Kuhn/Bachelard Problematic, in EXPERIMENTAL INQUIRIES 1* (Homer Le Grand ed., 1990); John Schuster, *The Scientific Revolution, in THE COMPANION TO THE HISTORY OF MODERN SCIENCE* 217 (G. Cantor et al. eds., 1990). Other historians (and engineers) have questioned traditional assumptions about the relationship between the sciences and technology. *E.g.*, DONALD MACKENZIE, *INVENTING ACCURACY: A HISTORICAL SOCIOLOGY OF NUCLEAR MISSILE GUIDANCE* (1990); ANDREW WEBSTER, *SCIENCE, TECHNOLOGY AND SOCIETY*

that historically there are no sciences, but to claim that they were not homogeneous, universal, or susceptible to a description that resembles the modern idea of *Science*. Instead there seems to be only an assortment of often disparate activities—the sciences (plural)—that are resistant to standardization through shared epistemologies, attitudes, or commitments other than at the most abstract level. Consequently, the sciences ought not to be conceived as a unified, peculiarly rational pursuit. None of the foregoing should be understood to imply that images of method, norms, or peer review are irrelevant to the practice, pedagogy, and rhetoric of the sciences. Rather they might be considered to be part of a flexible and frequently contested repertoire.³⁹ Indeed, this dimension converts accounts endeavoring to delineate scientific progress into much more complex and contextually attuned narratives, requiring additional levels of analysis concerned with the discursive construction of the *Scientific* and the *Natural*.

Assumptions about the existence and utility of attributes such as methods, norms, and other institutional arrangements purporting to guarantee the vitality and reliability of modern *Science* tend to discourage detailed investigation of expert knowledge claims contested in agonistic legal settings.⁴⁰ This is an unfortunate feature of much contemporary legal historiography and commentary, because litigation supplies circumstances in which attitudes to scientific (un)reliability and some of its dominant social manifestations are—somewhat inconsistently—strategically inscribed.⁴¹ Because adversarial legal forums and formal judgments endeavor to resolve disputes in ways conceived as procedurally regular and epistemologically robust, representations of practices and knowledges capable of being characterized as cogent and persuasive—frequently equated with reliability—assume considerable import.⁴² This should not, however,

(1991); Edwin Layton, *Mirror Image Twins: The Communities of Science and Technology in 19th-Century America*, 12 *TECH. & CULT.* 562 (1971); Walter Vincenti, *Technological Knowledge Without Science: The Innovation of Flush Riveting in American Planes, ca. 1930-ca. 1950*, 25 *TECH. & CULT.* 540 (1984).

39. NIGEL GILBERT & MICHAEL MULKAY, *OPENING PANDORA'S BOX: A SOCIOLOGICAL ANALYSIS OF SCIENTISTS' DISCOURSE* (1984); JONATHAN POTTER, *REPRESENTING REALITY: DISCOURSE, RHETORIC AND SOCIAL CONSTRUCTION* (1996).

40. Pierre Bourdieu, *The Force of Law: Toward a Sociology of the Juridical field*, 38 *HAST. L.J.* 814 (1987); Pierre Bourdieu, *The Specificity of the Scientific Field and the Social Conditions of the Progress of Reason*, 14 *SOC. SCI. INFO.* 19 (1975).

41. Processes of inscription in other contexts are examined in *INSCRIBING SCIENCE: SCIENTIFIC TEXTS AND THE MATERIALITY OF COMMUNICATION* (Timothy Lenoir ed., 1998); TIMOTHY LENOIR, *INSTITUTING SCIENCE: THE CULTURAL PRODUCTION OF SCIENTIFIC DISCIPLINES* (1997). See also STEVEN SHAPIN & SIMON SCHAFFER, *LEVIATHAN AND THE AIR PUMP: HOBBS, BOYLE AND THE EXPERIMENTAL LIFE* (1985).

42. See BARRY BARNES, *ABOUT SCIENCE* 72-89 (1985); WYNNE, *supra* note 6.

prevent the historian from inquiring into how images of reliability, relevance, progress, or methodological and normative propriety are actually contested and resolved during periods of controversy.⁴³

So, instead of attempting to define what *Science* is, I have suggested that there is no such epistemologically predicated entity. Rather, there are historically situated sciences, lacking universal or essential attributes. Abandoning the commitment to predetermined, highly artificial, and historically untenable attributes introduces tremendous complexity into the attempt to understand scientific controversy and its resolution. For example, rather than explain the outcome of scientific disagreement on the basis of the use of the *appropriate* method, the historian is empowered to question why and how a particular method, experiment, or interpretation was strategically deployed, contested, and eventually accepted as (in)correct. This complexity can be accentuated when superimposed upon adversarial legal contexts that not only facilitate but encourage the critical treatment of antagonistic knowledge claims.⁴⁴

Together, the contextual response to whiggism, the symmetrical approach to expert knowledges from the Strong Program in the sociology of scientific knowledge, and the deployment of less idealized models of the sciences operate to discourage the historian from evaluating the evidence produced during the process of the litigation on the basis of a *correct* value attributed at the close of proceedings.⁴⁵ They should prevent the introduction of attitudes and preempt recourse to values that were not available or were less certain at earlier times, if they are to be employed as a yardstick against which to measure earlier performances. Symmetry enables the analyst to reflect on how the reliability attributed to specific knowledge claims is discursively produced, contested, and accepted in specific settings,

43. COLLINS, *supra* note 24, has argued that the closure of scientific controversy usually reinforces, occasionally even codifies, what constitutes methodological propriety. Sociologists who have studied expert disagreement in legal settings have provided some indication of the manner in which judges and lawyers have influenced the development of forensic scientific techniques such as the protocols surrounding the use of DNA typing evidence. See EXPERT EVIDENCE, *supra* note 6; *Contested Identities*, *supra* note 6; Gary Edmond, *The Law-Set: The Legal-Scientific Production of Medical Propriety*, 26 SCI. TECH. & HUM. VALUES 191 (2001). Commentators such as Wynne have argued that controversy, like failed technological systems, provides a means of access to normally "hidden" social and institutional processes. See Brian Wynne, *Unruly Technology: Practical Rules, Impractical Discourses and Public Understanding*, 18 SOC. STUD. SCI. 147 (1988).

44. This is not intended to diminish the importance of traditions, institutions, professional commitments, or pervasive public beliefs. All of these aspects of knowledge and practice are, potentially, available to assist with both the strategic representation as well as the interpretation of expert knowledges.

45. Pickstone, *supra* note 23. This tendency may be of less consequence in short and more isolated proceedings. But even in these settings the meanings attributed to the evidentiary record are frequently dialectical, unfolding across the litigation. See BUTTERFIELD, *supra* note 4.

without diminishing the contextual use and influence of factors such as procedural rules, professional and institutional commitments, or resource implications. Symmetry and contextualism both privilege *process* and encourage commentators to explain how some knowledges and images of proper practice become authoritative or hegemonic.

As will be illustrated below, failure to adopt a more contextually and temporally sensitive symmetrical approach prevents the historian from accounting for a whole realm of nuances and debates around the meaning of expert knowledges and practices and their role in the maintenance of social order. In addition, failure to use such a symmetrical approach seems to render some of the highly sophisticated forensic techniques and rules employed by lawyers, judges, and scientists as perverse or irrational. To some extent this may appear curious: notwithstanding that commentators tend to attribute a *proper* value to purportedly clear and convincing scientific records that are eventually understood to constitute *reliable* evidence, a sizeable proportion of the litigation to be examined required decades to conclude.

Historians and legal commentators who assess performances on the basis of retrospective values, and even to some extent extra-legal practices, can avoid overlooking the dynamic features of adversarial litigation and the contingency of not only law and legal procedure but scientific and expert knowledges. To accomplish this, they should consider approaching scientific and expert disagreement on the assumption that experts, lawyers, and the parties are actively engaged in attempts to construct and represent knowledge in order to persuade—deliberately diminishing the value of some types of knowledge and expertise while simultaneously reinforcing or privileging others.⁴⁶ Adopting a more historically nuanced approach to expert knowledges examining their integration into legally conditioned narratives, and remaining sensitive to opposing cases, legal institutions, and constraints, shifts the analysis from a search for the most reliable knowledge and its earliest articulation in order to criticize individual

46. See *Contested Identities*, *supra* note 6, at 865; Gary Edmond, *Azaria's Accessories: The Social (Legal-Scientific) Construction of Guilt and Innocence*, 22 MELB. U. L. REV. 396 (1998); Steven Fuchs & Stephen Ward, *What Is Deconstruction, and Where and When Does It Take Place? Making Facts in Science, Building Cases in Law*, 59 AM. SOC. REV. 481 (1994). The strategic representations reflect the use of contingent and empiricist (or constitutive) repertoires described by GILBERT & MULKAY, *supra* note 39, at 56-57; HARRY COLLINS & TREVOR PINCH, *FRAMES OF MEANING* 188 (1982).

and institutional performances⁴⁷ to an examination of the processes involved in strategic productions, interaction and contestation.⁴⁸

The following examples have been drawn from the writings of historians, legal academics, and judges in order to demonstrate the broad utility of the methodological tools. Though not randomly selected, the examples do, nevertheless, exemplify a range of responses to scientific evidence that, on the basis of my own work and the work of others cited throughout the Article, appear to be ubiquitous.⁴⁹ These diverse examples illuminate theoretical limitations to the study of expert evidence in legal settings that occur in multiple jurisdictions, different types of litigation, and various time periods. The first of the studies examines medical evidence in a Victorian England sodomy trial. The second explores the last two appeals in the British case of the Birmingham Six. Finally, the third describes a recent series of trials and appeals associated with birth defects allegedly caused by a pharmaceutical product in the United States. Each of the studies examines a relatively detailed account of its subject; far more sophisticated than news-media and popular versions. Despite this, each will be shown to suffer from historiographical defects with serious implications for our understanding of the specific encounters and the more general issues of the operation of legal systems and the possibility of law reform predicated upon historical analysis.

II. MEDICAL DISAGREEMENT IN A NINETEENTH-CENTURY SODOMY TRIAL

The first example examines the uses of expert medical evidence in the historiography of the arraignment and trial of Ernest Boulton and Frederick Park in 1870-71. Focusing upon a recent paper by Charles Upchurch,⁵⁰ it is intended to demonstrate that simplistic or uncritical conceptualizations of expert evidence cannot capture the complex, situated negotiations and contests over medical knowledge

47. This amounts to a version of the whig quest for *origins*. BUTTERFIELD, *supra* note 4, at 43.

48. Several criminologists and legal theorists have explored the construction of criminality, particularly conviction, but they have focused their attention on the role of expert evidence in prosecuting rarely and in the process of exoneration very rarely. See, e.g., PAT CARLEN, *MAGISTRATES' JUSTICE* (1976); DOREEN MCBARNETT, *CONVICTION: LAW, THE STATE AND THE CONSTRUCTION OF JUSTICE* (1983); MIKE MCCONVILLE, ANDREW SANDERS & RICHARD LENG, *THE CASE FOR THE PROSECUTION: POLICE SUSPECTS AND THE CONSTRUCTION OF CRIMINALITY* (1991).

49. BENT FLYBJERG, *MAKING SOCIAL SCIENCE MATTER* 66-87 (2001) (discussing the benefits of selective case studies and the difficulties inherent in attempts to identify representative examples).

50. Charles Upchurch, *Forgetting the Unthinkable: Cross-Dressers and British Society in the Case of The Queen vs. Boulton and Others*, 12 *GEND. & HIST.* 127 (2000).

and propriety unfolding during the course of adversarial litigation. This discussion will show that an inability to appreciate the processes involved in the strategic articulation of legal rules and standards as well as contests over the meaning of evidence can produce tendentious interpretations.⁵¹ Upchurch's account of the litigation is particularly illuminating because he devotes, quite deliberately, little attention to the medical evidence, institutional traditions, and rules of practice and proof. Together, assumptions about the medical evidence and strategic decisions in relation to the conduct of the prosecution and defense lead Upchurch to postulate a conspiracy among the Crown, defense, and judges, to conceal the extent of cross-dressing and sodomy in Victorian England. As we shall see, that perspective is seriously compromised by combining knowledge of contemporary Victorian legal practice with a more diachronic assessment of the development of the expert opinion evidence presented during the trial.

In April 1870 Earnest Boulton and Frederick Park were arrested in drag. Having been under police surveillance for more than a year, they were finally arrested after entering a women's dressing room—a public indecency offense—at the Lyceum Theatre in London. Boulton and Park were detained in police custody, where they were examined by police surgeon Dr. Paul, ostensibly to ascertain their sex. Paul, whose interest and suspicions seem to have been aroused, extended the scope of his medical examination to include the anuses.⁵² Later, in his testimony, Paul described the defendants' penises as elongated and anuses as the most dilated he had ever encountered. The prisoners and a number of associates were sent to trial for buggery and conspiracy to commit buggery, in addition to the original public disorder offense. The bases for these charges were Paul's medical assessment and a range of other supposedly incriminating evidence,⁵³ including a considerable body of apparently

51. This discussion is adapted from a more detailed account: Gary Edmond, *The Importance of Being Ignorant: Interpreting Strategy, Legal Procedure and Medical Evidence in the Arraignment and Trial of R v. Boulton and Park and Others* (1871) (2001) (unpublished).

52. Harvey Sacks, *Notes on Police Assessment of Moral Character*, in *STUDIES IN SOCIAL INTERACTION* 280 (David Sudnow ed., 1972). Paul explained his curiosity on the basis of exposure to Professor Alfred Taylor's *MEDICAL JURISPRUDENCE* (1861).

53. In the historiography of Boulton and Park, most historians accept, some without explanation, that notwithstanding the eventual not guilty verdict, the accused were obviously sodomites. This assumption raises questions about the meaning of the contemporary behaviours, circumstantial evidence, and the ability of historians and various Victorian publics to sensitively interpret them. *See, e.g.*, Upchurch, *supra* note 50, at 137, 140. *See also* ROGER BAKER, *DRAG: A HISTORY OF FEMALE IMPERSONATION IN THE PERFORMING ARTS* 147-51 (1994); ROGER BAKER, *DRAG: A HISTORY OF FEMALE IMPERSONATION ON THE STAGE* 122-27 (1968); NEIL BARTLETT, *WHO WAS THAT MAN? A PRESENT FOR MR. OSCAR WILDE* 128-

intimate letters between Boulton and Park and others, some of which implied they were married to men, and the evidence of cross-dressing and associated antics including the sharing of beds.



