
Communicating and Interpreting Statistical Evidence in the Administration of Criminal Justice

4. Case Assessment and Interpretation of Expert Evidence

Guidance for Judges, Lawyers, Forensic Scientists and Expert Witnesses

Graham Jackson, Colin Aitken and Paul Roberts

PRACTITIONER GUIDE NO 4

Case Assessment and Interpretation of Expert Evidence

Guidance for Judges, Lawyers, Forensic Scientists and Expert Witnesses

By

Graham Jackson, Professor of Forensic Science, University of Abertay Dundee

Colin Aitken, Professor of Forensic Statistics, University of Edinburgh

Paul Roberts, Professor of Criminal Jurisprudence, University of Nottingham

Prepared under the auspices of the
Royal Statistical Society's *Working Group on Statistics and the Law*
(Chairman: Colin Aitken)

Contents

0	Introduction	3
	Membership of the Royal Statistical Society's Working Group on Statistics and the Law	12
1	Introduction to Case Assessment and Interpretation	13
1.1	What is Case Assessment and Interpretation (CAI)?	13
1.2	Two Contrasting Styles of Forensic Investigation	13
1.6	Potted Institutional History	16
2	Principles of CAI and the Role of the Forensic Expert	23
2.1	Fundamental Principles of Case Assessment and Interpretation	23
2.2	The Role of the Forensic Expert (in Reducing Uncertainty)	24
2.4	Uncertainty, Probability and Inferential Reasoning	25
2.10	Distinguishing Investigative Advice from Evaluative Expert Opinion	29
2.11	A Taxonomy of Expert Evidence, from Facts to Opinions	30
2.23	How Should Expert Evidence be Formulated?	40
2.24	The Hierarchy of Issues	41
2.25	Hypothetical Illustration: DNA Tights	41
2.34	Combining Evidence through the Hierarchy of Issues	48
2.41	Activity Level Propositions in <i>R v Barry George</i>	54
3	Case Assessment and Interpretation Protocol	60
3.1	Seven Stages of Case Assessment and Interpretation	60
3.14	Case Assessment and Interpretation in Report-writing	67
3.22	Appraising CAI: Practical Benefits, Preconditions and Challenges	72
4	CAI in Forensic Case Work: Three Detailed illustrations	76
4.2	Example #1: 'Sniffing for Death' - How to Evaluate the Response of a 'Cadaver Dog'	76
4.14	Example #2: 'The Robber's Mask' - How to Decide which Scientific Techniques to Employ	83
4.41	Example #3: 'The Case of the Older Offender' - Moving from Investigative to Evaluative Opinion	102

5	How to Read Forensic Reports and Interpret Expert Opinions	124
5.2	Stage 1: Classify the Type of Expert Opinion	124
5.8	Stage 2: Identify the Level in the Hierarchy of Issues	128
5.9	Stage 3: Testing the Class and the Level of an Expert Opinion	128
5.12	Stage 4: Assessing whether the Examination Strategy was Appropriate to Address Facts in Issue before a Court	133
Appendix A	CAI Aide-memoire	134
Appendix B	Bibliography	137

Introduction to *Communicating and Interpreting Statistical Evidence* in the Administration of Criminal Justice

0.1 ***Context, Motivation and Objectives***

Statistical evidence and probabilistic reasoning today play an important and expanding role in criminal investigations, prosecutions and trials, not least in relation to forensic scientific evidence (including DNA) produced by expert witnesses. It is vital that everybody involved in criminal adjudication is able to comprehend and deal with probability and statistics appropriately. There is a long history and ample recent experience of misunderstandings relating to statistical information and probabilities which have contributed towards serious miscarriages of justice.

0.2 Criminal adjudication in the UK's legal jurisdictions is strongly wedded to the principle of lay fact-finding by juries, magistrates and justices of the peace employing their ordinary common sense reasoning. Notwithstanding the unquestionable merits of lay involvement in criminal trials, it cannot be assumed that jurors, lay magistrates or justices of the peace will have been equipped by their general education to cope with the forensic demands of statistics or probabilistic reasoning. This predictable deficit underscores the responsibilities of judges and lawyers, within the broader framework of adversarial litigation, to ensure that statistical evidence and probabilities are presented to fact-finders in as clear and comprehensible a fashion as possible. Yet legal professionals' grasp of statistics and probability may in reality be little better than the average juror's.

Perhaps somewhat more surprisingly, even forensic scientists and expert witnesses, whose evidence is typically the immediate source of statistics and probabilities presented in court, may also lack familiarity with relevant terminology, concepts and methods. Expert witnesses must satisfy the threshold legal test of competency before being allowed to testify or submit an expert report in legal proceedings.¹ However, it does not follow

¹ *R v Atkins* [2010] 1 Cr App R 8, [2009] EWCA Crim 1876; *R v Stockwell* (1993) 97 Cr App R 260, CA; *R v Silverlock* [1894] 2 QB 766, CCR.

from the fact that the witness is a properly qualified expert in say, fingerprinting or ballistics or paediatric medicine, that the witness also has expert – or even rudimentary – knowledge of statistics and probability. Indeed, some of the most notorious recent miscarriages of justice involving statistical evidence have exposed errors by experts.

There is, in short, no group of professionals working today in the criminal courts that can afford to be complacent about their existing levels of knowledge and competence in using statistical methods and probabilistic reasoning.

- 0.3. Well-informed observers have for many decades been arguing the case for making basic training in probability and statistics an integral component of legal education (e.g. Kaye, 1984). But little tangible progress has been made. It is sometimes claimed that lawyers and the public at large fear anything connected with probability, statistics or mathematics in general, but irrational fears are plainly no excuse for ignorance in matters of such great practical importance. More likely, busy practitioners lack the time and opportunities to fill in persistent gaps in their professional training. Others may be unaware of their lack of knowledge, or believe that they understand enough already, but do so only imperfectly (“a little learning is a dang’rous thing”²).
- 0.4. If a broad programme of education for lawyers and other forensic practitioners is needed, what is required and how should it be delivered? It would surely be misguided and a wasted effort to attempt to turn every lawyer, judge and expert witness (let alone every juror) into a professor of statistics. Rather, the objective should be to equip forensic practitioners to become *responsible producers* and *discerning consumers* of statistics and confident exponents of elementary probabilistic reasoning. Every participant in criminal proceedings should be able to grasp at least enough to perform their respective allotted roles effectively and to discharge their professional responsibilities in the interests of justice.

² Alexander Pope, *An Essay on Criticism* (1711).

For the few legal cases demanding advanced statistical expertise, appropriately qualified statisticians can be instructed as expert witnesses in the normal way. For the rest, lawyers need to understand enough to be able to question the use made of statistics or probabilities and to probe the strengths and expose any weaknesses in the evidence presented to the court; judges need to understand enough to direct jurors clearly and effectively on the statistical or probabilistic aspects of the case; and expert witnesses need to understand enough to be able to satisfy themselves that the content and quality of their evidence is commensurate with their professional status and, no less importantly, with an expert witness's duties to the court and to justice.³

0.5 There are doubtless many ways in which these pressing educational needs might be met, possibly through a package of measures and programmes. Of course, design and regulation of professional education are primarily matters to be determined by the relevant professional bodies and regulatory authorities. However, in specialist matters requiring expertise beyond the traditional legal curriculum it would seem sensible for authoritative practitioner guidance to form a central plank of any proposed educational package. This would ideally be developed in conjunction with, if not directly under the auspices of, the relevant professional bodies and education providers.

The US Federal Judicial Center's *Reference Manual on Scientific Evidence* (Third Edition, 2011) provides a valuable and instructive template.⁴ Written with the needs of a legal (primarily, judicial) audience in mind, it covers a range of related topics, including: data collection, data presentation, base rates, comparisons, inference, association and causation, multiple regression, survey research, epidemiology and DNA evidence. There is currently no remotely comparable UK publication specifically addressing statistical

³ *R v B(T)* [2006] 2 Cr App R 3, [2006] EWCA Crim 417, [176]. And see Rule 33.2 of the Criminal Procedure Rules: 'Expert's duty to the court'.

⁴ The third edition of the *Reference Manual* is published jointly by the National Academy of Sciences and the Federal Judicial Center, and can be accessed at www.fjc.gov/public/home.nsf/autoframe?openform&url_l=/public/home.nsf/inavgeneral?openpage&url_r=/public/home.nsf/pages/1448.

evidence and probabilistic reasoning in criminal proceedings in England and Wales, Scotland and Northern Ireland.

0.6 In association with the Royal Statistical Society (RSS) and with the support of the Nuffield Foundation, we aim to fill this apparent gap in UK forensic practitioner guidance by producing a themed set of four *Practitioner Guides* on different aspects of statistical evidence and probabilistic reasoning, to assist judges, lawyers, forensic scientists and other expert witnesses in coping with the demands of modern criminal litigation. The *Guides* are being written by a multidisciplinary team principally comprising a statistician (Aitken), an academic lawyer (Roberts), and a forensic scientist (Jackson) with help for the DNA guide from Roberto Puch-Solis and Sue Pope. They are produced under the auspices of the RSS's Working Group on Statistics and the Law, whose membership includes – or has included since 2008 – representatives from the judiciary, the English Bar, the Scottish Faculty of Advocates, the Crown Prosecution Service, the National Policing Improvement Agency (NPIA),⁵ the Scottish Police Authority and forensic science providers,⁶ as well as academic lawyers, statisticians and forensic scientists.