Figure 1: Boulton and Park being led from Bow Street Police Station, 1870.⁵⁴

Anticipating that the trial would be sensational, like the arraignment, the Attorney-General was persuaded to undertake the prosecution in person. Before the trial he invoked his prerogative and dismissed the buggery charge on the basis of insufficient evidence. Instead, the accused were confronted with a charge of conspiracy to commit buggery among themselves and with others. The charge of

49 (1988); VERN BULLOUGH & BONNIE BULLOUGH, CROSS DRESSING, SEX AND GENDER 189-91 (1993); WILLIAM COHEN, SEX SCANDAL: THE PRIVATE PART OF VICTORIAN FICTION 129-73 (1996); H. MONTGOMERY HYDE, THE OTHER LOVE: AN HISTORICAL AND CONTEMPORARY SURVEY OF HOMOSEXUALITY IN BRITAIN 94-98 (1970); CYRIL PEARL, VICTORIAN PATCHWORK 105-20 (1972); RONALD PEARSALL, THE WORM IN THE BUD: THE WORLD OF VICTORIAN SEXUALITY 459-66 (1969); WILLIAM ROUGHEAD, BAD COMPANIONS 149-83 (1930); ALAN SINFIELD, THE WILDE CENTURY: EFFEMINACY, OSCAR WILDE AND THE QUEER MOVEMENT 6-8 (1994); JEFFREY WEEKS, AGAINST NATURE: ESSAYS ON HISTORY, SEXUALITY AND IDENTITY 46-47, 49-50 (1991); JEFFREY WEEKS, COMING OUT: HOMOSEXUAL POLITICS IN BRITAIN, FROM THE NINETEENTH CENTURY TO THE PRESENT 14-15, 37 (1977); Neil Bartlett, *Evidence: 1870*, in RECLAIMING SODOM 288-99 (J. Goldberg ed., 1994); Jim Davis, *Androgynous Cliques and Epicene Colleges: Gender Transgression on and off the Victorian Stage*, 26 NINETEENTH CENTURY THEATRE 50 (1998).

54. Originally published in the May, 20, 1871 issue of *The Day's Doings*, a weekly London news magazine.

conspiracy may have been more difficult to prosecute, but it was correspondingly more difficult to defend. During the trial the meaning of the medical evidence, the correspondence, and cross-dressing were actively contested. In a complicated exercise in social hermeneutics—there were few popular registers linking cross-dressing to sodomy—the Crown endeavored to link Paul’s medical opinion about the condition of the anuses to the other circumstantial evidence to encourage a particular apprehension of its meaning.

In response, eminent and highly competent defense lawyers endeavored to weaken the Crown’s links among the correspondence, cross-dressing, and sodomy. Drawing upon the defendants’ keen interest in amateur theatricals enabled lawyers to provide an innocent explanation for the many dresses and accessories as well as a means to explain the perpetuation of roles (such as those of husband and wife) from the stage into familiar correspondence.⁵⁵ In addition, the defense called four surgeons to contest Paul’s opinion and the attempt by the Crown to establish a link between the circumstantial evidence and the allegedly more objective opinions of its medical witnesses.

In his account of the case, Upchurch contends that the prosecution eventually failed because the Attorney-General and the defense conspired (notwithstanding the absence of written records) in order to contain knowledge of the extent of cross-dressing and sodomy in London, to protect both middle-class women and Britain’s international reputation. Commenting briefly though significantly on the medical opinions, Upchurch is critical of the Attorney-General for not preparing more incriminating medical evidence. This is important because the Crown acknowledged toward the close of the trial that it could no longer expect a jury to rely upon the incriminating medical evidence of Dr. Paul. So for Upchurch, and some others, the Crown’s medical evidence was inadequate, evidently before the trial commenced. Otherwise, criticism of the Attorney-General would be unfounded. Upchurch equates the evidence as it was understood toward the end of the trial with pre-trial perspectives and implies that any apparent, contemporaneous differences should be explained on the basis of a conspiracy among the legal protagonists. In other words, the Attorney-General was cognizant of the weaknesses in the

55. A number of social historians have examined connections between cross-dressing, same-sex relations, and the stage. See NETTA MURRAY GOLDSMITH, *THE WORST OF CRIMES: HOMOSEXUALITY AND THE LAW IN EIGHTEENTH CENTURY LONDON* 1-28 (1998); Lenard Berlanstein, *Breeches and Breaches: Cross-Dress Theatre and the Culture of Gender Ambiguity in Modern France*, 38 *COMP. STUD. SOC. & HIST.* 338 (1996); Davis, *supra* note 53, at 58-63; Lawrence Senelick, *Mollies or Men of Mode? Sodomy and the Eighteenth-Century London Stage*, 1 *J. HIST. SEXUALITY* 33 (1990).

Crown's medical evidence and made no serious attempt to augment it and thereby strengthen the prosecution's case.

If, however, the meaning of the medical opinion evidence can be shown to be, to a considerable degree, developed across the proceedings diachronically, as skillful Crown and defence barristers and surgeons strategically contested its meaning in relation to the other evidence, procedural constraints, and their own professional identities, then rather than ascribing bad faith or conspiracy to the Attorney-General and several others, we might obtain some insight into how the knowledge claims were developed and some were privileged. We might also want to examine some of the processes involved in the selection, presentation, emphasis, comparison, and simplification of the various expert opinions. Upchurch's account oversimplifies the dynamic and localized semantic contest surrounding the expert evidence, particularly the strategic production of medical propriety and reliable knowledge. A few examples will help to illustrate how the status of Paul's evidence was transformed from convincing to unreliable *during* the course of the proceedings, in a manner that tends to compromise Upchurch's conspiracy theory.

The following examples, extracted from a far more complex series of interactions, illustrate how the defense narrative was strengthened through strategic representations of medical evidence and practice.⁵⁶ It is not my intention to privilege or make *a priori* assumptions about the relative value of the evidence on the basis of the subsequent Crown concessions, but rather to attempt the more fundamental task of explaining how those conditions were brought about.

Paul's description of the defendants' anuses has already been recounted. Before the trial Paul's opinions were effectively black-boxed.⁵⁷ It was not until he was cross-examined that the basis of his

56. Although they are arguably not strictly relevant to the extent and significance of anal dilation, the meanings of cicatrix from previous operations, the presence of anal warts, and Park's possible scarring from syphilis of the anus were strategically contested in relation to Paul's performance and the conclusions of the other surgeons. Similar assumptions and debates persist in contemporary jurisprudence, particularly in relation to the alleged sexual assault of children. See *R v. J-LJ* [2000] S.C.R. 51 (Canadian Supreme Court case); IAN HACKING, *THE SOCIAL CONSTRUCTION OF WHAT?* 148-51 (1999) (discussing the Cleveland affair).

57. Richard Whitley, *Black Boxism and the Sociology of Science: A Discussion of the Major Developments in the Field*, in *THE SOCIOLOGY OF SCIENCE* 61 (P. Halmos ed., 1972). There was some preliminary cross-examination of Dr. Paul at the arraignment, but the presiding magistrate (Flowers), like Chief Justice Cockburn at the trial, thought that there was sufficient evidence to warrant proceeding with the prosecution. Both of these judges suggested that Paul's medical evidence was admissible. In the *Daily Telegraph*, the following determination was attributed to Flowers:

"My mind is made up. I hope I have not listened with any prejudice; I hope from the first my mind has been open; but I think on the evidence before me my duty is to send the matter for trial upon the graver charge and upon the misdemeanour

conduct and the propriety of his examination was strategically exposed (deconstructed), and developed in ways that were designed to make his earlier performance appear inadequate, especially when contrasted with the (known) performances of the more experienced and eminent surgeons. During cross-examination Paul conceded that he had examined the prisoners in a small, poorly lit room, behind a screen, for a period of one to three minutes each.⁵⁸ It was also made apparent that in extending his examinations beyond the genitalia, Paul had transcended his legal entitlement. It was no coincidence that the defence sought to ascertain from the other surgeons, particularly the other Crown witnesses, in what circumstances and for how long they had undertaken their examinations. Clark, Hughes, Johnston, and Harvey (for the defense) and Barwell, Gibson, and Taylor (for the Crown) had all conducted more leisurely and legally sanctioned medical examinations ranging from fifteen minutes to an hour while the prisoners were in custody awaiting trial.

In addition to the duration, the defense lawyers strategically differentiated the equipment produced by the other surgeons. At the trial, the defense emphasized their possession of a powerful magnifying glass and a speculum of the rectum and contrasted them to Paul's use of his naked eyes and hands. The contest over the adequacy and meaning of Paul's examination was intimately related, by the defense, to an image of the ideal or proper examination, in order to restrict what could be seen and consequently the conclusions that could be legitimately drawn.⁵⁹ Representations of the ideal examination combined extra-legal images with the opinions and performances of the other medical experts involved in the proceedings. Medical propriety and the nature of reality were simultaneously at stake—significantly, they were being negotiated in the courtroom *after* the surgeons had undertaken their examinations.⁶⁰

also, and that I must do that without bail." *The Boulton and Park Prosecution*, DAILY TELEGRAPH, May 31, 1870, at 5.

See also *The Boulton and Park Prosecution*, DAILY TELEGRAPH, May 12, 1871, at 6, for the comments attributed to Chief Justice Cockburn.

58. During re-examination, the Attorney-General endeavored to suggest that regardless of a range of apparent limitations to Paul's performance, the brief examination was sufficient to observe severe dilation.

59. See Lorraine Daston & Peter Gallison, *The Image of Objectivity*, 40 REPRESENTATIONS 81 (1992); Charles Goodwin, *Professional Vision*, 96 AM. ANTHROPOLOGY 606 (1994); Charles Goodwin, *Seeing in Depth*, 25 SOC. STUD. SCI. 237 (1995).

60. It is not my intention to suggest that courtroom contests are insulated from medical traditions and standards of practice, but rather to place emphasis on the circumstances attending litigation that tend to encourage and facilitate the strategic emphasis on differences and highly purposive representations of complex evidentiary matrices suited to the purposes of the case and sensitive to rules and traditions surrounding evidence and procedure. These considerations are irreducible to the conditions surrounding extra-legal scientific and medical controversies and on occasion even extra-legal standards or protocols. Perhaps a good example

The only surgeon, apart from Paul, to publicly acknowledge familiarity with a dilated anus was the Crown's witness, Professor Taylor. Thirty-eight years earlier Taylor had encountered what he identified as a cross-dressing sodomite, "Eliza" Edwards, during a routine, and therefore institutionally and professionally sanitized, autopsy. This encounter, in conjunction with his testimony indicating that there was nothing abnormal about the appearances of Boulton and Park, led the defense to endeavor to portray Taylor as not only the most authoritative expert witness but also the most experienced. In order to amplify these distinctions, the defense adverted to the fact that Paul had attended Taylor's lectures and deferred to Taylor's authority on forensic medicine. This enabled another strategic contrast in which the opinion of the eminent professor and author could be juxtaposed with that of his more modestly positioned student.⁶¹

In addition to the account of "Eliza" recorded in Taylor's celebrated text, a range of French and German publications were discussed during the course of proceedings. At the trial Paul and Gibson were in possession of medical works on sodomy from the Continent. None of the surgeons who testified professed much interest in the subject or prior interest in the Continental or local literature.⁶² The general English medical, legal, and social disdain enabled the defense to suggest Paul's possession of such works reflected an inappropriate interest in the subject. None of the foreign texts was admitted as evidence. The defense lawyers and Chief Justice Cockburn stressed the difficulty of testing or ascertaining the credibility of these foreign works and authorities in their construction of a preference for oral evidence in English law. As with his conduct of the examinations, Paul's possession of a work of questionable authority by Tardieu⁶³ could be portrayed as suspicious. In this context, where the possession of texts or even interest in some types of knowledge was vulnerable to aspersion, medical ignorance could be portrayed as a virtue. One interesting feature of the trial was that in the absence of extensive experience, all of the surgeons were willing and able, some in very confident terms, to provide testimony on the basis of

is the Australian High Court's response to expert medical opinion evidence in medical negligence litigation, *Rogers v. Whittaker* (1992) 175 C.L.R. 479. In *Rogers* the High Court explained that judges rather than doctors would determine whether specific actions and practices constituted medical negligence. See also *Rosenberg v. Percival* (2001) 75 A.L.J.R. 734.

61. See TAYLOR, *supra* note 52.

62. *E.g.*, WILLIAM ACTON, A PRACTICAL TREATISE ON DISEASES OF THE URINARY AND GENERATIVE ORGANS (IN BOTH SEXES) (2d ed. 1851).

63. AUGUST TARDIEU, ETUDE MÉDICO-LÉGALE SUR LES ATTENTANTS AUX MŒURS (1857).

extrapolation from their more regular physiological knowledge. Ignorance and uncertainty were being actively negotiated in a context acutely sensitive to (reasonable) doubt.⁶⁴

Upchurch's conspiracy theory is also weakened by a lack of attention to legal strategy, rules, and procedures. Indeed the alleged conspiracy might be more directly dismissed by interpreting the Attorney-General's trial strategy as defensible under the circumstances. For instance, Upchurch is highly critical of the Attorney-General for not calling some of the cross-dressing witnesses who appeared at the arraignment to testify at the trial. But this oversimplifies the proceedings. An arraignment is not a trial. There were different lawyers, judges, and charges at the arraignment, and the testimony of the cross-dressing witnesses appears to have been conceived as possibly prejudicial to the Crown case. Arguing that cross-dressing was widespread offered few advantages to the Crown if it could not be directly equated with sodomy. Acknowledging the prevalence of cross-dressing may have made the asserted relationship with sodomy appear less plausible. In turn, it might have compromised the defense's contention that the defendants' cross-dressing was generally associated with dramatic performances and only very occasionally manifested in what were characterized as public frolics and larks. A degree of convergence between the interests of the Crown and defense, as neither had much incentive to acknowledge the existence of a cross-dressing fraternity, and a restricted presentation of the evidence might be explained on the basis of the adversarial system requiring the parties to adduce their own evidence, in combination with a mutual desire to minimize the publicity associated with the prosecution. These convergences may have been accentuated by the concerns of the defendants' families, particularly Park's father who, as a senior English judge, presumably financed and may have influenced the conduct of the defense.