0.7 ***Using the Four Practitioner Guides – Notes, Caveats and Disclaimers***

The four *Practitioner Guides* have been produced over a five-year period. They are intended to form a coherent package, but each *Guide* is also designed to function as a stand-alone publication addressing a specific topic or set of related issues in detail. Some of the material restates elementary principles and general background that every criminal

⁵ The NPIA seat became vacant, following NPIA's abolition and replacement by the National Crime Agency pursuant to the Crime and Courts Act 2013.

⁶ The Forensic Science Service ceased casework operations and was effectively closed down in March 2012: see House of Commons Science and Technology Committee, *Forensic Science*, Second Report of Session 2013-14, HC 610 (TSO 2013); & *The Forensic Science Service*, Seventh Report of Session 2010–12, HC 855 (TSO, 2011). Members of the Service assisted with earlier *Guides*.

justice practitioner really ought to know. More specialist sections of the *Guides* might be dipped into for reference as and when occasion demands. We hope that this modular format will meet the practical needs of judges, lawyers and forensic scientists for a handy work of reference that can be consulted, possibly repeatedly, whenever particular statistical or probability-related issues arise during the course of criminal litigation.

0.8 *Guide No 1* was published in December 2010 as **Colin Aitken, Paul Roberts and Graham Jackson, *Fundamentals of Probability and Statistical Evidence in Criminal Proceedings* (RSS, 2010)**. This first *Guide* provides a general introduction to the role of probability and statistics in criminal proceedings, a kind of *vade mecum* for the perplexed forensic traveller; or possibly, ‘Everything you ever wanted to know about probability in criminal litigation but were too afraid to ask’. It explains basic terminology and concepts, illustrates various forensic applications of probability, and draws attention to common reasoning errors (‘traps for the unwary’).

Guide No 2 was published in March 2012 as **Roberto Puch-Solis, Paul Roberts, Susan Pope, and Colin Aitken, *Assessing the Probative Value of DNA Evidence* (RSS, 2012)**. Building on the general introduction to statistical evidence and probabilistic reasoning in criminal proceedings provided by the first practitioner guide, *Guide No 2* explores the probabilistic foundations of DNA profiling evidence and considers how to evaluate its probative value in criminal trials. It explains the basic procedures for producing a DNA profile and the methods for calculating its probability in simple and more complex cases. This *Guide* also briefly describes different types of DNA profiling, including ‘low template’ LTDNA, and discusses some issues surrounding the presentation and interpretation of DNA evidence in criminal trials.

Guide No 3 was published in May 2014 as **Paul Roberts and Colin Aitken, *The Logic of Forensic Proof: Inferential Reasoning in Criminal Evidence and Forensic Science* (RSS, 2014)**. Its broad topic is the inferential logic of judicial evidence and proof. Having elucidated the simple, but powerful, basic principles of inferential logic, it goes on to explain how inferential reasoning can usefully be encapsulated and summarised in

graphical models (including Wigmore Charts), some of which are capable of incorporating conditional probabilities (especially Bayesian Networks). Probabilistic methods for analysing inferential chains are becoming more influential in contemporary forensic science. Other models promote more rigorous evidential analysis and improve the construction of forensic arguments without explicit quantification.

All three previously published *Guides* are available free to download from the RSS website: www.rss.org.uk/statsandlaw.

- 0.9 The current *Practitioner Guide* No 4 completes the quartet. It addresses principles of forensic case assessment and interpretation, with particular regard to how forensic scientists should approach the task of assessing what types of forensic examination or other testing should be commissioned, and for what purposes, in particular criminal investigations, so that practitioners can advise their clients and conduct their forensic inquiries accordingly. It also considers how forensic science and other expert evidence is, or should be, presented and evaluated in criminal trials.

Each *Guide* focuses on topics of major practical importance in the administration of criminal justice, meriting sustained investigation in their own right. The individual *Guides* are free-standing publications that can be read as a narrative exposition or dipped into as works of reference. Taken together, the series of four *Guides* is intended to illuminate the general themes, questions, concepts and issues affecting the communication and interpretation of statistical evidence and probabilistic reasoning in the administration of criminal justice.

- 0.10 We should flag up at the outset certain methodological challenges confronting this ambitious undertaking, not least because it is unlikely that we have overcome them all entirely satisfactorily.

First, we have attempted to address multiple professional audiences. Insofar as there is a core of knowledge, skills and resources pertaining to statistical evidence and probabilistic

reasoning which is equally relevant for trial judges, lawyers, forensic scientists and other expert witnesses involved in criminal proceedings, it makes sense to pitch the discussion at this generic level. All participants in the process would benefit from improved understanding of other professional groups' perspectives, assumptions, concerns and objectives. For example, lawyers might adapt and enhance the ways in which they instruct experts and adduce their evidence in court by gaining insight into forensic scientists' thinking about probability and statistics; whilst forensic scientists, for their part, may become more proficient as expert witnesses by gaining a better appreciation of lawyers' assumptions and expectations of expert evidence, in particular regarding the extent and implications of its probabilistic underpinnings.

We recognise, nonetheless, that certain parts of the following discussion may be of greater interest and practical utility to some criminal justice professionals than to others. Our hope is that judges, lawyers and forensic scientists will be able to extrapolate from the common core to their particular interests and professional concerns. We have stopped well short of presuming to specify formal criteria of legal admissibility or attempting to formulate boilerplate instructions for judges to direct juries in criminal trials. It is not for us to make detailed recommendations on the law and practice of criminal procedure.

0.11 The following exposition is also generic in a second, related sense. The *Guides* are intended to be useful, and to be widely used, in all of the United Kingdom's legal jurisdictions. It goes without saying that the laws of probability, unlike the laws of the land, are valid irrespective of geography. It would be artificial and sometimes misleading when describing criminal litigation to avoid any reference whatsoever to legal precepts and doctrines, and we have not hesitated to mention legal rules where the context demands it. However, we have endeavoured to keep such references fairly general and non-technical – for example, by referring in gross to “the hearsay prohibition” whilst skating over jurisdictionally-specific doctrinal variations with no particular bearing on probability or statistics. Likewise, references to points of comparative law – such as Scots law's distinctive court structure or verdict rules – will be few and brief. Readers should not expect to find a primer on criminal procedure in the following pages.

0.12 The preparation of this *Guide* has benefited enormously from the generous (unpaid) input of fellow members of the RSS's Working Group on Statistics and the Law and from the guidance of our distinguished international advisory panel. The *Guide* also incorporates helpful suggestions and advice received from HHJ John Phillips and, in relation to criminal litigation in Scotland, Sheriff John Horsburgh, QC. Whilst we gratefully acknowledge our intellectual debts to this extraordinarily well-qualified group of supporters and friendly critics, the time-honoured academic disclaimer must be invoked with particular emphasis on this occasion: ultimate responsibility for the contents of this *Guide* rests entirely with the three named authors, and none of our Working Group colleagues or other advisers and commentators should be assumed to endorse all, or any particular part, of the text.

The vital contribution of the Nuffield Foundation, without whose enthusiasm and generous financial support this project could never have been brought to fruition, is also gratefully acknowledged.

The Nuffield Foundation is an endowed charitable trust that aims to improve social well-being in the widest sense. It funds research and innovation in education and social policy and also works to build capacity in education, science and social science research. Whilst the Nuffield Foundation is our primary funder, the named authors take sole responsibility for the views expressed in this *Guide*, which are not necessarily endorsed by the Foundation. Further information regarding the Nuffield Foundation's policies and activities is available at www.nuffieldfoundation.org.

0.13 We welcome further constructive feedback on all four published *Guides*. We are keen to hear about practitioners' experiences of using them and to receive suggestions for amendments, improvements or other material that could usefully be incorporated into revised editions.

All correspondence should be addressed to:

Royal Statistical Society
Chairman of the Working Group on Statistics and the Law
12 Errol Street
London
EC1Y 8LX

Alternatively, responses by email may be sent to c.g.g.aitken@ed.ac.uk, with the subject heading “Practitioner Guide No. [1, 2, 3 and/or 4, as appropriate]”.

Our intention is to revise and reissue all four *Guides* as a consolidated publication, taking account of further comments and correspondence. The latest date for submitting feedback for this purpose will be 1 March 2015.

Graham Jackson
Colin Aitken
Paul Roberts

8 December 2014

Membership of the Royal Statistical Society's Working Group on Statistics and the Law⁷

Working Group

Colin Aitken, University of Edinburgh, *Chairman*
Iain Artis, Faculty of Advocates
Graham Cooke, Kings Bench Chambers, Bournemouth
Mark Fenhalls, Criminal Bar Association
Nina Fisher, Law Society of England and Wales
Andrew Garratt, Royal Statistical Society, *Secretary to the Working Group*
Peter Gill, University of Oslo
HHJ Anna Guggenheim QC
Graham Jackson, University of Abertay Dundee
Tim Moloney, Criminal Bar Association
Tom Nelson, Scottish Police Authority
Niamh Nic Daeid, University of Dundee
Roberto Puch-Solis, LGC Forensics
Mike Redmayne, LSE
Paul Roberts, University of Nottingham
Jim Smith, Royal Statistical Society and University of Warwick
Karen Squibb-Williams, Acorn Chambers, Oxfordshire
Gertie Wallace, Crown Office and Procurator Fiscal Service (COPFS)

International Advisory Panel

Joe Cecil, Federal Judicial Center, Washington DC
Stephen Fienberg, Carnegie-Mellon University
James Franklin, University of New South Wales, Sydney
Joseph Gastwirth, George Washington University
Jonathan J. Koehler, Northwestern Law School, Chicago
Richard Lempert, University of Michigan
Nell Sedransk, National Institute of Statistical Science, Research Triangle Park, NC
Franco Taroni, Institute of Police Science, University of Lausanne
Peter Tillers, Cardozo Law School, New York

⁷ This *ad hoc* Working Group was wound up in October 2014, when it was succeeded by a permanent new Section of the Royal Statistical Society devoted to Statistics and Law.