In addition, the decision to prosecute for conspiracy is potentially subversive to the analysis by Upchurch. Appropriately, Upchurch explains that the Attorney-General's decision to dismiss the charge (*nolle prosequi*) of sodomy thereby made "the 1871 case hinge on issues of conspiracy to commit sodomy, rather than sodomy itself."⁶⁵ However, in this context it is important to recognize that proving a charge of sodomy required direct evidence. Even uncontested

64. See BARBARA SHAPIRO, BEYOND REASONABLE DOUBT AND PROBABLE CAUSE: HISTORICAL PERSPECTIVES ON THE ANGLO-AMERICAN LAW OF EVIDENCE (1991); Charles Nesson, *Reasonable Doubt and Permissive Inferences: The Value of Complexity*, 92 HARV. L. REV. 1187 (1979); Note, *Reasonable Doubt: An Argument Against Definition*, 108 HARV. L. REV. 1955 (1995).

65. Upchurch, *supra* note 50, at 144.

medical evidence may have been insufficient to sustain the charge. It is possible that the Crown case for conspiracy may have been the strongest case that could be assembled against the defendants (and may have been stronger without medical evidence). Upchurch's analysis overlooks the possibility that an allegation of sodomy was probably easier to defend than an accusation of conspiracy, so the Attorney-General's decision to drop the charge may have actually strengthened the Crown's position relative to the defense.⁶⁶

Although not primarily oriented toward legal strategy or the medical evidence, the account by Upchurch exploits several simplified images to its advantage. His account ascribes a particular value to the medical evidence of Paul and his critics (Clark, Hughes, Johnston, Harvey, and variously Taylor and Gibson). Notwithstanding its simplicity, the value accorded to the medical evidence infuses it with explanatory purchase, as a purportedly stable post-trial value is transported to the pre-trial situation in order to make sense of the Attorney-General's decisions, which appear disingenuous and serve to fuel the conspiracy theory.

If, in contrast, the meaning of Paul's opinion and the propriety of his conduct was not fully pre-determined or reducible to clear and incontrovertible meanings among the surgeons (or lawyers) before the unfolding of the trial, then the version of the litigation provided by Upchurch is seriously flawed on the basis that the processes involved in the production of meaning are not considered. Before his cross-examination, Paul's observations might have been considered to have been sufficiently reliable to support the Crown's case. Explanation of how and why particular images and practices are presented and preferred and how some opinions are epistemologically compromised or effaced is absent. That approach elides the contingency linking professional, institutional, and social alignments, media reporting and related concerns about public

66. See VERN BULLOUGH, *HOMOSEXUALITY: A HISTORY* 31-45 (1979); COHEN, *supra* note 53, at 85; GOLDSMITH, *supra* note 55, at 31-48. Prior to the enactment of the Offences Against the Person Act, 1828, 9 Geo. 5, ch. 31, conviction required proof of both penetration and emission. The Act, designed to reduce the onerous burden of proof, required "proof of penetration only," § 18. See *The King v. Wiseman*, 92 Eng. Rep. 774 (1718); *Rex v. Robert Reekspears*, 168 Eng. Rep. 1296 (1832); *Rex v. Cozins*, 172 Eng. Rep. 1272 (1834); Arthur Gilbert, *Buggery and the British Navy 1700-1861*, 10 J. SOC. HIST. 72 (1976); A.D. Harvey, *Prosecutions for Sodomy in England at the Beginning of the Nineteenth Century*, 21 HIST. J. 939, 941 (1978). It should be acknowledged that the decision to use the charge of conspiracy was criticized by the defence barristers, and eventually the Chief Justice. See *Law Report*, TIMES, May 16, 1871, at 11; *The Boulton and Park Prosecution*, DAILY TELEGRAPH, May 16, 1871, at 3. Notwithstanding that criticism, neither Flowers (presiding at the arraignment) nor Chief Justice Cockburn (at the trial) were willing to grant bail or dismiss the prosecution. The decision to proceed in the attempt to prosecute the more difficult-to-prove charge of conspiracy may have also been a consequence of the difficulty in dropping a politically sensitive case. See *The Young Men in Women's Clothes*, TIMES, May 16, 1870, at 9.

interest, and epistemological relations to the performances of eminent surgeons, barristers, and judges (and jury). Instead of conspiracy or the attribution of a proper value to the evidence, a more diachronically nuanced and symmetrical assessment of the contest surrounding the medical evidence may help to transform our understanding of the case, particularly the interpretation of the actions of not only the Attorney-General, but the various judges, the lawyers and surgeons (especially Paul) involved throughout the proceedings, and the strategies they employed to prosecute and defend the case, their professions and themselves.

III. CHANGING ASSESSMENTS OF SCIENTIFIC EVIDENCE IN THE CASE OF THE BIRMINGHAM SIX

Readers will presumably be more familiar with the two remaining examples, though perhaps in alternative forms to the narrative possibilities suggested in this Article. Both cases feature more recent and more protracted contests over the reliability and legal instantiation of expert evidence and, based on the perspectives presented in almost the entire range of popular, media, scholarly, and judicial accounts, provide fairly robust targets for methodological reflection.

The Birmingham Six were a group of Irish men convicted of bombing two Birmingham pubs in 1974. The prosecution case consisted of three basic evidentiary dimensions: (1) confessions, (2) scientific evidence that two of the six had been handling nitroglycerine, and (3) other circumstantial evidence. After their conviction in 1975, the Six were refused leave to appeal.⁶⁷ It was not until a 1987 television documentary publicly challenged some of the tests used by a forensic scientist involved in the investigation and trial that the case was referred by the Home Secretary to the Court of Appeal.⁶⁸ In dismissing that appeal the Court concluded, somewhat infamously, that “the longer this case has gone on, the more this court has been convinced that the jury were correct.”⁶⁹ In 1990 the case was again referred to the Court, this time on the basis of apparent irregularities in the records of confessions. On this later occasion the convictions were quashed and the appellants released. The following discussion will draw upon some aspects of the judicial reversal and subsequent

67. Leave for appeal was denied in 1976 by the then Chief Justice, Lord Widgery.

68. Hugh Callahan, Patrick Hill, Robert Gerard Hunter, Richard McKenny, William Power & John Walker, 88 Crim. App. R. 40, 44 (C.A. 1989) [hereinafter *Callahan & Others*].

69. *Callahan & Others*, 88 Crim. App. R. at 47. The Court in 1987, like Lord Denning commenting on the civil action in 1980, had little idea what would happen in 1991.

academic assessment of the scientific evidence in what is now generally accepted as a notorious miscarriage of justice.⁷⁰

A. Judicial Representations of the Scientific Evidence

Being a judge requires institutional commitment and the use of tacit skills which often constrain actions and pronouncements.⁷¹ When judges try or review cases they are invariably involved in some type of historical exercise, but their attentions are rarely directed exclusively to historiographical subtleties as they are simultaneously concerned with rules, procedures, evidence and rights, limited resources, implications which may flow from any decision as well as maintaining the socio-political legitimacy of the legal system.⁷² Judges, therefore, have different concerns and constituencies from historians but are frequently engaged in an historical enterprise, trying to make sense of the past on the basis of an imperfect evidentiary record and through institutions and procedures that often restrict the ways in which evidence, and even what can count as evidence, can be accessed.⁷³

70. This account of the Birmingham Six is adapted from a more detailed discussion of the uses of scientific evidence in miscarriages of justice. See Gary Edmond, *Misunderstanding the Uses of Scientific Evidence in High Profile Criminal Appeals: The Social Construction of Miscarriages of Justice*, 22 OXFORD J. LEG. STUD. 53 (2002).

71. See MALCOLM FEELEY & E. RUBIN, *JUDICIAL POLICY MAKING AND THE MODERN STATE: HOW THE COURTS REFORMED AMERICA'S PRISONS* 204-52 (1999); STANLEY FISH, *DOING WHAT COMES NATURALLY: CHANGE, RHETORIC AND THE PRACTICE OF THEORY IN LITERARY AND LEGAL STUDIES* (1989); MICHAEL POLANYI, *PERSONAL KNOWLEDGE* (1958); Duncan Kennedy, *Freedom and Constraint in Adjudication: A Critical Phenomenology*, 36 J. LEG. EDUC. 518 (1986).

72. The use of historical evidence, particularly in the realm of constitutional interpretation, has spawned a considerable literature. See, e.g., Richard Epstein, *History Lean: The Reconciliation of Private Property and Representative Government*, 95 COL. L. REV. 591 (1995); Martin Flaherty, *History "Lite" in Modern American Constitutionalism*, 95 COL. L. REV. 523 (1995); William Forbath, Hendrik Hartog & Martha Minnow, *Introduction: Legal Histories from Below*, WISC. L. REV. 759 (1985); Peter Hoffer, *Blind to History*, 23 RUTGERS L.J. 271 (1992); N. Hull & Peter Hoffer, *Historians and the Impeachment Imbroglia: In Search of a Serviceable History*, 31 RUTGERS L. J. 473 (2000); Peter Irons, *Clio on the Stand: The Promise and Perils of Historical Review*, 24 CAL. W. L. REV. 337 (1988); Alfred Kelly, *Clio and the Court: An Illicit Love Affair*, 1965 SUP. CT. REV. 119; Buckner Melton, *Clio at the Bar: A Guide to Historical Method for Legists and Jurists*, 83 MINN. L. REV. 377 (1998); Wendie Schneider, Note, *Past Imperfect*, 110 YALE L.J. 1531 (2001); Suzanna Sherry, *The Indeterminacy of Historical Evidence*, 19 HARV. J.L. & PUB. POL'Y 437 (1995); Cass Sunstein, *The Idea of a Useable Past*, 95 COL. L. REV. 601 (1995); William Wiecek, *Clio as Hostage: The United States Supreme Court and the Uses of History*, 24 CAL. W. L. REV. 227 (1988).

73. Indigeneous land claims and heritage protection litigation provide instructive examples of such difficulties. Judges are often required to hear and evaluate various types of aboriginal, anthropological, historical, documentary, archaeological, linguistic, and genealogical evidence. For some indication of the difficulties, consider the Canadian case of *Delgamuukw v. The Queen*, [1991] D.L.R. 185, *Delgamuukw v. B. C.*, [1997] S.C.R. 1010, and the Australian case of *Yorta Yorta Aboriginal Cmty. v. Victoria*, (1998) F.C.A. 1606, (2001) F.C.A. 45. See also Michael Asch & Catherine Bell, *Definition and Interpretation of Fact in Canadian Aboriginal Title Litigation: An Analysis of Delgamuukw*, 19 QUEEN'S L.J. 503 (1993).

The case of the Birmingham Six provides an instructive example of some of the techniques employed by judges to rationalize decision-making where they are confronted with an extensive evidentiary record developed over a considerable period of time. Focusing on the final two appeals affords an opportunity to examine how two senior appellate courts came to very different conclusions, one dismissing and one upholding the appeal. The following discussion examines differences between the representation of scientific evidence in the two appellate judgments. The tremendous publicity generated by the campaign to free the Birmingham Six and serious concerns with sections of the criminal justice system, including the Court of Criminal Appeal, might provide some degree of social explanation for the changing legal and evidentiary perspectives though, perhaps not surprisingly, such considerations are largely absent from the purportedly evidence-based and technical legal rationalizations produced by judges.⁷⁴

Notwithstanding allegations of forced confessions and police brutality, the earliest sustained public response to the convictions was directed toward the scientific tests undertaken by the Home Office forensic scientist, Dr. Skuse. During the trial Skuse had testified that, on the basis of his results from a Griess test, two of the defendants' hands had tested positive for nitroglycerine (NG). Later, another forensic scientist, Dr. Drayton, using the far more sensitive gas chromatography/mass spectroscopy (GC/MS) technique, confirmed Skuse's findings for one of the samples. This second result led Skuse to increase his confidence in the presence of NG from 99 to 100 percent.

Doubt surrounding the specificity of Skuse's Griess tests was resolved by the Court during the 1987 appeal, on the basis that even if the tests were not specific for NG, they had identified some substance, and the possibility that it was NG was strongly supported by the result of the (approximately 1,000 times) more sensitive GC/MS technique. Perhaps not surprisingly, during that appeal the appellants' counsel also challenged Drayton's GC/MS result.

74. Accepting that the preferred outcome may drive or even occasionally determine the reasoning and what constitutes relevant and reliable evidence, the evidence and the rules and their sui generis interactions with the production and assessment of evidence should not be overlooked. Approaching the evidence, procedure, rules, and traditions as well as the preferred outcome without pre-determined commitments may enable the construction of plausible explanations that provide some indication of the nature of the legal-judicial closure of controversy and the strategic use of expert evidence. In the case of the Birmingham Six, there was protracted and highly critical media coverage of the Court of Appeal and individual judges, which led the Chief Justice to resign prematurely, politicians to re-refer the case to the Court of Appeal, and public inquiries that all eventually resulted in the reform of criminal justice institutions and practices.

Accepting that Drayton's positive GC/MS result was expressed differently in a variety of contexts—her initial finding of “possible ng present” in the sample was used collaboratively, with Skuse's then reliable evidence, at the trial—when Skuse's testimony and techniques were impugned the GC/MS result assumed center-stage.⁷⁵ This remained as some type of corroborative interpretation although the balance between the findings of Skuse and Drayton was now reversed. During the 1987 appeal, scientists for the defense produced a number of possible explanations for Drayton's “positive” GC/MS result that were not consistent with the presence of NG. These suggestions were rejected by Drayton and subsequently by the Court of Appeal as not worthy of “sensible rational consideration.”⁷⁶ Upholding the convictions, the Court concluded that “the presence of nitroglycerine on Hill's left hand, for which there could be no innocent explanation, was proved beyond reasonable doubt.”⁷⁷

The interpretation of the NG evidence was quite different in the changing social and evidentiary circumstances leading to the second appeal in 1991. The circumstances were different, not because there were profoundly *new* perspectives on the explosives evidence, but as we shall see, many of the previously dismissed concerns were resurrected in the later appeal. What counts as *new* or *fresh* evidence can be invested with considerable legal import.⁷⁸ The major differences between the appeals concerned the concurrent investigation by Justice May (the May Inquiry)⁷⁹ into other contemporaneous terrorist convictions which raised serious doubts over the reliability of some of the scientific evidence; evidence suggesting serious anoma-

75. For an indication of the changing status, or tractability, of knowledge across contexts, consider JEROME RAVETZ, *SCIENTIFIC KNOWLEDGE AND ITS PROBLEMS* (1971). For a study of reification of scientific “knowledge,” see BRUNO LATOUR & STEVE WOOLGAR, *LABORATORY LIFE: THE SOCIAL CONSTRUCTION OF SCIENTIFIC FACTS* (1979).