1. Introduction to Case Assessment and Interpretation

1.1 *What is Case Assessment and Interpretation (CAI)?*

Part 1 of this *Guide* introduces a framework for structuring forensic examinations and reporting their results, known as ‘Case Assessment and Interpretation’ (CAI). The framework originated from within the mainstream forensic science disciplines, and features a strictly logical, probabilistic approach to addressing the facts in issue in criminal investigations, prosecutions and trials. Although, to date, most casework applications of CAI have been in the mainstream forensic sciences, the framework itself is perfectly general. As this *Guide* will show, the logical protocols of CAI can be adapted to enhance examination strategies and inform evidential analysis and interpretation in relation to any kind of expert evidence received in criminal proceedings.

This *Guide* is concerned with ‘forensic experts’, ‘expert evidence’ and ‘expert witness testimony’ in the generic sense.⁸ These terms include, but are by no means restricted to, mainstream forensic science. Unless the context specifically suggests otherwise, any references to ‘forensic scientists’ or ‘scientific evidence’ in the following pages are intended *mutatis mutandis* to encompass forensic expertise - i.e. scientific or other technical expertise applied to the administration of criminal justice - in general.

In this Part, we explain the intellectual impetus behind CAI and briefly sketch out its institutional origins. Some preliminary background context may assist fuller comprehension of the method, which is fully elucidated, with detailed illustration, in this *Guide*.

1.2 *Two Contrasting Styles of Forensic Investigation*

Consider two very different approaches to undertaking forensic examinations and reporting their results. The first, roughly corresponding to traditional forensic science methodology, we will label ‘Method 1’. The second orientation to forensic inquiries, anticipating CAI’s methodological assumptions and practical aspirations, we call ‘Method 2’.

⁸ Cf. Rule 33.1 of the Criminal Procedure Rules, defining ‘expert’ as any ‘person who is required to give or prepare expert evidence for the purpose of criminal proceedings, including evidence required to determine fitness to plead or for the purpose of sentencing’.

Method 1: Adopting the first approach, the forensic scientist reads the request from the police or prosecutor,⁹ opens the evidence bags and begins examination of the items using whichever analytical methods or techniques she feels are appropriate. (Throughout this *Guide* we will use ‘technique’ broadly to encompass any method of examination, including observational, physical, chemical, instrumental, etc.) The scientist needs very little in the way of circumstantial information about the case before starting the work. Having made observations using the selected technique(s), the forensic scientist interprets the results of the examination and reports her findings. The scientist is very much led by whatever is found, adjusting her interpretations and examination strategy as she proceeds.

Method 1 seems well-suited to *investigative* situations. It is redolent, perhaps, of the classic, popular view of the forensic scientist’s role as auxiliary detective. This approach to forensic inquiry requires minimal assessment of which (if any) forensic examinations will really address effectively the issues in the case, nor does it adopt a methodical approach to interpreting the evidential significance of any eventual findings. This suggests a major limitation to Method 1. Interpretation of the results of the examination takes place only after those results are already known. Method 1 is consequently especially prone to instrumental justification and *post hoc* rationalisation of scientific findings in the light of extraneous case information. Scientific evidence might, for example, be interpreted to fit in with prevailing case theories or police views of the most likely culprit. It has been shown that unconscious biases of one kind or another can surreptitiously influence forensic findings (see e.g. Budowle et al 2009; Risinger et al 2002). Unconscious bias is not necessarily a serious problem in investigative work (depending on the type of forensic investigations undertaken) but may become a contentious issue in later court proceedings.

The methodological weaknesses inherent to Method 1 suggest the need for a different approach.

⁹ Or if the scientist is being instructed by the defence, from defence legal representatives. However, the vast majority of forensic examinations in the UK’s legal jurisdictions are in fact commissioned by the police or by prosecutors.

1.3 **Method 2:** This involves a markedly different approach to conducting forensic investigations. Before carrying out any substantive examinations or analytical tests, the scientist makes a conscious effort to think through, in a structured and systematic way, the contribution that forensic inquiries might make to clarifying the issues in the case. To this end, she carefully considers what kinds of examinations or investigative strategies will address those issues most effectively, and tries to assess the potential significance of the various observations that might be made by adopting alternative or complementary approaches. In this way, the scientist's choices of specific techniques and overall investigative strategy will be rationally defensible. They will be based on documented, justifiable scientific appraisals tailored to the demands of particular investigations or prosecutions. And, moreover, the scientist's interpretations of her findings and any opinions she might express regarding their probative significance in criminal proceedings will be informed by a balanced *ex ante* appraisal of all logically conceivable potential observations. Risks of confirmation bias or circular, *post hoc* justification should consequently be minimised.

In contrast to the first method, Method 2 is predicated on a formal assessment of the examination strategy and a systematic forecast of the value of the potential results of examinations.

1.4 Whichever general method is adopted by forensic practitioners (and it may be that either or both will be used at different stages in some types of case), finite resources must always be deployed wisely and effectively. This may appear self-evident; and it might be assumed that individual scientists, and the organisations to which they belong, adhere routinely to this requirement. However, if commissioners of expert forensic services are not discerning clients, or if the providers of these services have personal or institutional preferences for certain techniques, then the work carried out may not be effective or economically justifiable in answering to the needs of particular cases.

1.5 'Case Assessment and Interpretation' (CAI), as we have said, provides a model for a structured approach to forensic investigations and the provision of expert evidence. It helps forensic scientists, other forensic experts, investigators, lawyers and managers to make rational decisions about two related but conceptually distinct phases of forensic inquiries, namely:

- ***The assessment phase:*** during which the forensic scientist formulates an effective investigative strategy and selects appropriate techniques tailored to the requirements of a particular case; and
- ***The interpretation phase:*** during which the forensic scientist interprets the meaning of any findings and, possibly, attempts to assess their probative significance for the case.

CAI provides a coherent intellectual framework and associated practical protocols that promote, within each individual case, intelligent decision-making about which items to examine and which techniques to employ. It guides the scientist in report writing and helps experts to formulate and communicate their findings, including any interpretational opinions, in a manner best calculated to support the proper administration of justice.

1.6 ***Potted Institutional History***

The concept of CAI was first developed in the late 1990s, primarily in response to two, potentially competing pressures generated by the introduction of a commercial market for forensic science services in England and Wales (Touche Ross 1987; Cook et al 1998; Jackson and Jones 2009). The first pressure was the requirement to improve the form and reliability of scientific opinions presented in court; the second was the need to provide value-for-money for clients and stakeholders.

During the 1970s and 1980s, the contribution and impact of forensic science in criminal investigations and prosecutions began to expand markedly, partly owing to the application and development of innovative, powerful, scientific techniques culminating, a little later, in the advent of DNA profiling. But at the same time, very little attention was being paid to interpretational models or to the practical development of efficient, effective ways of communicating scientific findings and expert opinion. Forensic scientists employed a variety of words and phrases to express their opinions. There was little consistency across different disciplinary specialisms, or even between individuals within the same scientific field. Such terminological inconsistencies remain pervasive within forensic science to this day, and

continue to generate concern (e.g. NAS 2009; Kaufman 1998) and pose difficult operational issues for courts.¹⁰

Happily, there is a viable alternative to this manifestly unsatisfactory state of affairs. Having originally been developed within the – since disbanded – Forensic Science Service (FSS) in England and Wales, the generic principles and methods of CAI are suitable for extension to all areas of forensic expertise and, arguably, within all legal jurisdictions.

- 1.7 Communication of reliable scientific evidence should be the aim of every expert witness. This is now an explicit requirement of the Criminal Procedure Rules in England and Wales.¹¹ Yet in some notable (and now notorious) cases, flawed scientific evidence has contributed to successful appeals against conviction. In a string of cases originally tried in the 1970s – John (Jack) Preece,¹² Judith Ward,¹³ the Birmingham Six,¹⁴ the Maguire Seven¹⁵ – it later transpired that the strength of the scientific evidence presented at trial had either been overstated or misinterpreted. Successful appeals against conviction in these cases, after the accused had already spent many years in gaol, were based partly on convincing argument that scientific evidence had been misleading or unreliable. While unreliable evidence could be the result of poor scientific techniques, or poor technical knowledge on the part of the scientist, deficient interpretation of analytical results or poor communication of the evidence by expert witnesses at trial are equally, or sometimes the more, significant factors.

By the early 1990s, implication in high profile miscarriages of justice had dented forensic science's previously-held and rarely questioned reputation for reliability and evidential

¹⁰ See e.g. *R v T* [2011] 1 Cr App R 9, [2010] EWCA Crim 2439; *R v Smith (Peter)* [2011] 2 Cr App R 16, [2011] EWCA Crim 1296; *State of Western Australia v Dair* [2006] WADC 157.

¹¹ Rule 33.2 Crim PR provides that '[a]n expert must help the court... by giving objective, unbiased opinion on matters within his expertise.... This duty overrides any obligation to the person from whom he receives instructions or by whom he is paid'. Rule 33.3 proceeds to list a series of factors that experts' reports must address, most of which are, in one way or another, concerned with ensuring the validity and reliability of the expert's evidence and the manner in which it is expressed.

¹² *Preece v HM Advocate* [1981] Crim LR 783, CA.

¹³ *R v Ward* [1993] 1 WLR 619, (1993) 96 Cr App R 1, CA.

¹⁴ *R v McIlkenny* (1991) 93 Cr App R 287, CA.