76. Richard McIlkenny, Patrick Hill, William Power, John Walker, Robert Gerard Hunter & Hugh Callaghan, 93 Crim. App. R. 287, 302 (C.A. 1991) [hereinafter *McIlkenny & Others*].

77. *Id.* Some of these extracts pertaining to the 1987 appeal are taken from the 1991 judgement. Despite widespread commitment to the existence of an accountable and publicly accessible criminal justice system, the trial and first appeal were not fully reported on the public record.

78. These classifications do not always reflect a literal chronology as much as a chronology of judicial appropriation and use. Despite their description as entirely new or fresh, that characterization is frequently a consequence of a novel use in judicial rationalizations. It is here that, despite the legal system's purported commitment to openness and access, trials and the majority of judicial rationalizations routinely benefit from the restricted number of readers and a general commitment, among legal commentators and in media reports, to idealized images of the sciences.

79. SIR JOHN MAY, INTERIM REPORT ON THE MAGUIRE CASE (1990) [hereinafter *MAY, INTERIM REPORT*]. SIR JOHN MAY, FINAL REPORT: RETURN TO AN ADDRESS OF THE HONOURABLE THE HOUSE OF COMMONS DATED 30 JUNE 1994, FOR A REPORT INTO THE CIRCUMSTANCES SURROUNDING THE CONVICTIONS ARISING OUT OF THE BOMB ATTACKS IN GUILDFORD AND WOOLWICH IN 1974 (1994) [hereinafter *MAY, FINAL REPORT*].

lies in the police record and description of confessions that the Six had always insisted were forced or fabricated; and even greater media, political, and popular concern at what was now widely accepted as a serious miscarriage of justice.

The case of the Birmingham Six was referred to the Court of Appeal, for a second time, ostensibly on the basis of the results of document analysis (ESDA) of the records of confession. In short, it was alleged that investigating police had fabricated the confessions and/or used brutality to extract them. Specialist examination of the confessions revealed a number of discrepancies between some of the documents central to the prosecution and statements from police about the circumstances of their production. This evidence was extremely important for three reasons. First, the defendants had steadfastly maintained that they were beaten and forced to confess, having unsuccessfully and controversially attempted to recover damages for some of their injuries through civil action against police and prison warders while in prison. Second, because the police and judges had simply dismissed these allegations—during one appeal Lord Denning had declared, “This is such an appalling vista that every sensible person in the land would say: ‘It cannot be right that these actions should go any further’”—any evidence tending to support their allegations would be extremely important.⁸⁰ The judges who heard the 1987 appeal described the confessions as “the true foundation of the prosecution case.”⁸¹ Finally, because the scientific evidence pertaining to the explosives had been weakened in the 1987 appeal—though significantly not below the threshold necessary to sustain the convictions—if other elements of the Crown’s case were also compromised, then it would be more likely that the Court might find the convictions “unsafe and unsatisfactory.”⁸²

The ESDA evidence, and the perhaps begrudgingly decision to allow the appeal in the face of considerable public pressure, led the Court to produce a judgment quashing the convictions.⁸³ To appear institutionally consistent, the second judgment had to be capable of

80. *McIlkenney v. Chief Constable of West Midlands Police Force & Another*, 2 All E.R. 227, 240 (1980); *Hunter v. Chief Constable of the West Midlands Police & Others*, A.C. 529 (1982).

81. *Hugh Callahan & Others*, 88 Crim. App. R. 40, 42 (C.A. 1989).

82. Under Section 2 of the Criminal Appeal Act (1968), judges were empowered to quash convictions considered “unsafe and unsatisfactory” or where there had been a mistaken decision on a question of law or a material irregularity in the trial. The meaning of the phrase “unsafe and unsatisfactory” was developed across decades of criminal appeals. See RICHARD NOBLES & DAVID SCHIFF, *UNDERSTANDING MISCARRIAGES OF JUSTICE: LAW, THE MEDIA, AND THE INEVITABILITY OF CRISIS* (2000); ROSEMARY PATTENDEN, *ENGLISH CRIMINAL APPEALS 1844-1994* (1996).

83. Questions of what constitutes evidence and how it should be interpreted are, for appellate judges, inseparable from socio-political and institutional concerns.

eroding the apparent reliability of the NG evidence, in a manner that insulated the earlier 1987 decision from retrospective recrimination. To this end, the Court of Appeal emphasized differences between the evidentiary matrix available in 1987 and that developed by 1991 so that the earlier commitment to an array of evidence upholding the conviction to the “beyond reasonable doubt” standard could be infused with reasonable doubt, ostensibly on the basis of the *new* evidence. As we shall see, the attempt to achieve this result produced an arguably weak justification in relation to the NG evidence, because the 1991 Court accepted some of the arguments rejected in the earlier appeal. Together, the appeals can be interpreted to suggest complex linkages between the interpretation of different types of evidence, institutional politics (particularly the maintenance of judicial credibility), and sensitivities to other purportedly extra-legal factors—those that are usually understood as extrinsic to the production and assessment of scientific knowledge.

During the 1987 appeal, Skuse’s methods and results had received considerable criticism from the defense. As we have seen, however, in the judgment the Court accepted that Skuse had detected some substance which they combined with Drayton’s GC/MS result to provide conclusive evidence for the presence of NG. In the 1991 appeal, one of the scientists involved in the May Inquiry (Dr. Lloyd) offered a potential explanation for Skuse’s positive Griess test results, as well as the trouble he had originally encountered cleaning his testing equipment after the positive results, that could be consistent with the innocence of the defendants. Lloyd suggested that the soap used to clean the testing bowls may have led inconsistently to nitrite contamination, such that if “he [Skuse] rinsed the first bowl twice, and the second bowl five times, he would get nitrite contamination in the first bowl but not the second.”⁸⁴ The detergent contamination theory, which in this new narrative was leading inexorably to a re-examination of the standing of the GC/MS result, also provided an explanation for why there had been no positive thin layer chromatography (TLC) results from the samples. Earlier recourse by the Court and the Crown to the rapid decomposition of NG was implicitly abandoned.⁸⁵ Notwithstanding the fact that Skuse had not used a nitrite-based soap, the Court found—in

84. *McIlkenny & Others*, 93 Crim. App. R. 287, 299 (C.A. 1991).

85. TLC was a technique more sensitive than the Griess test but less sensitive than GC/MS. The Crown had previously explained the negative TLC result on the basis that NG decomposed rapidly. This meant that it was possible to produce a positive Griess result, and subsequently, a negative TLC result and under some conditions still detect NG using GC/MS. This provides an illustration of how according to the Duhem-Quine thesis, modification in frameworks can accommodate results and data.

conjunction with a vague allusion to scientific developments between 1974 and 1991—that “Skuse’s evidence at the trial based on his Griess test is now shown to be unreliable.”⁸⁶ Having rendered Skuse’s evidence unreliable, the Court was left with Drayton’s positive, but now uncorroborated, GC/MS finding as the only remaining NG evidence.

Because of its centrality in the 1987 appeal, it was institutionally important for the 1991 Court to address Drayton’s single positive result if it intended to quash the convictions. The Court appeared eager, in its 1991 judgment, to locate and endorse arguments which compromised the reliability of the GC/MS result. Drawing upon one new finding, and presumably influenced by the changing evidentiary and social context, the 1991 Court determined that even the GC/MS result should now be considered unreliable. In her original testing of the sample, Drayton had identified a small peak at the appropriate retention time for NG. She testified that there was no other substance which produced such a peak under similar circumstances. Once again the issue of the specificity was challenged. Lloyd testified that during some tests, again for the May Inquiry, he had identified an unknown substance from the hand of a smoker who “clearly had not been handling explosives” purporting to show “a distinct peak on 46 at about the same retention time as nitroglycerine.”⁸⁷ With no one able to identify the substance, the judges accepted the existence of this other entity, explaining that the test chart was there “for all to see.”⁸⁸ Confronted with this new evidence, in conjunction with an ever-changing case, Drayton modified (perhaps sensibly renegotiated) her former position, now recognizing an alternative to NG as a possibility.⁸⁹ An unidentified substance that produced a peak “at *about* the same time” as NG, observed using different equipment, now led to the rejection of the previously strongest element of the scientific case against the accused.⁹⁰

The fact that the Crown did not contest the ESDA evidence on appeal and, given the history of the litigation, its inability to defend inconsistencies in the police testimony, led to this new evidence being effectively accepted without challenge. Evidence of police misrepresentation and fabrication were sufficient to introduce serious doubts into review of the safety of the original convictions, even if the ESDA findings did not translate into evidence directly

86. *McIlkenny & Others*, 93 Crim. App. R. at 300.

87. *Id.* at 302.

88. *Id.*; see also Daston & Gallison, *supra* note 59.

89. *McIlkenny & Others*, 93 Crim. App. R. at 302-03.

90. Despite claims to the contrary, it is possible that the unknown substance was NG.

supporting every element of the defense narrative. For example, the ESDA evidence was used to suggest serious irregularities in records of interviews, even some that recorded no admissions, and some of the interviews now alleged to have been partially fabricated were originally alleged not to have occurred. These concerns arise even before the issues of the general reliability of the ESDA technique and its specific application are considered. Perhaps ironically, because there was no cross-examination from the Crown, the evidence was accepted ostensibly as reliable and invested with an epistemic warrant in some ways equivalent to that of the original NG evidence before and immediately after the trial.⁹¹

Based on a very succinct overview of Skuse's evidence, the GC/MS result, and the fresh evidence about the confessions, it is worth noting that the changing interpretation of the opinions of Skuse and Drayton seem to be difficult to explain entirely on the basis of the NG evidence, even when combined with claims about interim advances in related sciences. Presumably, other evidence and social factors, particularly the confessions and the popular calls for the prisoners to be released, influenced the interpretation and representation of the evidence. The difficulties encountered by the Court and police force in dealing with the ESDA evidence seem to have influenced the interpretation of the previously highly incriminating NG evidence. The Court's explanation of the reversal of its interpretation of the NG evidence between 1987 and 1991 may not appear entirely convincing. Here, not only is the evidence unable to speak for itself, but very similar NG evidence is interpreted quite differently on separate occasions, and this interpretation appears to be inextricably linked to a range of subsidiary social and epistemological changes—such as the contextually devastating impact of the ESDA findings—which are not explicitly articulated in the judicial rationalization.

This section is not intended as a critique of judges and judging, but rather to provide some sense of the historiographical approaches to expert evidence applied by judges. It provides some indication of the manner in which decision-making can be linked to the maintenance

91. There is often a conspicuous similarity in the manner in which confidence is vested in some knowledges to prove guilt and later in other knowledges to quash the convictions. The later tendency is often equated with factual innocence, especially in popular accounts. In practice this often leads to a type of reverse asymmetry, where new evidence or techniques are accepted, even if contested, and used to quash convictions. In such cases we often encounter recourse to differing burdens of proof—between prosecution and appeal—and their judicial articulation. Consider, for example, the use of DNA typing evidence by those associated with the Cardozo Law School Innocence Project. JIM DWYER, PETER NEUFELD & BARRY SCHECK, *ACTUAL INNOCENCE: FIVE DAYS TO EXECUTION AND OTHER DISPATCHES FROM THE WRONGLY CONVICTED* (2000); Geoffrey Rapp, *DNA's Dark Side*, 109 *YALE L.J.* 163 (2000) (book review).

of judicial credibility and the legitimacy of legal institutions.⁹² Judicial assessments of evidence are not purely epistemological. In the case of the Birmingham Six there was tremendous media, public, and political pressure on the Court to acquit. Accepting that it may be difficult to reconstruct the extent of social and institutional influences upon judges, they presumably shape the articulation of a range of evidentiary and procedural interpretations. However, notwithstanding the explanations that commentators (such as myself) might attribute, judges are peculiarly reluctant to account for their decisions on the basis of social pressures and almost without exception revert to legalistic explanation. Drawing on (and perpetuating) the existence of a law/fact dichotomy, they tend to explain their decisions on the basis of law and/or evidence. As in traditional models of *Science*, reified images of *Law* and *Fact* operate as if they expunge egregious social contaminants from judicial decisions and fact-making.

As we have seen, courts are sensitive to the findings of earlier tribunals and not every explanation (or reinterpretation) of the evidence will appear convincing. Frequently this will require the production of evidence capable of being plausibly represented as new or fresh. The ability to deem some evidence as new offers the institutional advantage of absolving judges and juries of responsibility for earlier inconsistent decisions because they were not apprised of a more complete evidentiary record. The earlier performances can be defended, or excused, because it is the new evidence that requires the later court to reverse. Judges vigilantly guard the ability to determine what constitutes new evidence and have developed a jurisprudence around the subject, ostensibly to prevent counsel from withholding elements of a case in order to have a second attempt on appeal. In this way judicial historiography is highly selective. Admissible facts are inextricably bound to legal, procedural, and institutional considerations such as the need to publicly rationalize decision-making that frequently involves not

92. For example, earlier decisions are defended or insulated from subsequent reversals and in most notorious miscarriage of justice cases, responsibility is ordinarily apportioned to witnesses, police, and scientists rather than judges or juries. Because of the constitutional centrality of the jury, appellate judges are generally reluctant to overturn jury convictions, though this does not prevent them from doing so. When judges reverse jury verdicts they sensitively manage the explanation: Either the constitutionally entrenched function of the jury is downplayed or differences between the trial and appeal, such as the production of new evidence, are used to explain the reversal. There is a general reluctance among the judiciary to imply that a central component in modern adversarial legal systems—the jury—is irrational or seriously deficient, perhaps with some minor exceptions such as in extremely protracted complex cases or in occasional extra-curial pronouncements.

only the resolution of expert disagreement but occasionally the reversal of earlier resolutions.⁹³

B. Legal Commentary: (Mis)Understanding Miscarriages of Justice

Judicial accounts of law and evidence are not invariably accepted, especially by non-legal commentators. Where an appellate court explains that the quashing of a criminal conviction is the legal equivalent to a finding of not guilty, non-legal discourses frequently equate them with innocence.⁹⁴ A finding of not guilty often leads to a reinterpretation of the entire proceedings from a perspective that privileges the accused's factual innocence and any evidence corroborating that interpretation. Usually this involves, as in the examples of Boulton and Park and the Birmingham Six, an evaluation of all the earlier expert evidence from the standpoint of the evidentiary record available and accepted at the conclusion of proceedings. Adopting this whiggish approach leads historians and commentators to embark on retrospective recriminations. Whereas judges maintain an institutional interest in emphasizing differences between proceedings with different outcomes, we will now turn to an example of how historians and legal commentators sometimes emphasize similarities—that the evidence at the outcome was ostensibly available at the trial—in order to demonstrate flaws in legal processes.⁹⁵ This type of account is designed to demonstrate how, and sometimes why, the legal system is considered to distort scientific evidence and regularly fails to identify the appropriate value of evidence.⁹⁶

93. Sir John May provides some indication of diachronic shifts in the evidence against the accused in the case of the Birmingham Six. He explains: "In some of the representations made to me, reliance has been placed on some abstract notion of justice, by reference to which it has been argued that the convictions of the Guildford Four were clearly wrong and indeed should never have occurred. Such reliance is impractical." MAY, FINAL REPORT, *supra* note 79, at 298.