¹⁵ *R v Maguire* (1992) 94 Cr App R 133, CA

trustworthiness. Various institutional measures were canvassed, and some significant reforms were introduced (e.g. the creation of the Criminal Cases Review Commissions), to try to insulate the criminal justice process from further miscarriages of justice (RCCJ 1993; House of Lords Select Committee on Science and Technology 1993). Within the forensic science community, serious attention began to be devoted to improving the ways in which scientists routinely interpreted their findings and presented their evidence in court.

- 1.8 With few exceptions, forensic scientists had traditionally undergone very little formal training in the skills of interpretation, beyond ‘learning by apprenticeship’. Trainee scientists worked alongside more experienced colleagues, watched what they did, listened to how they expressed their opinions, and attempted to emulate their more experienced colleagues in their own evolving practices. Reporting protocols were devised and adapted to meet local requirements, inevitably producing an array of divergent practices. In particular, there was limited consistency in the formats or language used to express scientific findings or to interpret their (potential)¹⁶ probative significance for the proceedings. It is a moot point how much has really changed in the ensuing years, in the UK or overseas.

Furthermore, until the 1980s, forensic science training in the UK contained no formal instruction on the logic of inferential reasoning (the general topic addressed by *Practitioner Guide* No 3 in this series). Statistical training was also very limited in scope, focussing chiefly on the application of ‘significance techniques’ (e.g. in comparisons of measurements of physical or chemical properties of glass fragments). A few isolated early papers explored potential forensic applications of Bayes’ Theorem and likelihood ratios in modelling inductive inference (e.g. Finkelstein and Fairley 1970). However, it was a paper by a UK-based statistician, Dennis Lindley (1977), that stimulated Ian Evett, a forensic scientist and statistician working for the FSS in England and Wales, to develop a systematic approach to interpreting transfer evidence in forensic practice based on likelihood ratios (Evett 1983, 1991). The likelihood ratio approach to evaluating evidence is explained from first principles in *Practitioner Guide* No 1. It provides a logically justifiable, coherent intellectual framework

¹⁶ In one, non-trivial sense, the probative value of the evidence is whatever value is placed on it by the fact-finder in any particular case. Forensic scientists aim to assist factfinders in criminal adjudication to arrive at sound verdicts. In this sense, the forensic scientist’s own interpretation of the strength of scientific findings is always provisional and subject to the factfinder’s ultimate appraisal.

for forensic scientists, or anybody else, to think about and manage the uncertainties that must be addressed in formulating forensic strategies and providing effective assistance to the administration of justice. This covers, in particular, giving practical advice to police or other investigators involved in criminal proceedings and providing expert evidence suitable for presentation in criminal trials.

1.9 A second, reinforcing impetus for change was more prosaic. Government initiatives in the late 1980s encouraged public services to be more cost-effective and efficient in their operations and service provision. Direct charging of ‘clients’ for services rendered was extended to parts of the public service that had not previously been subjected to market pressures. Forensic science services in the UK had hitherto been centrally funded from local and national government, with very little direct connection between the supply of services and perceived consumer demand (Ramsay 1987). An influential report by management consultants Touche Ross (1987) contained proposals for the re-organisation of forensic science services along more commercial lines. This report kick-started a process of transformation in England and Wales, which began with direct payment for services and has produced, some 25 years later, a completely commercial marketplace for the purchase of forensic scientific assistance. In such commercial environments, the attention of providers and clients naturally turns to achieving ‘value-for-money’ in purchasing decisions, based on rational cost-benefit analysis. However, while costs can be monitored relatively easily, the benefits of forensic science have proven difficult to measure.

1.10 Forensic science providers responded in various ways to the challenges of forging new relationships with clients and driving the internal changes necessary to make that relationship work. As part of this package of measures, the FSS set up a project team comprising forensic scientists with complementary skills in casework, statistics and client-relations, to develop guidance on determining case strategy and interpreting findings. The overarching aim of the resulting guidance, by now known as the model for Case Assessment and Interpretation, was ‘to enable decisions to be made that will deliver a value-for-money service meeting the needs of our direct customers and the criminal justice system’ (Cook et al 1998).

The CAI model developed by the FSS initially focussed on scientific evidence presented in court. The benefit of a likelihood ratio approach to interpreting forensic results was naturally conceptualised in terms of its effectiveness in communicating the probative value of the

evidence to criminal justice professionals and, ultimately, to fact-finders. The model comprised three broad stages, each with its own prompts to guide the scientist. The first, critical, stage involved identifying the issue(s) for which the scientist was expected to provide assistance. This would necessitate good communication and effective co-operation between forensic science providers and their criminal justice ‘clients’. The second stage comprised the scientist’s assessment as to whether the requested work was likely to produce useful findings; which is not the same as asking whether the findings would enhance the instructing client’s case. The results of scientific testing could favour either the prosecution or the defence.¹⁷ This assessment required a formal, concise appraisal prior to any substantive work being undertaken. It was anticipated that such an appraisal would both inform clients’ decisions on the desirability of commissioning particular forensic inquiries, and also assist scientists to avoid *post hoc* rationalisation of their results. Having agreed a schedule of work with the client, forensic testing could be conducted and results obtained in the normal way. The third and final stage of the CAI model was an evaluation of the extent to which those scientific findings helped to make progress in resolving the factual issue(s) under investigation. The expert’s input might help to shape subsequent police inquiries (possibly involving further forensic testing), or be communicated to fact-finders in the form of a scientific report or as expert witness testimony adduced in a criminal trial.

- 1.11 The details of the CAI model were tested, developed and refined through numerous workshops attended by scientists representing all mainstream disciplines across the Forensic Science Service. It was calibrated to reflect the practice of the most effective scientists. Presentations on the model were also given to representatives of the police, Crown Prosecution Service, Criminal Bar Association and the judiciary. Feedback received at these sessions provided further helpful insights into the views and expectations, both of direct users of forensic science services and of the broader constituency of criminal justice professionals. The model was tweaked accordingly.

As well as helping to build up a more useful and resilient service, better able to meet customer requirements and to withstand adversarial changes in criminal proceedings,

¹⁷ Note that para.3.5 of the Code of Practice issued pursuant to Part II of the Criminal Procedure and Investigations Act 1996 stipulates that ‘[i]n conducting an investigation, the investigator should pursue all reasonable lines of inquiry, whether these point towards or away from the suspect’.

additional unpredicted benefits flowed from wider implementation of the CAI model. The role and range of forensic scientists' distinctive contributions to the proper administration of justice were clarified. Engagement with the model has produced fuller and deeper understanding of the skills and knowledge required of a competent scientist. Above all, CAI has informed operational forensic science by clarifying what information is required to provide balanced, logical, transparent and robust evaluation of scientific evidence.

1.12 The aspirations of CAI are closely aligned with developments in criminal procedure law and practice during the ensuing years. Rule 33 of the Criminal Procedure Rules, applicable in England and Wales, requires forensic scientists (and other expert witnesses) to spell out various factors potentially bearing on the interpretation of their reports, including 'details of any literature or other information which the expert has relied on', 'all facts given to the expert which are material to the opinions expressed in the report, or upon which those opinions are based', any qualifications to the expert's conclusions, and – where pertinent – a summary of the range of expert opinion on any contested question and an explanation of the expert's own views and conclusions relative to that range of opinion. Expert reports must also 'contain a statement that the expert understands his duty to the court, and has complied and will continue to comply with that duty'. Compliance with these formal legal duties anticipates a logical approach to forensic inquiry, with a fully documented audit trail. Expert witnesses should be prepared to explain the significance of their findings, and justify their decision-making in developing a rational forensic strategy in consultation with their clients' instructions, if challenged by an adversarial opponent or called upon to do so by the trial judge.

1.13 In 2013, ACPO and the Crown Prosecution Service agreed a new national 'Streamlined Forensic Reporting' (SFR) initiative, with a view to reinforcing the policies of proactive judicial trial management and more effective pre-trial preparation implemented by the Criminal Procedure Rules since 2005. Specifically:

SFR is a revised case management procedure for producing forensic evidence at court and seeks to reduce unnecessary costs, bureaucracy and delays in the criminal justice system. The process takes a more proportionate approach to forensic evidence through the early preparation of a short report that details the key forensic evidence the prosecution intend to rely upon (ACPO and CPS 2013: 3).

Early identification of the facts in issue and the development of effective scientific examinations adapted to the demands of the instant case are integral features of CAI. Thus, CAI is fully compatible with SFR, and directly supports the broader objectives of the Criminal Procedure Rules, including ‘early identification of the real issues’, ‘ensuring that evidence, whether disputed or not, is presented in the shortest and clearest way’, and the ‘overriding objective... that criminal cases be dealt with justly’.¹⁸

¹⁸ CrimPR, rr. 1.1 and 3.2

2. Principles of CAI and the Role of the Forensic Expert

2.1 *Fundamental Principles of Case Assessment and Interpretation*

The CAI model for forensic case assessment and interpretation rests on certain foundational notions and fundamental principles, which were clarified and refined as the model was developed, taking appropriate account of criminal practitioners' input and feedback. Successful implementation and effective uses of the CAI method, described in detail by Part 3 of this *Guide*, can be greatly facilitated by an appreciation of the underlying principles reviewed in this Part.

Specifically, CAI:

- clarifies the role of forensic expertise in criminal investigations, highlighting a vital distinction between *investigative* advice and *evaluative* opinions;
- identifies the different forms of logical reasoning characteristic of expert assistance in its investigative and evaluative modes;
- provides an illuminating taxonomy of the formulations currently routinely employed in forensic practice to report scientific findings, covering a spectrum ranging from hard scientific facts to evaluative expert opinions;
- rests on a rigorous logical method for evaluating the results of forensic examinations probabilistically;
- explains how the form in which evaluative opinion is expressed maps onto a 'hierarchy of issues', such that the probative value of the evidence may change according to the issue addressed; and
- enables discrete evaluations of particular scientific inquiries to be amalgamated into a single evaluative opinion, addressed to issues at activity level (the level which may, at least sometimes and in certain circumstances, provide greatest assistance to criminal investigators and fact-finders).