94. For example, a judge presiding at the final (successful) appeal in the *Chamberlain* (dingo baby) case, *Re Conviction of Chamberlain* (1998) 93 F.L.R. 239, 254 (Kearney J.), explained: "It is rarely that a criminal trial positively establishes the innocence of an accused person. If it does so, it does so by accident. . . . It is not the court's function to establish innocence because, in the absence of a conviction, innocence is presumed: no finding is required."

95. The strategic emphasis on similarities and differences has resonances with Schuster's discussion of historians' commitment to either "revolution" or "continuity" in accounts of early modern natural philosophy on the basis of the selection and emphasis of their materials, *see* Schuster, *supra* note 38. For more abstract treatment consider the discussion of similarity/difference relationships in Collins's discussion of replication, COLLINS, *supra* note 24; *see also* GEOFFREY BOWKER & SUSAN STARR, *SORTING THINGS OUT: CLASSIFICATION AND ITS CONSEQUENCES* (1999); LUDWIG WITGENSTEIN, *PHILOSOPHICAL INVESTIGATIONS* (1963). As we shall see in the following section, examining the Bendectin litigation, emphasis on differences between trials, diachronic changes across litigation, and the description of evidence as significant, similar, or different are all unavoidably linked to images of scientific knowledge and practice.

96. *See* Edmond, *supra* note 43.

An example of this practice, in an otherwise legally and theoretically informed discussion of miscarriages of justice, is a recent account of scientific evidence in the case of the Birmingham Six produced by Richard Nobles and David Schiff.⁹⁷ Notwithstanding a commitment to some of the critical approaches to the sciences developed in this Article,⁹⁸ their account reveals a lacuna between theory and practice that is neatly captured in the treatment of the testimony of the defense scientist Dr. Black during the Birmingham Six trial.

Nobles and Schiff refer to Black's testimony in the following capacities:

Dr Skuse's evidence was challenged by Dr Black, a scientist and former chief inspector of explosives at the Home Office. He questioned the specificity of the Griess test, and was not satisfied with Dr Skuse's *failure to reproduce the positive results* when the samples in question were *subjected to other tests*.⁹⁹

97. Other examples of recent work on miscarriages of justice sharing similar conceptual constraints, even though not of an autopoietic persuasion, include CRIMINAL JUSTICE UNDER STRESS 5 (Eric Stockdale & Silvia Casale eds., 1992); RONALD HUFF ET AL., CONVICTED BUT INNOCENT: WRONGFUL CONVICTION AND PUBLIC POLICY (1996); JUSTICE IN ERROR (Clive Walker & Keir Starmer eds., 1993); JUSTICE, SCIENCE AND THE ADMINISTRATION OF JUSTICE (1991); LEGAL ACTION GROUP, PREVENTING MISCARRIAGES OF JUSTICE: A SUMMARY AND INITIAL RESPONSE TO THE REPORT OF THE ROYAL COMMISSION ON CRIMINAL JUSTICE (1993); MANSFIELD & TONY WARDLE, PRESUMED GUILTY: THE BRITISH LEGAL SYSTEM EXPOSED (1994); MISCARRIAGES OF JUSTICE: A REVIEW OF JUSTICE IN ERROR (Clive Walker & Keir Starmer eds., 1999); CHRIS MULLIN, ERROR OF JUDGEMENT: THE TRUTH ABOUT THE BIRMINGHAM BOMBS (1986); VISCOUNT RUNCIMAN, THE ROYAL COMMISSION ON CRIMINAL JUSTICE: REPORT (1993); BOB WOFFINDEN, MISCARRIAGES OF JUSTICE (1987); David Bell, *The Expert Misleads. The Court Follows*, 27 AUST. J. FOR. SCI. 59 (1995); David Bell, *Whose Accountability, Judges or Experts?* 2 AUST. J. FOR. SCI. 74 (1994); S. Edwards, *From Scapegoats to Sacrificial Lambs: The Guildford Four Affair*, 1989 NEW L.J. 1449 (Oct. 1989); Michael Kirby, *Miscarriages of Justice – Our Lamentable Failure?*, DENNING L.J. 97 (1991); Jeffrey Miles, *Forensic Science: In the Spotlight or Under the Microscope*, 3/4 AUST. J. FOR. SCI. 3 (1991); Michael C. Porter, *The Evidence of Experts*, 27 AUST. J. FOR. SCI. 53 (1995); Alec Samuels, *Forensic Science and Miscarriages of Justice*, 34 MED. SCI. & LAW 148 (1994); Stuart Tipple, *Forensic Science: The New Trial by Ordeal?*, LAW SOC. J. 44 (Aug. 1986); John Wadham, *Unravelling Miscarriages of Justice*, 1993 NEW L.J. 1650 (Nov. 1993); Adrian Zuckerman, *Miscarriage of Justice—A Root Treatment*, CRIM. L. REV. 323 (1992); Adrian Zuckerman, *Miscarriage of Justice and Judicial Responsibility*, CRIM. L. REV. 492 (1991).

98. At some points in their chapter on scientific evidence, Nobles and Schiff postulate the temporal contingency of scientific evidence and assert that the presentation of scientific evidence in legal settings “accords with its popular perception.” NOBLES & SCHIFF, *supra* note 82, at 174, 183. On other occasions they provide an essentialized image of the scientific method and a range of norms and commitments, citing works from the sociology of scientific knowledge that are generally highly critical of such approaches. *Compare id.* at 179, 185-86. *with* EXPERT EVIDENCE, *supra* note 6. Nobles and Schiff also suggest the existence of a culture clash between the systems of science and law. *See also* STEVEN GOLDBERG, CULTURE CLASH: LAW AND SCIENCE IN AMERICA (1994); Gary Edmond & David Mercer, *Manifest Destiny: Law and Science in America*, 10 METASCIENCE 40 (1996).

99. NOBLES & SCHIFF, *supra* note 82, at 189-90 (emphasis added).

They later note as follows:

Dr Hugh Black gave evidence of the *cumulative* nature of the tests used for NG. *As a scientist, he would not have accepted single positives unconfirmed by other tests.* Even with the GC/MS test, which is the most sensitive, *he would not have accepted a single test, which confirmed* the presence of one mass. He wanted all three masses associated with NG to be confirmed. In his evidence to the 1987/8 appeal Dr Black was fairly *uncompromising* in asserting that the tests which had been carried out were insufficient, and that *he did not himself have to carry out such tests* in order to reach this conclusion. . . . On Dr Black's view neither the evidence of Dr Skuse, who had none of his Griess test positives confirmed by any other test, nor of Dr Drayton, who had done only one run and seen a mass of 46 at a time associated with NG, would suffice. The Court of Appeal never accepted Dr Black's view even though other scientists at the 1987/8 appeal supported it. And even at the 1991 appeal the Court accepted that the single GC/MS result must mean something, and only rejected it when a plausible innocent explanation was provided. Whilst the 1991 Court formally accepted Dr Black's view on the need for Griess tests to be confirmed, this was largely lip service; they concentrated on the presence of an alternative explanation for the Griess test results, rather than rejecting them as inadequate data.¹⁰⁰

These passages repeatedly draw upon and reinforce an idealized methodological and normative ethos for their effect. Unlike Skuse, Black is portrayed as a good scientist.¹⁰¹ Subsequent analysis of Black's evidence is developed in a subsection entitled *Good Science*. He is characterized as uncompromising, committed to the cumulative nature of *Science* and the reproduction of results. As a *Scientist* Black could not accept a single positive result. Carefully managing these appearances and attributes in an attempt to distinguish the circumstances leading to the 1987 decision, the authors provide little indication of the strategic nature of Black's evidence, its integration into a defense case, its changing reception based on the prevailing socio-legal climate or the introduction of other evidence such as ESDA, and the Court's institutional imperative in 1991. Rather, Black's testimony and the methods underpinning it are portrayed as methodologically rigorous and obvious. This approach to Black's

100. *Id.* at 199-200 (emphasis added).

101. The category of "good science" shares many similarities with the empiricist repertoire described by GILBERT & MULKAY, *supra* note 39, and COLLINS & PINCH, *supra* note 46.

evidence raises a fundamental question: why was Black's evidence not originally accepted if it was so demonstrably reliable, predicated as it was upon the proper scientific attitude and techniques? For Nobles and Schiff the answer is found in the irrationality of legal procedures, particularly the continuing use of lay judges and juries.¹⁰² This leads them to lament that despite the support of other scientists, "The Court of Appeal never accepted Dr Black's view" but merely provided "lip service."¹⁰³ Here the converse of Upchurch's assessment of Paul's evidence is occurring. Just as the Attorney-General should have realized the weakness in Paul's evidence, a rational system of adjudication would have recognized the intrinsic value of Black's opinion.

Perhaps curiously, in other parts of their discussion, Nobles and Schiff refer to extracts from the 1987 appeal in which Black acknowledged, during cross-examination at the trial, limits to his evidence.¹⁰⁴ These observations play no discernible role in the assessment of the value of his testimony. Instead they are deployed to illustrate that a practical scientist may be more familiar with the operation of some techniques than a more qualified chemist like Skuse.¹⁰⁵ Similarly, the longer extract above provides little sense of expert disagreements over preferences in relation to testing and interpretations of tests—some of which it alludes to—nor the practical difficulties confronting a forensic scientist, dealing with imperfect and deteriorating samples, or the corroboration between Skuse and Drayton. Images of scientific practice are flexibly linked to claims for (un)reliability.

This analysis is reinforced if we consider the reception of Black's evidence in the Courts of Appeal. Nobles and Schiff are highly critical that neither Court of Appeal accepted the opinions of Dr. Black nor dedicated much attention to them. Rather, as we shall see, his approach was actually criticized by the trial judge, and that criticism found endorsement even in the 1991 judgment in which the convictions were finally quashed. The Courts, in contrast to Nobles and Schiff, encountered an institutional impediment to accepting Black's opinion or infusing it with transhistorical value. In 1976 an application for appeal by the Birmingham Six had been refused. Any subsequent endorsement of Black's original testimony would have

102. NOBLES & SCHIFF, *supra* note 82, at 186-87.

103. *Id.* at 188. Here, Nobles and Schiff acknowledge some of the diachronic changes in the evidence and the ability of appellate courts to resist evidence, even where it may "have left the original prosecution case in tatters" because they determine what constitutes fresh evidence and its legal effect on the safety of the original conviction. *Id.*

104. *Id.* at 204-05.

105. *Id.* at 203.

compromised that decision and any other decision, like the 1987 appeal, that upheld the convictions. To protect the integrity of the trial and processes of judicial review, the Court had an incentive to emphasize, as we have seen, the novelty of the exonerating evidence. These observations are borne out in the 1987 and 1991 judgements.

Black and his evidence are not discussed in the extracted 1987 judgment; the later judgment discussed Black's evidence but inoculated it.¹⁰⁶ Rather than incorporate Black's evidence from the trial into an ironic reassessment of earlier performances, the Court of Appeal effectively endorsed the trial judge's original summation: "The judge [Justice Bridge] then drew attention to Dr. Skuse's experience, Dr. Black's relative inexperience, the absence of any text book authority in favour of Dr. Black's theory, and his failure to carry out any experiments."¹⁰⁷

This presents a very different portrait from the image of Black presented by Nobles and Schiff. Regardless of the (retrospective) merit of Black's testimony, there may, as Justice Bridge implied, have been good reasons to question it at the trial. Justice Bridge's criticism of Black's methodology and evidence raises precisely the type of predicament that symmetry encourages the analyst to address and illustrates the flexibility available in the representation of what comes to be understood as proper scientific practice.

To a considerable degree, the approach adopted by Nobles and Schiff stems from their commitment to autopoietic theory and their attendant belief in system-specific communication, including specialized scientific communications. This approach privileges *Science* as a peculiarly rational pursuit. However, if we cannot always identify—especially *a priori*—what constitutes *good Science*, or if there are serious and intractable disagreements over whether expert disagreement is genuine—what is often at stake in litigation—then it seems inappropriate to criticize jury and judicial decisions as irrational, particularly when we have no independent means of ascertaining the proper value of the evidence or how it should be evaluated

106. ROLAND BARTHES, MYTHOLOGIES 84 (1972). The tendency to "inoculate"—admitting a little evil to prevent a greater one—provides a means of influencing interpretations. In this instance, Black's evidence receives some judicial attention and this makes criticism more difficult than if he were simply ignored.

107. *McIlkenny & Others*, 93 Crim. App. R. 287, 296-97 (C.A. 1991). Nobles and Schiff also accept the proposition advanced by an appellant's barrister, Michael Mansfield, in his address to the Court of Appeal in the 1991 appeal:

The problem that it appears has faced Dr. Black in the first place and Drs. Caddy and Baldock in the second, is that they are scientists with, if I put it this way, immaculate theory. They are right, but because they have not been able to produce an example, the court, in a sense has turned its back on that.

NOBLES & SCHIFF, *supra* note 82, at 201. It is easy to imagine this objection being cursorily dismissed in litigation with alternative social circumstances or other plausible evidence.

rationality. These observations are strengthened when we consider that other groups, such as scientists and journalists, also accepted and even endorsed the original conviction for over a decade.