This Part explains each of these principles and underlying assumptions, and identifies their practical implications for criminal proceedings.

2.2 *The Role of the Forensic Expert (in Reducing Uncertainty)*

What do forensic scientists (and other expert witnesses) do? Standard responses to this question run along the lines that forensic experts provide assistance to the police and to the courts on matters of scientific, technical or other specialist knowledge beyond the general competence and ordinary commonsense experience of judges, magistrates, justices of the peace, or lay fact-finders. Hence, expert evidence is admissible in England and Wales when it is anticipated to be *helpful* to the fact-finder, whereas ‘[i]f on the proven facts a judge or jury can form their own conclusions without help, then the opinion of an expert is unnecessary’.¹⁹ The basic idea, as reflected in Part 33 of the Criminal Procedure Rules,²⁰ is that experts provide objective unbiased assistance to the courts on technical questions.²¹ Most people are well aware that this model of forensic expertise is simplistic and that notions of ‘objectivity’ and ‘bias’ can be problematic in practice (especially, perhaps, in adversarial systems of criminal procedure), but the model has been historically influential and still today represents, broadly speaking, the prevailing popular view of the forensic scientist’s role.

2.3 A different way of thinking about the role of the forensic expert, and one which has proved highly illuminating across a range of applications, focuses on the expert’s (epistemic) contribution to reducing the uncertainties with which criminal justice process decision-makers must grapple. That is to say, forensic experts help by reducing uncertainty. We might then reformulate the expert’s primary role as a duty:

To provide information that helps reduce the uncertainty of a material fact in forensic (investigative or judicial) settings.

¹⁹ *R v Turner* [1975] 1 QB 834, 841-2, CA.

²⁰ ‘An expert must help the court to achieve the overriding objective by giving objective, unbiased opinion on matters within his expertise’: CrimPR r.33.2(1).

²¹ Also see *Davie v Edinburgh Magistrates* [1953] SC 34, 40, where Lord President Cooper formulated the classic common law conception of the role of scientific experts: ‘Their duty is to furnish the judge or jury with the necessary scientific criteria for testing the accuracy of their conclusions, so as to enable the judge or jury to form their own independent judgment by the application of these criteria to the facts proved in evidence’.

Two aspects of this definition require further explanation. First, how – in what sense – do forensic experts contribute to reducing uncertainty? Secondly, the distinction between investigative and judicial settings merits emphasis.

2.4 ***Uncertainty, Probability and Inferential Reasoning***

Forensic uncertainty is reduced when experts supply new information, knowledge or opinions which assist criminal practitioners and fact-finders in their decision-making. Investigators, lawyers, judges and lay fact-finders arrive at conclusions about contested issues of fact through inferential reasoning from evidence (combined with their pre-existing, socially influenced, knowledge, beliefs and expectations). Part of their evidence base is supplied, in many cases, by forensic scientists or other expert witnesses. Forensic scientists arrive at the conclusions presented in their evidence employing essentially the same processes of inferential reasoning. Inferences are built up on inferences, until the relevant decision-maker – e.g. a police investigator deciding which lines of inquiry to pursue; a prosecutor deciding what charges to bring; a judge making an evidentiary ruling on the *voire dire*; or a criminal jury determining its verdict – arrives at factual conclusions informing the decision-maker's executive judgement. Inferential conclusions are (only) as good as the evidence (data) and intermediate inferential conclusions on which they rest. A structure of inferential reasoning could be a box-iron girder bridge or a house of cards.

Although the basic forms and procedures of human inferential reasoning are universal and common to all human beings, particular reasoning tasks attract additional institutional demands and constraints. Enhanced epistemic demands are characteristic of criminal adjudication. For example, judges are generally obliged to give reasons for their decisions (Ho 2000), and certain kinds of reasons may be mandated, or excluded from consideration, on normative grounds. In this sense, 'legal rationality' imposes more onerous conditions than ordinary, common sense inferential reasoning. Likewise, forensic scientists are subject to enhanced duties of rigour and transparency in their inferential reasoning. This *Guide* spells out the core requirements of these enhanced duties of rationality.

Practitioner Guide No 3 provides a comprehensive general introduction to inferential reasoning in forensic contexts. Here, we briefly recapitulate the characteristics of the three principal forms of inferential reasoning, and indicate their applications in forensic expertise.

2.5 (i) **Deduction** is the process of forming categorical inferential conclusions from fixed premisses in ‘closed-set’ contexts. It is a rule-bound, deterministic process, supplying certain knowledge provided that specified preconditions are met. For example, if we know that a man committed the crime, and we also know that Adam is the only man in the world, we can deduce (infer by deduction) that Adam is the criminal.

This exceedingly artificial example flags up the practical limitations of deduction in forensic contexts. Very few, if any, forensic science questions exist in closed-set situations. The inferences drawn by forensic scientists are conditional, probabilistic and subject to revision in the light of further information. Most forensic science questions are therefore not amenable to exclusively *deductive* answers (though deduction is routinely employed in a more informal sense, often as a loose synonym for induction).

2.6 (ii) **Induction** is the process of inferring generalisations from observed data. For example, if ten cars are observed to stop at a red light whilst traffic continues to flow when the lights are green, one might infer that the eleventh car to approach a red light will likewise also stop. This conclusion (prediction) is arrived at by formulating a generalisation, known by inductive inference from discrete observations, of the following type: ‘Cars (generally) stop when the lights are red (but go on green)’. This is *not* certain knowledge of the kind produced by deduction.²² Rather, it produces a reasonable working assumption, which may need to be revised in light of further pertinent data (e.g. the eleventh vehicle turns out to contain an escaping felon, who runs the red light). Inductive inference is inherently fallible and revisable.

Another important way of describing inductive inference is that it is *probabilistic*. It is sometimes possible to quantify the probability of inductive inferences in quite precise ways, as where we have robust statistical data showing that 23% of forensic scientists are avid watchers of *CSI*, or that eight out of ten cats prefer a particular brand of pet food. Probabilistic judgements are also routinely expressed using vague linguistic expressions

²² A generalisation of the form ‘Cars always stop on red (but go on green)’ would be fallacious, since one cannot validly infer an exception-free rule from past observations (a thought often attributed to Hume). Even if every one of the ten million swans you have ever seen were white, this does not exclude the possibility that the next one you see could be black.

employing ‘fuzzy quantifiers’ like ‘usually’, ‘generally’, ‘often’, ‘sometimes’, ‘occasionally’, etc. Even ‘more likely than not’ is a statement of probability. So is ‘equally likely’.

The probabilistic nature of inductive inferences makes them suitable for evaluation using formal methods and models, including Bayes Theorem. *Practitioner Guide No 1* explained how Bayes Theorem, a logical derivation of elementary mathematical axioms, enables one’s assessment of the uncertainty of any (past, present or future) event to be updated through the incorporation of relevant new information.

2.7 Bayes Theorem is the embodiment of inductive inference, with significant forensic applications. It offers a strictly logical method for revising subjective views of the truth of any proposition or legally salient fact in issue in the light of new evidence (generally, see Redmayne 1998; Lempert 1977). Bayes Theorem can be applied over and over to up-date existing probabilities to take account of additional pieces of information, as new discoveries emerge over time or existing information takes on new significance as an investigation unfolds.

Bayes Theorem features two key probabilistic components:

- the ratio of ‘prior probabilities’ of (i) a disputed fact in issue (or other matter of interest) and (ii) some specified mutually exclusive alternative; and
- the likelihood ratio (a ratio of two probabilities) for relevant new evidence, e.g. the expert’s observations or interpretative findings.

These two components are combined together, in accordance with standard probability axioms, to produce an updated ratio of ‘posterior probabilities’ assessing the truth of the disputed fact in issue relative to the specified alternative. This is only a skeleton summary. The rudiments of Bayes Theorem are more fully explained in *Practitioner Guide No 1*.

2.8 For many of the issues addressed by expert evidence, forensic scientists are not in a position to assign appropriate, informed or realistic prior probabilities for material facts in issue; nor is it their role to do so. Prior probabilities are typically informed by inferential conclusions that can appropriately be made only by fact-finders in criminal adjudication. It is not the expert’s

role to pre-empt or displace jury reasoning on these issues (and any institutional pressures that they may experience in this direction should be resisted by conscientious forensic scientists.)

What an expert should be in a position to offer is the assessment of a likelihood ratio (LR) for the evidence. The LR is the ratio of two probabilities, conditioned on mutually exclusive (but not necessarily exhaustive) propositions. In a forensic context, the LR can be explained generically as the ratio of: (1) the probability of the forensic scientist's observations, if the postulated fact in issue were true; to (2) the probability of the same observations on some alternative hypothesis (e.g. that the fact in issue is false). The expert presenting a likelihood ratio makes no assumption either way about the truth or falsity of the fact in issue, and therefore cannot be criticised for deliberately or inadvertently usurping any of the jury's responsibility for fact-finding. (Likelihood ratios were also explained in *Practitioner Guide No 1*, and employed in the evaluation of DNA profiling evidence in *Guide No 2* in this series.)

2.9 **(iii) Abduction**²³ describes the intellectual and imaginative process of generating possible explanations to account for an expert's actual or anticipated scientific observations. Abduction is a special case of induction, whereby potential explanations for events in the world are improvised through a blend of experience, creative thinking and intuition. Abductively-generated explanations may then be tested against observed data in the real world. Systematic testing of imaginative experimental hypotheses through empirical scientific methods is a familiar, and socially very important, species of abductive inference.