Without intending to address the question of factual guilt or innocence, and suspending judgment on the question of the proper value of the evidence and its refraction through legal frameworks, the foregoing analysis provides some indication of the manner in which not only criminal convictions are constructed, but also how miscarriages of justice are the product of a range of interrelated social, legal, and epistemic contingency. Instead of an allegedly epistemologically warranted reversal, the politics of institutions and inseparably expert knowledges enable the processes of conviction and exoneration to be opened to historical and sociological scrutiny.

IV. PATHOLOGICAL SCIENCE IN CONTEMPORARY MASS TORT LITIGATION

The final example is drawn from recent and controversial mass tort litigation in the United States.¹⁰⁸ It shifts our attention from the criminal to the civil sphere—to product liability, the law of negligence, and the difficulties associated with establishing legal causation.¹⁰⁹ Unlike the previous examples concerned with just a few isolated individuals, trials, and appeals, the Bendectin litigation involved thousands of plaintiffs and many trials and appeals.

Bendectin was marketed as an anti-nausea drug and commonly prescribed for women with morning sickness.¹¹⁰ After several decades of use (from 1956 to 1983), a few epidemiological studies produced in the late 1970s raised, somewhat inconclusively, concerns about its safety. From the early 1980s, some of the many families of women who had consumed Bendectin during the first trimester of pregnancy endeavored to sue the pharmaceutical manufacturer, Merrell Dow, alleging that it had caused a range of birth defects, particularly limb reduction abnormalities. To sustain their actions, the plaintiffs were required to prove that on the balance of probabilities Bendectin had caused their injuries. In the wake of a few initial “victories,” the

108. This overview of the Bendectin litigation is adapted from a longer and more comprehensive discussion by Gary Edmond & David Mercer, *Litigation Life: Law-Science Knowledge Construction in (Bendectin) Mass Tort Litigation*, 30 SOC. STUD. SCI. 265 (2000).

109. The seminal work in the field, routinely cited by judges, is HERBERT HART & TONY HONORÉ, *CAUSATION IN THE LAW* (1985). For alternative analyses, see Margaret Berger, *Eliminating General Causation: Notes Towards a New Theory of Justice and Toxic Torts*, 97 COLUM. L. REV. 2117 (1997); Gary Edmond & David Mercer, *Rebels Without a Cause? Judges, Scientific Evidence and the Uses of Causation*, in *CAUSATION IN LAW AND MEDICINE* 83 (Ian Freckelton & Danuta Mendelson eds., 2002).

110. Bendectin was marketed as Debendox in the United Kingdom and Australia.

number of plaintiffs increased dramatically, resulting in the filing of numerous suits and the consolidation of cases into joint trials. As it prepared to defend an escalating number of cases, Merrell Dow removed Bendectin from the market in 1983.

Considered as a whole, the Bendectin litigation is interesting because it provides an example of sustained legal and scientific attention focused on evidence of causation across numerous trials and appeals from the early 1980s to the mid-1990s. The cases provide an extensive public record of trial and appellate judgments and have generated considerable legal commentary. Judges and scholars routinely cite the Bendectin litigation and its historiography to support the need for alternative means of dealing with the refractory problems raised by the use of expert evidence in legal settings.

Almost all scholarly accounts describe the Bendectin litigation as an extreme example of legal pathology.¹¹¹ In the most detailed study to date, Green characterized it as a demonstration “that the tort system can go awry.” According to Green, “Bendectin is the Taj Mahal of horror stories about the tort system: the single most criticized piece of large-scale litigation of all time.”¹¹² As a result, commentators tend to be highly critical of the civil justice system for its inability to resolve the Bendectin controversy expeditiously in favor of the defendant, Merrell Dow. Generally, the concerns of commentators are focused on the inability of juries, and to a lesser degree judges, to *properly* value the great weight of scientific evidence that demonstrated no legally or scientifically meaningful correlation between the incidence of birth defects and the consumption of Bendectin.¹¹³ What is presented as the overwhelming weight of scientific evidence in favor of Bendectin’s safety operates ironically to condemn the performance of the U.S. civil justice system.

Despite some differences in the extent of detail and emphasis, accounts of the Bendectin litigation by Huber, Green, and Sanders all attribute responsibility for the protracted litigation to the inability of the courts to properly value the scientific evidence. For these commentators the epidemiological record clearly vindicated Merrell Dow, but the courts and particularly juries failed or took too long to

111. The term is a derivation of the title of Langmuir’s influential essay, Irving Langmuir, *Pathological Science*, reprinted in 42 *PHYSICS TODAY* 36 (Oct. 1989).

112. MICHAEL GREEN, *BENDECTIN AND BIRTH DEFECTS* 328 (1996). See also NORMAL LEVITT, *PROMETHEUS BEDEVILLED: SCIENCE AND THE CONTRADICTIONS OF CONTEMPORARY CULTURE* 211 (1999).

113. Most of the commentaries insinuate that the legal and scientific outcomes should be identical, that the legal system can and should identify and reflect what is presented as a non-problematic, proper scientific valuation.

appreciate that *reality*.¹¹⁴ Each of the authors expresses a belief, exemplified in the following statement by Green, in the existence of a hierarchy among the various types of evidence presented in Bendectin trials, and all locate epidemiological evidence at the top: "There plainly is a hierarchy to these different forms of toxic effect evidence. Epidemiology is at the top, and structural similarity, in vitro testing, and case reports are at the bottom."¹¹⁵ These accounts rely upon and privilege a purportedly epistemologically based conjunction between the defendant's case strategy and the eventual judicially-led closure around the primacy of published epidemiological evidence. Consequentially, it caricatures the efforts of the plaintiffs, their lawyers, and scientists as irredeemable. In this way it operates as a sociology of error predicated upon an historically insensitive approach to the contests and development of a judicial consensus around not only the meaning of the evidence, but also what should be recognized as evidence in legal fora. The defendant's case, and implicitly the safety of Bendectin, is vindicated on the basis of a purportedly clear and convincing epidemiological (read scientific) record.¹¹⁶

By now a familiar pattern involving the retrospective organization of evidence into *good* (reliable) *Science* and *bad* (unreliable) *Science* should be emerging. Once again, the adoption of more flexible and contingent models of scientific knowledge and practice will illustrate how even this leading example of pathological litigation with the evidence overwhelmingly in favor of the defendant exhibits some indication of the difficulties endemic to endeavors to broker closure in response to expert disagreement. As in the previous examples, opening-up and examining the processes involved in the judicial

114. See KENNETH FOSTER & PETER HUBER, *JUDGING SCIENCE: SCIENTIFIC KNOWLEDGE AND THE FEDERAL COURTS* (1998); PETER HUBER, *GALILEO'S REVENGE: JUNK SCIENCE IN THE COURTROOM* 111-29 (1991); JOSEPH SANDERS, *BENDECTIN ON TRIAL: A STUDY OF MASS TORT LITIGATION* (1998); Michael Green, *Expert Witnesses and Sufficiency of Evidence in Toxic Substances Litigation: The Legacy of Agent Orange and Bendectin Litigation*, 86 NW U. L. REV. 643 (1992); Joseph Sanders, *The Bendectin Litigation: A Case Study in Life Cycles of Mass Torts*, 43 HASTINGS L.J. 301 (1992) [hereinafter Sanders, *Bendectin Litigation*]; Joseph Sanders, *From Science to Evidence: The Testimony on Causation in the Bendectin Cases*, 46 STAN. L. REV. 1 (1993). See also RICHARD GOLDBERG, *CAUSATION AND RISK IN THE LAW OF TORTS: SCIENTIFIC EVIDENCE AND MEDICINAL PRODUCT LIABILITY* 102-31 (1999); Bert Black, *A Unified Theory of Scientific Evidence*, 56 FORDHAM L. REV. 595 (1988); Louis Lasagna & Sheila Shulman, *Bendectin and the Language of Causation*, in *PHANTOM RISK: SCIENTIFIC INFERENCE AND THE LAW* 101 (K. Foster et al. eds., 1993).

115. GREEN, *supra* note 112, at 37.

116. This is similar to equating the quashing of a conviction in the miscarriage of justice trials (legal innocence) with factual innocence. In the terminology of White, epidemiology acts as synecdoche purporting, in the defence account and many judicial reviews, to stand for all scientific knowledge. See HAYDEN WHITE, *THE CONTENT AND THE FORM: NARRATIVE DISCOURSE AND HISTORICAL REPRESENTATION* 31-38 (1987).

closure of the Bendectin litigation introduces a more complex realm of socio-epistemic contests than the leading accounts suggest. Instead of according a privileged position to the published epidemiological record, this approach indicates how the legal closure becomes intertwined with a historical sociology of competing knowledges and an archaeology of the legal uses of epidemiological evidence. That is, investing epidemiology with legal credibility provided an evidentiary resource to explain and justify resolution.¹¹⁷

In undertaking an examination of the judicial treatment of the scientific evidence in the Bendectin litigation, the most striking feature is the primary evidential role eventually accorded to published epidemiological studies that failed to identify statistically significant associations between Bendectin consumption and birth defects.¹¹⁸ From the very beginning of the Bendectin litigation, Merrell Dow and their highly regarded and expensive legal advisers concentrated the legal defense upon the published epidemiological record. The perceived strength of this record, in conjunction with earlier judgments endorsing the value of epidemiological evidence, facilitated this approach.¹¹⁹ By the mid- to late 1980s, a number of federal appellate courts had come to accept published epidemiological evidence as the appropriate body of knowledge to resolve the Bendectin litigation. Many of these decisions were outcome-dispositive, effectively dismissing the plaintiffs' evidence and any chance of success in court. If the plaintiffs were unable to adduce published epidemiological studies linking Bendectin with birth defects, then there would be no further litigation. In practice, the preference for epidemiological evidence acts as a type of epistemological tautology, a consideration disregarded by Huber, Green, and Sanders. Published epidemiological studies are characterized as the appropriate type of evidence to resolve the controversy over causation, and an increase in the number of published studies unable to establish a statistically significant association purports to strengthen the appearance of a scientific (and simultaneously judicial) consensus. This approach

117. As in the example of images of the ideal medical examination in the case of Boulton and Park, this is not independent of extra-legal approaches to epidemiology or the historical emergence and recognition of epidemiological evidence in federal courts.

118. None of the forty studies undertaken by 1990 led the authors to confidently conclude that there were statistically significant associations between Bendectin and birth defects. For a discussion of the qualified reporting of epidemiological studies, consider David Rier, *The Versatile Caveat Section of an Epidemiology Paper; Managing Public and Private Risk*, 21 SCI. COMM. 3 (1999).

119. One example is the contest over causation in relation to personal injury in the Agent Orange litigation, *Agent Orange Prods. Liab. Litig.*, 597 F. Supp. 740 (E.D. N.Y. 1984). See also PETER SCHUCK, *AGENT ORANGE ON TRIAL: MASS TOXIC DISASTERS IN THE COURTS* (1986).

might have been more convincing if published epidemiological studies were the only type of evidence adduced by the parties. But they were not.

Although some of the more ambiguous epidemiological studies had actually contributed to anxieties about the safety of Bendectin and were referred to in many trials, plaintiffs, their lawyers, and experts were aware that the expanding epidemiological record did not, at least in terms of statistically significant published results, support their civil actions. In response, they developed their cases on the basis of other types of evidence. They drew upon the evidence, routinely used in regulatory decision-making, of toxicology, *in vivo* and *in vitro* studies, chemical structure analyses between Bendectin and known teratogens, and the allegedly checkered institutional history of Merrell Dow, particularly the fact that it had produced and sponsored questionable research and had earlier associations with MER-29 and Thalidomide.¹²⁰ If we adopt a symmetrical approach and endeavor to explain the production of the judicial-defendant consensus around the primacy accorded to the published epidemiological record, then the relative weight and meanings of the different types of evidence, especially as the number of studies and findings changed over time, becomes far more complex than is suggested in the secondary accounts and many of the judgments. The complexity, however, extends beyond the legitimacy of interdisciplinary hierarchies.

Not only did judges, lawyers, and scientists attribute different values to the broader evidentiary record beyond the published epidemiological studies, but there were considerable differences in the values ascribed to the meaning of the epidemiological evidence. While the defense and many judges emphasized that virtually every published study had failed to locate statistically significant associations, closer inspection indicates variation between the studies in relation to the strength of associations found in different types of alleged birth defects.¹²¹ Some types of study were portrayed as more reliable or relevant than others. In addition, several of the plaintiffs' experts, including epidemiologists, undertook re-analysis and meta-analysis of published results and claimed their re-calculations

120. *In re Richardson-Merrell, Inc., "Bendectin" Prods. Liab. Litig.*, 624 F. Supp. 1212 (S.D. Ohio 1985).

121. This provides an example of how legal concerns influence the decision to litigate in ways that transcend narrow epistemological issues. Contingency-fee lawyers acting for the plaintiffs generally sought to try particular cases, such as those with limb reduction deformities, because they offered the largest potential damage awards. But these deformities did not necessarily display the strongest of the associations between Bendectin and birth defects in the epidemiological literature.

demonstrated stronger statistical associations.¹²² There was also criticism of the judicial insistence on the conventional standard of statistical significance set at 0.05 or a relative risk of 2.0.¹²³ Ultimately, most of these issues were resolved in a manner that favored the defendants, but the closure was not as easy or epistemologically convincing as accounts of the litigation routinely suggest.

Some indication of the difficulty involved in endeavoring to value the epidemiological evidence properly is provided through the different approaches adopted not only by the parties, but also by judges across the decade when Bendectin litigation raised logistical, legal, and social concerns, and by scientists, even those not directly engaged in the litigation. Despite the claim that most, if not all, of the Bendectin litigation was without merit, most commentators seem to encounter difficulty identifying a precise moment when the published epidemiological record was sufficiently compelling to warrant the exclusion of all Bendectin cases from courts. Without arbitrarily adopting some point late in the litigation—which would be to accept the primacy of the epidemiological evidence—there was disagreement among the protagonists over the meaning of the published epidemiological record. Contrary to claims by Huber, Green, and Sanders, it was not only juries who were inconsistent.¹²⁴ Scientists from different fields and even some from similar fields interpreted the scope and meaning of the evidentiary record inconsistently. For example, although Green and Sanders would presumably locate their moment of epidemiological evidentiary sufficiency somewhere around 1985 or 1986, in 1983 editorials in scientific journals were highly critical of the litigation, describing it as “lacking scientific merit.”¹²⁵ This reaction raises the following question: on what basis did these scientists decide, given that Green and Sanders accept some initial ambiguity in the evidence, before the bulk of the epidemiological studies were published? Another difficulty is the submission of an *amicus curiae* brief to the U.S. Supreme Court in relation to the *Daubert* appeal over admissibility standards from a

122. In the 1993 *Daubert* appeal to the United States Supreme Court, the majority opinion was complimentary with respect to the formal reputations and qualifications held by some of the plaintiffs’ expert witnesses. They were not, even at that date, dismissed as charlatans.