Abductive reasoning also plays a vital role in the administration of criminal justice, especially at the earlier stages of an investigation when relevant information is invariably incomplete and may turn out to be unreliable (generally, see Tillers and Schum 1991). Forensic experts can assist police detectives in formulating explanations that could account for whatever evidence is currently in-hand, and which may point to promising avenues for further investigation likely to test these initial hypotheses or generate pertinent new information. Given almost literally limitless empirical possibilities, some means of narrowing down the

²³ Terminology is not entirely consistent, especially across different disciplines. We offer here what we take to be the standard, and most useful, conception. Also see Schum (2001).

scope of an investigation is absolutely essential. Forensic scientists can draw on their past experience and knowledge to offer investigators realistic explanations for scientific findings, and thus assist in the development of rational and efficient investigative strategies. But the limitations of abductive reasoning, grounded as it is in a mixture of analytical technique and educated guesswork, should be highlighted and explained to those whose role it is to evaluate the probative value of the expert's observations at different stages of the proceedings.

2.10 *Distinguishing investigative advice from evaluative expert opinion*

A key realisation emerging from the early development of CAI was that forensic experts routinely provide two, significantly different kinds of assistance to other criminal justice practitioners, depending on the stage that a particular investigation or prosecution has reached from crime-scene through to court. Moreover, these two different functions require different kinds of inferential reasoning, leading to distinct analytical products and forms of communication (Jackson 2009; Jackson et al 2006).

Generally, when in *investigative* mode, experts (consciously or otherwise) utilise *abductive* reasoning to offer *explanations*. If prior probabilities for such explanations can be made explicit and appear realistic, the expert may also offer *posterior probabilities* for those explanations. These two elements – abductively-generated *explanations* and *posterior probabilities* for the explanations – constitute *investigative advice* of a type that is generally appropriate before a suspect is apprehended. They help with the perennial investigative question: who done it? Investigative opinions may also be offered postulating explanations for *how* the crime was committed. ‘How done it?’ questions and hypotheses remain in play even after questions of identity have been settled, e.g. the only realistic suspect is already in police custody but the prosecution will still need to prove how the crime was committed.

When a case proceeds to trial, expert reports or testimony should be focussed on disputed facts in issue (such as the identity of the perpetrator or the nature or source of physical evidence). Expert opinion incorporating a *likelihood ratio* representing the expert's considered evaluation of the evidence should, in principle, be helpful to the fact-finder in assessing the evidence as a whole – provided, of course, that such an opinion is germane to the proceedings and is presented in a comprehensible form. (Note that LRs can be expressed in numbers or words, or both.) Expert evidence of this type, incorporating an evaluation of a likelihood ratio, is appropriately characterised as *evaluative opinion*.

2.11 *A Taxonomy of Expert Evidence, from Facts to Opinions*

We have established that the role of forensic experts can be conceptualised in terms of their contribution to reducing uncertainty about facts in issue, and that this function is discharged through two distinct modes of inferential reasoning and outputs, corresponding to different phases of criminal proceedings. We labelled these two modes of expert assistance, respectively, investigative advice and evaluative opinion.

‘Opinion’ is a difficult concept in the law of criminal evidence (see Roberts and Zuckerman 2010: § 4.3). In one sense, all testimony is a matter of opinion, inasmuch as it reports witnesses’ inevitably subjective impressions of the world. However, the common law adopts a pragmatically robust approach in distinguishing ‘facts’ from ‘opinions’. What the law treats as ‘fact’ is in reality an unproblematic, uncontested inferential conclusion. So the witness is allowed to testify that he saw Eve, whom he knows, provided that the identity of the person whom he saw on the occasion in question is not a disputed fact in the litigation. If it is, the witness will be cross-examined on the more basic evidential data which he used to draw the inference that the person he saw was Eve; e.g. that the person he saw is the same height as his friend Eve, has the same hairdo, was wearing distinctive clothing similar to Eve’s wardrobe, or whatever. When it is said (not very helpfully) that witnesses must testify to facts not opinions, it is the forensically appropriate granularity of inference that is really at stake. Equally, when it is said (equally unhelpfully) that expert evidence is an exception to the opinion evidence rule, what is really meant is that experts are sometimes permitted to present complex inferential conclusions as composite, unelaborated, coarse-grained ‘facts’ (mainly because decomposing the evidence into its more fine-grained components – data and inferences – would not assist the fact-finder, given the specialist technical nature of expert evidence).

Further practical experience with the CAI model and reflection on its implications for existing forensic practice suggested the following, more refined five-fold classification of the types of information that may be contained in expert evidence. This taxonomy comprises: (i) factual information; (ii) categorical opinions; (iii) explanations; (iv) posterior probabilities, and (v) probabilities for observations (preferably the combination of two probabilities expressed as an LR). An expert’s report or testimony at trial may, and often does, combine two or more types of information. There is also a sixth category of expert opinion writing, which should be disfavoured for the reasons we explain.

2.12 **(i) Factual information:** This is the simplest kind of information provided by forensic scientists, comprising factual reports (undecomposed coarse-grained inferential conclusions) stating what the expert did, saw or found. For example:

- *I attended the scene and I noted extensive bloodstaining on the wall above the head of the bed.*
- *I examined the balaclava and recovered ten fragments of glass.*
- *I found deposits of soil embedded in the sole of the left shoe.*

Such statements are generally made and accepted as a matter of routine practice and convention. Of course, their veracity could in principle always be challenged, e.g. in terms of evidential continuity. But they involve very little, if any, active interpretation on the part of the expert. They are of some help to investigations and courts in providing evidential material and filling out the factual background to the case but, by themselves, do not offer any guidance on the significance, impact or weight that could be attached to the bare facts.

2.13 **(ii) Categorical opinions:** There are two basic variations on categorical opinions. One kind categorically excludes a relevant evidential possibility. The second type affirmatively asserts the truth of a factual proposition with a material bearing on the case. The nature of factual propositions as statements with truth-responsive predicates, and some of the complexities of formulating them in forensic contexts, are discussed in *Practitioner Guide No 3*. The essential idea is that factual propositions can be *either* true or false (not simultaneously both or neither). Categorical opinions assert that certain propositions are definitively, absolutely true (or as the case may be, definitively, absolutely false).

Categorical exclusions are in principle justifiable on logical (epistemic) grounds, and may legitimately be offered in expert opinions provided that key stated assumptions are acceptable to the court. Examples of justifiable exclusions (with their assumptions spelt out) might include the following:

Illustration #1: Footwear mark exclusion

Finding: The footwear mark on the tiled floor at the scene of the crime is different from the pattern of the sole of the suspect's shoe

Conclusion: *This shoe cannot have made the mark* (categorical exclusion)

Assumption: the sole of the suspect's shoe has not been altered or changed appearance in the meantime

Illustration #2: DNA profile exclusion

Finding: The DNA profile obtained from semen-stained swabs is different from the DNA profile of the defendant's saliva sample

Conclusion: *The defendant cannot be the source of the semen on the swabs* (categorical exclusion)

Assumptions: the DNA profile that has been obtained from the swabs is actually from the semen on the swabs; the defendant's sample is actually from the defendant; all profiling has been done accurately and without any sample-switching or contamination

2.14 Categorical assertions take various forms in modern forensic science, but two characteristic types are commonly seen.

The first type of categorical assertion expresses expert judgements of classification. A standard question in forensic inquiry takes the form: 'What is this material?' Applying standard tests and protocols, the expert may be able to provide a definitive answer. Examples include:

- *The fibres found at the scene are polyester fibres.*
- *The material sent for chemical analysis is cannabis.*
- *The spent cartridge case is XXX brand ammunition.*

Strictly speaking, these conclusions are fallible inductions rather than unassailable and timeless deductive truths. They assume that the tests and protocols for classifying particular questioned materials or substances are sufficiently discriminating safely to rule out other possibilities. For example, there *could be* unknown substances that produce the same

chemical reactions as cannabis (but are not cannabis); there *could be* a brand of *XXY* ammunition which (perhaps unknown to the forensic scientist) is visually identical to *XXX*, but is not in fact *XXX*.

Categorical assertions in the form of classifications implicitly rule out alternative possibilities by negative inductive inference. Things that are not known are assumed to not exist. The classical logicians labelled this argument the fallacy *argumentum ad ignorantiam*, or argument from ignorance. In fact, such conclusions may represent perfectly sensible working assumptions in the absence of any better information indicating the contrary (Walton 1992). But whether this is a reasonable basis on which to base criminal verdicts is always, in principle, a contingent matter.

2.15 A second type of categorical affirmative proposition, known as ‘individualisation’, is (even) more problematic. Forensic examples of individualisation include:

- *This shoe made the mark at the scene.*
- *The piece of broken button recovered from the complainant’s bed came from the broken button seen on the suspect’s jacket.*
- *The fingermark is that of the left forefinger of the suspect.*

It is not uncommon to encounter expert witness reports or testimony asserting one or more categorical individualisations on this model. Another set of variations addresses ‘what happened?’ questions, providing categorical assertions or conclusions along the following lines:

- *The body must have been dragged through the hallway and into the kitchen.*
- *The origin of the fire was the electricity meter in the cupboard below the stairs.*

Categorical assertions, of identity or explanation, may well rest on a particular expert’s vast experience and proven competence. Yet all such expressions of categorical opinion are very difficult, if not impossible, to justify on strictly logical grounds. It has been argued that the ‘individualisation fallacy’ is a pervasive weakness infecting contemporary forensic practice

and that '[t]here is no scientific basis for the individualization claims in forensic sciences' (Saks and Koehler 2008: 202).

2.16 A categorical conclusion of individualisation logically rules out all other possible alternatives. The expert is saying *this individual person or item is the one*, to the exclusion of all others. To spell out the logical implications of our first three illustrations of individualisation, the expert would need to be completely sure that, respectively,

- It is *impossible* to obtain the degree of observed correspondence between the scene-mark and the suspect's shoe if a shoe other than the suspect's had made the mark at the scene.
- It is *impossible* to obtain the observed fit between the broken edges if the piece of broken button recovered from the complainant's bed had come from another broken button.
- It is *impossible* to see that number of matching features if the fingerprint is not that of the left forefinger of the suspect.

Events that are (mathematically) impossible have probability zero. However, real-world forensic issues rarely have zero probabilities. A particular probability may indeed be very small, even *virtually* zero, yet it remains strictly incorrect to describe an event with miniscule probability as 'impossible'. Inferential conclusions of individualisation thus remain personal, albeit possibly highly cultivated and reliable, expert *opinions*.²⁴ The process of arriving at judgements of individualisation seems to reflect the partly psychological progression of an expert's becoming 'personally convinced' that the questioned trace *must have* the same source as the reference trace or that one specific event must have occurred.

2.17 **(iii) Explanations:** Explanations in forensic opinions account for, or fit around, the observations that have been made in a case. They *explain* existing evidential information, in causal or other terms. As such, forensic explanations are generated and considered *after* relevant observations are known. They are generally produced through the expert's own

²⁴ www.thefingerprintinquiryScotland.org.uk/, p 741, Recommendations 1-3.

abductive reasoning, though on occasion they are suggested by others (e.g. in police requisitions for forensic testing) and the expert simply adopts and reiterates that explanation if her further observations are consistent with it.

Examples of explanations, given after expert or scientific observations have been obtained and compared, might include the following:

- *The mark could have been made by the defendant's shoe.*
- *The bloodstaining on the wall could have been caused by multiple blows to the deceased's head.*
- *The injuries are consistent with having been caused by a claw hammer.*
- *The defendant cannot be excluded as a source of the partial DNA profile seen in the mixture of DNA on the swabs.*

Characteristically, these examples, offer no assessment of the probability that the explanation is true. Such explanations are open-ended: there may be no limit to the number of competing or alternative explanations that could be generated to explain the findings. Of course, some of these explanations may be more or less speculative, implausible or fanciful than others. They may be helpful to investigators, in suggesting further lines of inquiry when the forensic expert is in investigative advice mode. However, 'bare' explanations are likely to be of limited assistance to fact-finders, and might even be regarded as potentially misleading and, sometimes, pernicious.²⁵

2.18 **(iv) Posterior probabilities:** Expert opinions formulated as posterior probabilities express the probability (in numbers or words) that a fact in issue or some other proposition of interest is true, given the scientific findings or observations that have been obtained. These are, in other words, probabilities *conditioned on* the information known to the expert (which should be spelt out clearly in the expert's report, in accordance with requirements imposed in England and Wales by the CrimPR). Examples would include:

²⁵ National Academy of Sciences (2009); Kaufman (1998); *State of Western Australia v Dair* [2006] WADC 157; *R v Puaca* [2005] EWCA Crim 3001. Cf *R v T* [2011] 1 Cr App R 9, [2010] EWCA Crim 2439.

- *The mark was very probably made by the defendant's shoe.*
- *It is likely that blood was deposited as a result of multiple blows to the head of the deceased.*
- *The type of weapon used to force open the door was highly likely to be a claw-hammer.*
- *The most likely cause of the partial, matching DNA-profile in the mixture of DNA on the swabs is that the defendant is a contributor to that mixture.*

These are aptly described as ‘posterior probabilities’ for contested issues, since they are derived and expressed *after* some new item of information, e.g. an expert’s observation or scientific result, becomes known. This form of opinion is unfortunately beset with a variety of problems, starting with the immediate challenge of assigning an appropriate prior probability for the relevant fact in issue.

2.19 *Practitioner Guide* No 1 explained that, in order to offer a posterior probability, the expert must have considered or implicitly assumed a (numerical or semantically fuzzy) value for the prior probability of the fact in issue. The risk here is that the expert’s prior probability may be inappropriate, given the circumstances of the particular case. The expert’s prior probability may fail to take account of relevant considerations (which might be unknown to the expert); it could be biased (consciously or otherwise) towards the instructing party; or it could be completely speculative.

Consider an expert engaged in comparing a sample of questioned handwriting with a specimen provided by a suspect. The expert observes what she considers reasonable agreement between the two samples, and reports a ‘match’. However, the questioned sample was only a small piece of writing and, let us say, it is not very distinctive. So the expert expresses the opinion (in the form of a semantically ‘fuzzy’, numerically unquantified posterior probability) that the suspect ‘*probably wrote the questioned item*’. The problem is that, in arriving at this opinion, the expert may have been subconsciously influenced by the logically extraneous knowledge that there was strong circumstantial evidence against this particular suspect. Her ‘prior probability’ for the proposition that the suspect is the writer of the questioned handwriting sample could be relatively high in this instance; certainly higher

than if the suspect were merely one member of a large pool of individuals trawled into a ‘mass screening’ operation.

2.20 Irrespective of the institutional role or identity of the person making the assessment, prior probabilities for issues in contention should be assigned based on all the relevant information known to that person prior to additional information or evidence being taken into account. It is essential that this stage in the reasoning process should be fully transparent: the adopted ‘priors’ should be clearly specified, and thus rendered open to later challenge or revision. It may be that, in some situations, the expert is able to provide further assistance to the court by providing data from which prior probabilities can be developed. The appropriateness of those data is always an important consideration for the court. However, in other situations, assigning prior probabilities is not an appropriate task for the expert, because it involves making selections between (potentially) contested facts, a role properly reserved to the fact-finder in criminal adjudication. (Of course a jury does not consciously articulate ‘priors’; rather, intuitive assessments of prior probability are subsumed within ordinary ‘common sense’ inferential reasoning.) In each and every case, lawyers, judges, and expert witnesses themselves should be alive to the fact that a prior probability has been adopted or assumed whenever it has been; and the appropriateness of that selection or assumption should be open to scrutiny, reconsideration and potential challenge – both in terms of its evidential (epistemic) soundness and its fairness, in favouring one side or another in adversarial proceedings. Failure to articulate or consciously scrutinise prior probabilities exacerbates the risk that an expert witness will inadvertently usurp the role of the fact-finder, by stating a conclusion conditioned on information for which the expert cannot vouch. In addition, concealed or inadvertent ‘priors’ pose the risk of evidential double-counting, whereby the fact-finder ascribes probative value to factors that have *already* been fully accounted for in the expert’s working assumptions. In this way, the probative value of the evidence may come to be overestimated.

Criminal lawyers and forensic practitioners also need to consider the possibility that expert opinions expressed in terms of posterior probabilities may be infected by fallacies of the transposed conditional, including the ‘prosecutor’s fallacy’, which were discussed in *Practitioner Guide* No 1. All of these errors or confusions share the same fundamental flaw: they imperil the rationality of criminal adjudication by presenting a false or misleading view

of the probative value of expert evidence, imperilling the fact-finder's efforts to utilise the evidence presented in the case to reach true conclusions of fact on contested factual issues.

2.21

(v) Probabilities for the observations (expressed as likelihood ratios): *Practitioner Guide*

No 1 explained the basic reasoning procedures for expressing conditional probabilities for observations, employing likelihood ratios. Expert evidence may be given in this form. With reference to relevant data, and drawing upon specialist knowledge and understanding, the expert is able to offer the court an assessment of the probability (or 'likelihood') of obtaining the expert's observed findings, conditioned on the assumption that either of two competing propositions were true. Such expert opinions might then take the form: 'my findings would have x probability if the prosecution's proposition were true; and y probability if the defence's countervailing proposition were true'. The probative value of the evidence is the ratio of these two probabilities, i.e. the likelihood ratio. The question then becomes one of how most effectively to convey this information to lawyers and courts, e.g. whether in the form of a numerical LR, in terms of a verbal scale of equivalents, or as a verbal explanation of the likelihood ratio ('this evidence would be z times more likely if the prosecution proposition [that...] were true, than if the defence proposition [that...] were true').²⁶

The pre-eminent virtue of reporting scientific findings in this conditional format is that the expert does not rely on concealed 'priors' or purport to advance conclusive factual findings, one way or another. The relevance and probative value of scientific findings are explicitly conditional on asserted facts that are not endorsed by the expert; and the consumers of such opinions are invited to arrive at their own informed evaluations of the probative value of the expert's observations, taking account of the likelihood ratio supplied by the expert (stating how many times more – or less – probable it would be to obtain the observations if the prosecution proposition rather than the alternative defence proposition were true). It is, of course, another matter whether the consumers of such opinions can interpret these expressions meaningfully. Lawyers and courts need to be able to understand the meaning of a likelihood ratio and to assess its evidential significance. Promoting such understanding is a central ambition of these *Practitioner Guides*. Whether lay fact-finders in criminal trials should be directly exposed to likelihood ratios, and whether jurors would understand them if

²⁶ Thus, z is the ratio of x/y . For example, if $x = 0.75$ and $y = 0.25$, $z = 0.75/0.25 = 3$.

they were, are difficult and controversial questions beyond the scope of the present discussion, which is concerned with professional education rather than institutional reform.