123. These are often succinctly explained in judgments. Judges frequently draw upon earlier legal decisions, as often as science textbooks, for their (authoritative) explanations of scientific evidence. This is common for epidemiological evidence and descriptions of DNA typing techniques.

124. The characterization of jury verdicts as inconsistent, when juries were exposed to different cases, different plaintiffs, different trial tactics, different lawyers, different judges, different scientists, and a changing body of evidence, and fulfilled a different constitutional role to appellate courts and trial judges is, at the very least, troublesome.

125. SANDERS, *supra* note 114, at 88 (citing Editorials published in CANADIAN MEDICAL ASSOCIATION JOURNAL and DRUG INTELLIGENCE & CLINICAL PHARMACY).

Bendectin trial as late as 1993.¹²⁶ In their brief, a group of eminent epidemiologists criticized the approach to publication and statistical significance adopted in several federal circuits. Scientists, including epidemiologists, did not speak with a unified voice on the appropriate way to approach the evidentiary record or even how to approach epidemiological evidence, and there was no overall consistency as to the point at which a particular body of knowledge should be seen to have particular legal effects.

A survey of judgments over the course of the entire litigation does not clarify the situation. At the beginning of the Bendectin litigation, before it had become a mass tort,¹²⁷ trial judges hearing Bendectin trials tended to admit all of the non-epidemiological evidence, and some federal appellate courts not only accepted the admissibility of the plaintiffs' evidence, but also endorsed verdicts based upon it. However, from the mid-1980s appellate courts began to exclude the plaintiffs' evidence, explaining that it was not sufficiently scientific or reliable to be admitted or that the evidence was insufficient to prove that Bendectin had caused the specific injuries.¹²⁸ Despite this emerging pattern of rejection, in the early 1990s, two appellate judgments, *DeLuca v. Merrell Dow Pharmaceuticals, Inc.*¹²⁹ and *Turpin v. Merrell Dow Pharmaceuticals, Inc.*,¹³⁰ endeavored to provide more extensive reviews of the evidence and the legal standards, and in 1990 a judge who had presided over a Bendectin trial in 1985 that had consolidated 800 cases admitted the plaintiffs' evidence.¹³¹ Further, in the 1993 *Daubert* judgment, the Supreme Court made no findings regarding the adequacy of the respective cases, even though Huber had already written his influential text and Green and Sanders had published preliminary articles that were highly critical of the Bendectin litigation.¹³² To some extent the judicial tendencies

126. Brief of *Amici Curiae* Professors Kenneth Rothman, Noel Weiss, James Robbins, Raymond Neutra, and Steven Stellman in Support of Petitioners, *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579 (1993) (No. 92-102). Earlier, in 1985, ten epidemiologists had appended their names to a letter submitted to the journal *TERATOLOGY* expressing their concerns about Bendectin.

127. Sanders, *Bendectin Litigation*, *supra* note 114.

128. *E.g.*, *Lynch v. Merrell-Nat'l Labs.*, 830 F.2d 1190 (1st Cir. 1987); *Richardson v. Richardson-Merrell, Inc.*, 857 F.2d 823 (D.C. Cir. 1988); *Ealy v. Richardson-Merrell, Inc.*, 897 F.2d 1159 (5th Cir. 1989); *Brock v. Merrell Dow Pharm., Inc.*, 884 F.2d 167 (5th Cir. 1989).

129. *DeLuca v. Merrell Dow Pharm., Inc.*, 911 F.2d 941 (3rd Cir. 1990).

130. *Turpin v. Merrell Dow Pharm., Inc.*, 959 F.2d 1349 (6th Cir. 1992).

131. *In re Bendectin Prods. Liab. Litig.*, 732 F. Supp. 744 (E.D. Mich. 1990); *In re Richardson-Merrell, Inc.*, 624 F. Supp. 1212 (S.D. Ohio 1985).

132. In addition to developing evidential and procedural rules, appellate courts are often concerned with the merits of certain actions as well as broader considerations about community acceptance, justice, and the ramifications of decisions. *But compare* *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 119 S.Ct. 1167 (Ala. 1999) (examining the experts' evidence in crit-

to admit evidence at the beginning of the litigation and then adopt a far more restrictive approach once some of the implications of those earlier cases had emerged might be understood to be consistent with managerial strategies and social and economic anxieties related to mass tort litigation as much as putatively epistemological orientations.¹³³

Neither courts nor commentators provide much explanation beyond recourse to authority (often legal) for preferring published epidemiological evidence to the mosaic of other factors. Their preference is further compromised by the ambiguous status of publications once the idealized image of rigorous skeptical peer review is tempered. This is not to suggest that courts and commentators do not distinguish between the different types of evidence, but rather, to open these meta-scientific distinctions to examination.¹³⁴

An example of some convergence between the interests of the defendant and the concerns of judges confronted with an expanding congregation of Bendectin cases might be developed in relation to their respective approaches to publication as an evidentiary or evaluative resource.¹³⁵ The defendants embraced published epidemiological studies because the results supported their position. That said we should not fall into the trap, on the basis of hindsight, of thinking that this was, at least initially, legally mandated.¹³⁶ Further, appealing

ical detail) *with* Kumho Tire Co., 526 U.S. at 257, 119 S.Ct. at 1180 (Stevens, J., dissenting) (criticizing the majority's decision to evaluate the evidence instead of remanding).

133. A few courts modified some of the more onerous requirements, such as the legal necessity of statistically significant published epidemiological studies, once the "threat" posed by Bendectin cases had dissipated. *See, eg., Deluca*, 911 F.2d 941; *Turpin*, 959 F.2d 1349. It may have been relatively easy to address the evidence more directly once the litigation had been effectively shutdown.

134. The *a priori* commitment to the value of the epidemiological evidence, particularly normative images of scientific practice such as skepticism and neutrality, seems to invest the evidence with its value. It might be that while the preferred outcome is not necessarily determinative in the evaluation of evidence, it remains, nonetheless, highly influential. For some discussion of the metaphysical frameworks associated with the emergence of the modern sciences, consider some of the early works by Kuhn and Burtt, EDWARD BURTT, *THE METAPHYSICAL FOUNDATIONS OF MODERN SCIENCE* (1924); THOMAS KUHN, *THE ESSENTIAL TENSION* (1977); THOMAS KUHN, *THE STRUCTURE OF SCIENTIFIC REVOLUTIONS* (1962).

135. Perhaps those like Callon and Latour, who embrace actor-network theory, might describe this as an instance of the defendant's enrolment of the judiciary and epidemiologists and epidemiological studies. *See* Michel Callon, *Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St. Brieuc Bay*, in *POWER, ACTION AND BELIEF: A NEW SOCIOLOGY OF KNOWLEDGE?* 196-233 (John Law ed., 1986); Callon & Latour, *supra* note 13; John Law & Michel Callon, *The Life and Death of an Aircraft: A Network Analysis of Technical Change*, in *SHAPING TECHNOLOGY/BUILDING SOCIETY: STUDIES IN SOCIOTECHNICAL CHANGE* 21 (W. Bijker & John Law eds., 1992).

136. *See* JASANOFF, *supra* note 35; John Abraham, *Distributing the Benefit of the Doubt: Scientists, Regulators, and Drug Safety*, 19 *SCI. TECH. & HUM. VALUES* 493 (1994); John Abraham, *Scientific Standards and Institutional Interests: Carcinogenic Risk Assessment of Benoxapfen in the U.K. and U.S.*, 23 *SOC. STUD. SCI.* 387 (1993); John Abraham & Julie

to pervasive though largely uncritical images of scientific practice, published studies had the benefit of having been formally peer reviewed. If we ignore the fact that litigation inspired the vast majority of studies, the claim that they were developed away from litigation enabled them to be presented free of apparent partisanship or the science-for-litigation stigma.¹³⁷ It was no coincidence that the term “junk science,” a highly pejorative characterization of purportedly unreliable litigation-driven knowledge, disproportionately applied to plaintiff’s evidence, was popularized in the wake of the Bendectin litigation.¹³⁸ Confronted with thousands of potential litigants and the need to publicly rationalize their decision-making, the appeal of published epidemiological studies, which identified no statistically significant association between Bendectin and birth defects and could be portrayed as impartial and inherently reliable, should not be underestimated.¹³⁹ Merrell Dow sought to restrict the debate about causation to the evidence that supported its case most directly. This strategic maneuver was simultaneously developed through the characterization of earlier judgments, including those from other mass torts, that had recognized epidemiological evidence as authoritative. For Merrell Dow, this strategy had the additional benefit that if judges began to accept epidemiological evidence as the appropriate means for assessing causation evidence in mass tort litigation, then plaintiffs, and more particularly their lawyers and scientists, would appear disingenuous. In effect they would appear to be, quite irrationally, arguing against what was being portrayed as consensual scientific reality.

The approaches adopted in the leading commentaries and some of the judicial decisions provide a moral framework and chronology through which to assess the performances of not only individual actors (lawyers, scientists, judges, and jurors), but also the workings

Sheppard, *Democracy, Technocracy, and the Secret State of Medicines Control: Expert and Nonexpert Perspectives*, 22 SCI. TECH. & HUM. VALUES 139 (1997); Rier, *supra* note 118.

Occasionally editors have to consider the social implications of their publications and how they might influence litigation and regulatory decisions. Consider, for example, the debates around the reliability of DNA typing evidence in the pages of *Nature* in the late 1980s and early 1990s. See also *Contested Identities*, *supra* note 6.

137. Edmond, *supra* note 31.

138. See Gary Edmond & David Mercer, *Trashing “Junk” Science*, STAN. TECH. L. REV. (1998), available at <http://stanford.edu/STLR/Articles/Index.htm>.

139. In some circumstances, as in this example, judges claim they are adopting scientific opinions directly from the scientific community. This usually involves an emphasis on reliability and a diminution in the role or agency granted to the application of legal standards, other than one concerned with the sufficiency of the evidence. On other occasions, however, judges temper the uses or the expediency of appropriations. See *supra* note 60 (discussing the role of medical opinion evidence in relation to medical negligence).

of the entire U.S. civil justice system.¹⁴⁰ They function to compromise the actions of plaintiff lawyers, plaintiff scientists, and especially the juries that produced pro-plaintiff verdicts. As in the examples of Boulton and Park and the Birmingham Six, assessing performances on the basis of a consensus that developed toward the end of proceedings—without attending to the production of consensus over time—is a highly questionable method of explaining the assessment and representation of evidence in mass tort litigation. It evaluates judicial decision-making and the various legal and social influences based on the assumption that the later judicial decisions were correct, without explaining why epidemiological evidence was privileged or why the parties adopted particular tactics, introduced a range of experts, and emphasized different types of evidence.¹⁴¹

An alternative approach, sketched tentatively above, suggests that rather than an instance of legal pathology, the Bendectin litigation represents an example of the contingency involved in developing expert knowledges for legal purposes, particularly the persuasive production of closure. One might explain why lawyers, scientists, judges, and juries could disagree over what constituted evidence by recognizing a raft of considerations such as evidentiary shifts over time; disagreements among scientists; difficulties in assessing a changing evidentiary record and evaluating different types of scientific evidence; changes in plaintiff and defence strategies; the scale and resource implications of the litigation; the example of earlier trials and appellate decisions; the wealth of the defendants; the reluctance of plaintiffs and lawyers to settle; and the socio-political composition of juries.

As in the example of the Birmingham Six, it seems curious, if the plaintiffs, their lawyers, and their scientists were so patently unbelievable or the published epidemiological evidence so convincing, that U.S. federal courts encountered such protracted difficulty in resolving the litigation. If the Bendectin litigation is the leading example of mass legal pathology, then courts will presumably continue to encounter extreme difficulties when confronted with evidentiary contests conceived as more evenly balanced.

Huber, Green, and Sanders suggest that there was an obvious extra-legal value to the scientific evidence and that the legal system

140. In some ways the criticism and concern resembles instances of alleged scientific impropriety like the recent episode concerning cold fusion. See GIERYN, *supra* note 37; Thomas Gieryn, *The Ballad of Pons and Fleischmann: Experiment and Narrative in the (Un)making of Cold Fusion*, in *THE SOCIAL DIMENSIONS OF SCIENCE* 217 (E. McMullin ed., 1992). For a discussion of some of the moral implications, see Trevor Pinch, *Opening Black Boxes: Science, Technology and Society*, 22 *SOC. STUD. SCI.* 487 (1992).

141. See Sanders, *The Bendectin Litigation*, *supra* note 114 (discussing changing evidentiary strategies in one of his earlier articles on the Bendectin litigation).

ought to have identified and adopted it. None of their analyses, because of their restricted epistemological orientation, enables them to address the question of how the legal system was actively and constitutively involved in the production and assessment of scientific evidence, nor how concerns about logistics and the enforcement of admissibility and sufficiency standards might influence the presentation and assessment of the evidence, and ultimately our understanding of the sciences or the existence of public problems.¹⁴²

In closing, it is important to acknowledge that Bendectin has recently returned to the market. The entrenched attitudes of senior appellate courts and an extensive and unsuccessful record of litigation, perhaps more than the extensive epidemiological record *supporting* its safety, may explain its return. Questions about the safety of Bendectin remain inextricably bound to what is considered as reliable evidence and how it should be interpreted.

V. CONCLUSION: LEGAL HISTORIOGRAPHY AND ITS CONSEQUENCES

The three cases provide recent examples of the legal, historical, and judicial assessment of controversial expert evidence. Notwithstanding the detailed scholarly treatment of these cases, I have endeavored to illustrate how questionable assumptions about expert and scientific evidence have constrained all of the analyses. Assumptions about scientific and other forms of expertise have radically simplified important evidentiary dimensions in each of the studies and inhibited “thicker” descriptions that might have facilitated a history and sociology of expert evidence, examining the production of expert opinion evidence not only in the local legal context, but extending it further afield in an attempt to trace relations to broader institutional and professional commitments, and social and political values.¹⁴³

142. See JOSEPH GUSFIELD, *THE CULTURE OF PUBLIC PROBLEMS: DRINKING-DRIVING AND THE SYMBOLIC ORDER* (1981) (discussing “public problems”). In relation to concerns about litigation explosions and insurance crises, see Theodore Eisenberg & James Henderson, *Inside the Quiet Revolution in Products Liability*, 39 *UCLA L. REV.* 731-810 (1992); Marc Galanter, *The Day After the Litigation Explosion*, 46 *MD. L. REV.* 3 (1986); Marc Galanter, *Reading the Landscape of Disputes: What We Know and Don't Know (and Think We Know) About Our Allegedly Contentious and Litigious Society*, 31 *UCLA L. REV.* 4 (1983); James Henderson & Theodore Eisenberg, *The Quiet Revolution in Products Liability: An Empirical Study of Legal Change*, 37 *UCLA L. REV.* 479 (1990); Michael Saks, *Do We Really Know Anything About the Behavior of the Tort Litigation System—and Why Not?*, 140 *U. PA. L. REV.* 1147 (1992); Michael Saks, *If There Be a Crisis, How Shall We Know It?*, 46 *MD. L. REV.* 63 (1986).

143. Clifford Geertz, *Thick Description: Toward an Interpretative Theory of Culture*, in *THE INTERPRETATION OF CULTURES* 3 (1973); see also Harry Collins, *Stages in the Empirical Programme of Relativism*, 11 *SOC. STUD. SCI.* 3 (1981).

Each of the case studies indicates how the assignation of a stable value to the expert evidence functions effectively to resolve controversy. That assignation does not operate impartially. Rather, it serves to problematize individual and institutional performances and limits recognition of the local, dynamic, and strategic contests over the meaning and application of knowledge routinely encountered in adversarial legal settings. For Upchurch, the Attorney-General's decision to obtain no further medical evidence is characterized as a personal failure and evidence of a Crown conspiracy. The possibility that the interpretation of the medical evidence changed or was open to different readings at different times, and that the non-medical dimensions of the case were considered sufficient for a successful prosecution, are consequently left unattended. In the final appeal of the Birmingham Six, the inclusion of institutional concerns and the broader evidentiary and social matrix may enhance our appreciation of the changing approach to the evidence of nitroglycerine that was previously represented as sufficient to sustain the convictions. Nobles and Schiff downplay some aspects of this remarkable reversal of the evidence—and the explanation—involved in quashing the convictions for murder. Consequently, the inability to recognize and properly value Black's evidence at that time becomes evidence of the irrationality of legal institutions, rather than an indication of the contest over propriety, reliability, and credibility occurring in the context of the broader narratives unfolding during the trial and subsequently through the course of appeals.¹⁴⁴ For Green, Huber, and Sanders, the privileging of statistically significant, published epidemiological studies converted plaintiffs' counsel into cash-crazed lawyers and their experts into hired guns, willing to testify to the truth of any theory. The apparent inability of judges and juries to rule appropriately leads the authors to advocate institutional reform.

In contrast to these approaches, more critical sociological images of the sciences combined with more historically contingent and sensitive approaches to evidence, including concern about what counts as evidence and why it does so, suggest the existence of a more fluid realm of contestation and negotiation as the actual content (or *in situ* meaning) of *objective*, *reliable*, and *good* evidence was strategically inscribed, contested, and endorsed.¹⁴⁵ Accepting that images of and commitment to methods, norms, standards, and protocols can pre-exist litigation, their precise configuration, in any

144. See LANCE BENNET & MARCIA FELDMAN, *RECONSTRUCTING REALITY IN THE COURTROOM* (1981); BERNARD JACKSON, *LAW, FACT AND NARRATIVE COHERENCE* (1991).

145. Edmond, *supra* note 46; Gary Edmond, *Law, Science and Narrative: Helping the "Facts" to Speak for Themselves*, 23 S. ILL. U. L.J. 555 (1999).

presentation or attempted closure, is often the result of interests and dynamics manifested in particular litigation, especially the evidence, experts, and legal avenues available to the parties. For judges rationalizing their decisions, appeals to ideals, authority, and impartiality are often explicitly reintroduced in the assessment of admissibility, sufficiency, weight of evidence, and the reliability of convictions. This is meant to enhance the plausibility of their decision and reinforce the rationality and legitimacy of legal institutions. The evidence does not have to be presented as determinative, for judges may privilege legal values and procedures, such as admissibility thresholds, articulated burdens of proof, or procedural concerns, in order to exclude or diminish its impact, even the impact of apparently cogent scientific and expert evidence.

Conventionally, legal procedures, interested parties, and partisan experts have been understood to *distort* scientific evidence.¹⁴⁶ If, however, the legal system is not conceived as a debased or inferior forum distorting pristine extra-legal *Science*, but as a partially separate world with its own concerns, constraints, and traditions—including the availability of cross-examination, the involvement of lay fact-finders, articulated admissibility standards, burdens of proof, and rules of procedure—then the production, presentation, and assessment of scientific evidence might be understood in relation to the exigencies of context. Indeed, the highly public and highly critical context may actually facilitate the investigation of some of the processes involved in the production and valuation of scientific knowledges in contemporary society.¹⁴⁷ As a working axiom, the historian should remember, without being deterministic or reductionist, that scientific knowledge and its manifestations are always contextually contingent.¹⁴⁸ It may be that we need to develop a special concept of *law-science knowledge*, that is, expert evidence developed in relation to not just the contingencies of a case and the rules of procedure and evidence, but also tacit assumptions influencing the production, presentation, and explanation of evidence and its representation in judgments and commentaries. We might then want to consider why in some contexts rhetorical emphasis is placed on *truth* or *reliability*, the purportedly epistemic side of legal strategy and decision-making, and at other times emphasis is laid upon what

146. See, e.g. GOLDBERG, *supra* note 98; HUBER, *supra* note 114; David Bernstein, *Junk Science in the United States and the Commonwealth*, 21 YALE J. INT'L L. 123 (1996); Note, *Confronting the New Challenges of Scientific Evidence*, 108 HARV. L. REV. 1481 (1995).

147. Given the serious attention devoted to the scientific evidence by the various parties, it seems to be a curious methodological tendency to disregard, at the beginning of any study, the local and strategic perspectives developed by the parties that eventually lose.

148. POTTER, *supra* note 39.

might be characterized as the *legal*, *ethical*, and *procedural* dimensions. We might also wish to examine the production of boundaries and maintenance of dichotomies.

Further, commentators adhering to fairly traditional images of *Science* may have unwittingly inverted the issue of distortion. Rather than lament the distortion of scientific evidence in adversarial jurisdictions, it may be more appropriate to interpret some of the highly influential representations of *Science* in legal settings, especially in the seminal decisions of superior courts such as the U.S. Supreme Court's *Daubert* and *Kumho* judgments,¹⁴⁹ as dominant representations, highly conspicuous and influential in a range of public debates, including legal and regulatory fora. The images of *Science* presented in Supreme Court judgments may reach a broader audience and be as influential in shaping public ideas and discourses as those appearing in the pages of the journals *Science* or *Nature*.¹⁵⁰

From the foregoing discussion, it seems that the explanation of scientific disagreement in litigation is inescapably connected to the image of the sciences underlying the analysis. In effect, once the model of *Science* is selected, the litigation more or less explains itself. Butterfield acknowledged this type of predicament when he implored historians to reflect upon their theoretical frameworks:

Our assumptions do not matter if we are conscious that they are assumptions, but the most fallacious thing in the world is to organise our historical knowledge upon an assumption without realising what we are doing, and then to make inferences from that organisation and claim that these are the voice of history.¹⁵¹

149. *Daubert v Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993); *Kumho Tire Co. v Carmichael*, 119 S.Ct. 1167 (Ala. 1999). Recently, in the highly influential *Daubert* judgment, the U.S. Supreme Court indicated that one of the key means of assessing the reliability of scientific evidence was by determining whether the knowledge had been tested (or falsified), referring directly to the philosophers Popper and Hempel. The majority provided a range of factors such as the known or potential error rate of a technique or theory, and whether the claims had been published subject to peer review and generally accepted in the relevant scientific community. Interestingly, this may be interpreted as a derogation from a strict approach to falsification, as these supplementary factors were some of the types of criteria that falsification was designed to eliminate. Very limited scholarly attention has been devoted to the sudden appearance of Popper and falsification in the *Daubert* judgment. See Adwina Schwartz, *A "Dogma of Empiricism" Revisited: Daubert v. Merrell Dow Pharmaceuticals, Inc. and the Need to Resurrect the Philosophical Insight of Frye v. United States*, 10 HARV. J.L. & TECH. 149 (1997); see also Gary Edmond & David Mercer, *Keeping "Junk" History, Philosophy and Sociology of Science Out of the Courtroom: Problems with the Reception of Daubert v. Merrell Dow Pharmaceuticals Inc.*, 20 U.N.S.W. L.J. 48 (1997); Gary Edmond & David Mercer, *Conjectures and Exhumations: Citations of the History, Philosophy and Sociology of Science in US Federal Courts* (2002) (unpublished manuscript, on file with the author).

150. Michael, *supra* note 28, develops the idea of popular idealized images of science, which I have denoted as *Science*, and more variable and contingent localized understandings.

151. BUTTERFIELD, *supra* note 4, at 23-24. From an explicitly phenomenological perspec-

We can assume that historians committed to the categories of *good* and *bad Science* will continue to discover not only charlatans but also scientists whose impeccable opinions were overlooked by irrational procedures and incompetent individuals. Those who approach litigation with more flexible tools may, at least, attempt to investigate how representations of *good* and *bad* are strategically produced in the courtroom and beyond, and explore the complexities involved in the resolution of expert disagreements and their public rationalizations. From a methodological perspective, legal commentators and historians should be reflexive about their descriptions of scientific practice and knowledge. They should be as skeptical of judicial representations of scientific evidence as non-originalist legal scholars examining the representations of the parties or the judges of the Supreme Court in relation to the interpretation of the U.S. Constitution.¹⁵²

While the article by Upchurch has the least currency in relation to contemporary debate about the legal system and law reform, the same claim cannot be made for the accounts of the Birmingham Six and Bendeclin litigation. Accounts of these cases, purporting to provide an historical overview of the trials and appeals, are responding to and actually part of ongoing debates over the meaning of the litigation and the appropriate practical responses to these legal failures or pathologies. The account by Nobles and Schiff contributes to an ongoing discourse about recent high profile miscarriages of justice, judicial and political responses, and institutional reforms such as changes to the Criminal Appeals Act (1968) and the creation of an independent Criminal Cases Review Commission in 1995. Nobles and Schiff are critical of some of the institutional procedures, particularly what they describe as the “irrational” lay treatment of expert evidence. Notwithstanding that their text may be used in ongoing debates in the U.K., their theoretical framework makes it difficult, perhaps inappropriate, for them to suggest reforms. Alternatively, those commenting on the Bendeclin litigation exhibit far less reticence in advocating reform. With the advantages conferred by hindsight, Huber, Green, and Sanders propose a variety of

tive, Merleau-Ponty spoke of the “retrospective illusion” of taking “objects constituted by our perceiving consciousness as pre-existent causes of perception.” MAURICE MERLEAU-PONTY, *SIGNS*, at xiii (Richard McCleary trans., Northwestern U. Press 1964) (1960).

152. See, e.g., Robert Gordon, *Historicism in Legal Scholarship*, 90 *YALE L.J.* 1017 (1981); Robert Gordon, *The Struggle over the Past*, 44 *CLEV. ST. L. REV.* 123 (1996); Mark V. Tushnet, *Following the Rules Laid Down: A Critique of Interpretivism and Neutral Principles*, 96 *HARV. L. REV.* 781 (1983); Tushnet, *supra* note 7. See also *supra* note 72.

reforms focusing particularly on admissibility thresholds as well as on the presentation and assessment of expert evidence.¹⁵³

Interestingly, it might be contended that in practice very few of the proposed reforms would have operated in the ways suggested when the contest over the legal teratogenicity of Bendectin was actually in process. They may have influenced the reception of evidence, but not necessarily in the epistemologically predicated manner the authors imply. The appearance of neutrality and independence, like contests over the appropriate methodology and interpretation of results, are always easier to determine once a controversy has been resolved.¹⁵⁴ What is missing from the proposals is recognition of a tautological loop. The reforms are predicated on questionable assumptions about the value and meaning of scientific evidence and the capability of lay participants. However, if the accounts of the evidence and lay competence are controvertible or fundamentally flawed, then the reforms may not operate in the manner anticipated or may actually be unnecessary.

Historical accounts of miscarriages of justice and mass torts have been central to debates over procedural and evidential reforms in a variety of jurisdictions during the last two decades.¹⁵⁵ In closing, it is my intention to provide an indication of some of the consequences of (mis)understanding the role played by expert evidence in contemporary litigation.¹⁵⁶ Without intending to be nostalgic, I suggest that responses to the Bendectin litigation and a number of other prominent mass torts have been gradually transforming the rights and duties between manufacturers and citizens. Concerns about “junk science” and spurious litigation, of which Bendectin suits are conventionally heralded as a leading example, have led to more restrictive legal admissibility standards. Higher thresholds reduce the number of plaintiffs able to “have their day in court.” Further, the complexity of expert evidence and anxieties about the (in)consistency of outcomes have been combined with ongoing concerns about jury competence. In response, recent U.S. Supreme and Federal Court decisions pertaining to expert evidence have invoked the language of resistance. When making admissibility decisions, trial judges are encouraged to

153. These include the use of court-appointed experts, expert panels, expert juries, and trial by judge alone. See FOSTER & HUBER, *supra* note 114; GREEN, *supra* note 112, at 193-211; SANDERS, *supra* note 114, at 299-351.

154. COLLINS, *supra* note 24.

155. See, for example, the recent change to Rule 702 of the Federal Rules of Evidence (1975) and the review of English civil litigation by Woolf, LORD WOOLF, ACCESS TO JUSTICE: FINAL REPORT (1996).

156. In playing with the idea of (mis)understanding, it is my intention to capture both those advocating law reform on the basis of some belief in the potency of their analysis and those led by perceived benefits, such as lower insurance premiums or compensation payments.

act as vigilant gatekeepers. Gatekeeping is ostensibly concerned with excluding unreliable evidence from courts and protecting jurors. These approaches elide the politics surrounding the selection of particular images of reliability, judicial distrust of American publics, or reduced public access to the legal system, and to a considerable extent achieve these results in reliance on simplistic models of scientific knowledge and controvertible historical accounts.