2.22

(vi) Other (vague and ambiguous) expressions: We observed in *Practitioner Guide No 1* that a range of other linguistic expressions has traditionally been used to express forensic scientific findings. Some of them have already been encountered, above, in the form of ‘explanations’. Stock phrases include: ‘provides a link between’; ‘there is evidence of association between’; ‘is consistent with’; ‘could have originated from’; ‘there are no significant findings’; ‘cannot be excluded’.

Some of these phrases have perfectly legitimate forensic uses. But they require careful contextual interpretation. Although they express conclusions that are strictly speaking true (or perhaps, ‘not false’), vague or ambiguous expressions may be incomplete to the point of becoming positively misleading to certain audiences. Thus, findings of consistency or potential origin may easily give the impression of strong association, whereas, in strict logic, findings may be ‘consistent with’ the prosecution’s allegation of guilt *and* with the defendant’s denials. To take an extreme case for the sake of illustration, a forensic scientist might accurately report that ‘the footwear mark is consistent with the accused’s shoe’, but fail to mention that ‘the footwear mark is also consistent with the co-accused’s shoe’, or with 90% of training shoes worn by young people, etc. Extremely unlikely possibilities may still be ‘consistent with’ the evidence. If all that is said is that two items ‘could have’ a common origin or that common origin ‘cannot be excluded’, it is doubtful that this would be understood, according to the ordinary canons of communication and conversation, to cover cases of vanishingly small probability.

Likewise, failure to define the meaning of key concepts such as ‘link’, ‘association’ or ‘significant’ leaves the receiver of the opinion to infer their own meaning, and hence the weight, that should be attached to the evidence. In light of the very real possibility that lay people will not fully comprehend forensic scientists’ technical usage, such phrases should be viewed critically. They would probably be best avoided altogether.

How Should Expert Evidence be Formulated?

We have reviewed six different presentation formats for expert evidence. Setting aside type (vi) as problematic for the reasons just given, the expert is still confronted with a choice between five basic alternatives (and most of them have further distinctive sub-variants). How should the expert choose between these alternatives?

There is no simple answer to that question, because the best way of expressing expert findings is likely to depend on the type of expertise in question, the stage reached in the proceedings, and the demands of the instant case. It also turns crucially on the particular expert's understanding of a forensic scientist's role in addressing the facts at issue. Some expert witnesses might prefer simply to 'state the facts' of their findings, as in type (i) above, and leave any further evaluation or inferential reasoning to the instructing professionals and the court. Others feel they should address the facts in issue more directly, either with a categorical answer (type ii) or in probabilistic terms (type iv). Still other forensic scientists believe it more appropriate to offer explanations (type iii); and finally, some experts strive to make the probative value of their findings more intelligible through the provision of a likelihood ratio (type v). There is plainly no broad consensus or internationally-agreed approach to formulating expert evidence across the forensic sciences, or even within particular forensic specialisms (albeit that influential voices in the field have positively endorsed type (v) opinions incorporating likelihood ratios (e.g. Association of Forensic Science Providers (2009) or discussed them with approval (National Research Council 2009: 6-3)). A variety of approaches can be seen whenever forensic experts publish articles, speak to the media, testify in court, or have their evidence summarised in counsels' speeches or judicial summaries.

Pluralism and contextual adaptations have merit, and there should be no attempt to impose a single unified approach across the entire, diverse, spectrum of expert evidence. However, certain generic considerations are worth emphasising. One significant dimension of expert evidence which the CAI model helped to clarify is that forensic scientists are engaged in addressing a range of different issues, which can helpfully be organised in terms of a conceptual hierarchy.

2.24 *The Hierarchy of Issues*

The notion of a ‘hierarchy of issues’ was introduced in a general way by *Practitioner Guide No 1*, and illustrated through its application to DNA profiling evidence in *Practitioner Guide No 2*. The proposed hierarchy comprises four ‘levels’ of issue: offence, activity, source and sub-source.

The CAI model facilitates more rigorous reflection on the different ‘levels’ of issue within a case to which expert evidence might be addressed. This point can also be grasped and visualised using formalised models of inferential networks, such as the Wigmore charts and Bayes nets discussed in *Practitioner Guide No 3*. The simple intellectual point is that the probative value of expert evidence is necessarily a function of the issue on which it bears, and this issue should be understood within the framework (or network) of issues within the case as a whole.

If we conceptualise ‘issues’ in terms of their granularity, we can think of a structural hierarchy running from coarse-grained ‘offence level’ issues at the top running down to very fine-grained ‘sub-source issues’ at the bottom. It makes sense to speak of a *hierarchy* of issues, in the sense that issues at the top of the hierarchy are logically closer to the ultimate question for the trier of fact – Is D guilty of *this* offence? – than issues near or nearer the bottom. In a Wigmore chart, offence level issues would typically overlap with ultimate and penultimate probanda, and they would be charted, literally, at the top; whereas activity, source and sub-source issues would appear, respectively, further down the chart. The proposed hierarchy is not, however, a measure of probative significance in particular trials. Sub-source level issues can be, and sometimes are, determinative of the contested facts in particular cases.

2.25 *Hypothetical Illustration: DNA Tights*

The operation, and heuristic value, of this framework are best illustrated through their practical application. This section examines an hypothetical case through the interpretative lens of the hierarchy of issues.

2.26 **Case circumstances:** A heterosexual couple (H + W) were confronted by two masked men in the bedroom of their marital home in the early hours of the morning. Armed with pick axe handles, the intruders demanded that H open up a safe located in the

bedroom. Whilst H fumbled with the lock, one of the masked men tied W's wrists together with a pair of tights that had been lying on the bedroom floor. After taking money and jewellery from the opened safe, both men departed. H used a pair of scissors to cut the tights from W's wrists and they called the police. On attending the scene, various items including the tights were recovered for possible forensic examination.

W informed police that she had worn the tights the previous evening whilst socialising at a busy city centre pub. She believed she recognised the voice of the man who tied her up but not the second man. Reference DNA samples were taken from both H and W.

Following enquiries, the police arrested a male suspect (S) approximately 10 hours after the incident. He did not possess any clothing fitting the description provided by the complainants of the offenders' apparel. A reference DNA sample was taken from him. Police intended to charge him with the offence. DNA examination of the tights and reference samples was requested with a view to building a case against the suspect. However, to save the cost of a full examination of the tights, the police submitted to their independent forensic science provider only one piece of fabric cut out by the police's in-house 'screening' laboratory, along with the reference samples from W, H and S.

2.27 **Assessment by the forensic DNA laboratory:** It would be relatively straightforward for the forensic science provider to comply with the police request for a DNA examination of the material provided. They could carry out the profiling techniques, produce a report describing any profiles that had been obtained from the fragment of submitted tights, and state whether there were any matches to the reference profiles. If a match with the suspect were obtained, this would generally be accompanied by some assessment of the significance of the match, usually in the form of a random match probability or likelihood ratio (see *Practitioner Guide No 2* for detailed explanation).

Whichever presentational format were adopted, the forensic test and its results would be directed to the *sub-source* level, i.e. the issue being addressed is the provenance of the profiled DNA (as opposed to the provenance of any biological material containing the DNA, which would be *source* level issue). In some situations this may be all that is required and would meet the needs of investigators and the court in a balanced, robust way. However, there are situations in which this straightforward approach to meeting 'consumer demand'

could potentially be misleading and productive of error and injustice. Likelihood ratios framed at sub-source level are capable of producing highly impressive-sounding values, of the order of millions or even billions (corresponding, for example, to the random match probability of 1 in a billion associated with full SGM Plus™ DNA profiles²⁷ for genetically unrelated individuals). However, reformulating the legally material issue at source, activity or offence levels is capable of dramatically reducing the value of the LR, depending on whether factors such as transfer and persistence of physical traces are being taken into account. These reformulations may *appear* to suggest very different probative values for the evidence, but this is a kind of forensic illusion: nothing about *the evidence* has changed. What is changing is the assessment of the (actual) probative value of the evidence *with regard to its impact on the probability of particular contested issues*. The crucial point to grasp is that ‘probative value’ describes a *relationship* between the evidence and the particular fact or facts it is taken to prove. A change of legally material factual issue routinely changes the probative value of the evidence in those proceedings; as one should logically expect.²⁸

2.28 Methodical application of the CAI model, requiring careful attention to the hierarchy of issues, concentrates investigators’ minds on the relevance and probative value of the evidence in relation to specified issues in the case. This intellectual discipline assists both client and forensic provider to focus on conducting scientific investigations and framing expert reports to provide maximum assistance (taking due account of the efficient use of resources) to the criminal justice system and to the courts. CAI’s emphasis on transparency and balanced evaluations is very much in keeping with the objectives of the Criminal Procedure Rules and the recently introduced programme of ‘Streamlined Forensic Reporting’ (ACPO and CPS 2013).

²⁷ The SGM Plus™ DNA profile has been ‘industry standard’ in the UK now for several years, though it is currently being replaced with newer generation profiling techniques employing 17 loci. Interestingly, the match probability for a full 17-locus matching profile will remain of the order of 1 in a billion.

²⁸ There are limiting cases. For example, evidence which is irrelevant (has zero probative value and a likelihood ratio of 1) in relation to one issue, may also be irrelevant (also have zero probative value, LR = 1) in relation to many other salient issues.

CAI anticipates, and should foster, a proactive attitude on the part of forensic scientists. On receipt of a request from the police or other criminal justice system ‘customer’, the CAI practitioner should try to gain a clear understanding of the needs of the case, where necessary looking beyond the narrow terms of particular requests. In our *DNA Tights* hypothetical, the CAI model would direct the provider to contact the police and obtain further details of the circumstances of the case. One important consideration in this type of scenario would be what, if anything, the suspect will say in his defence, if this is already known or can reasonably be surmised (e.g. by considering the suspect’s police interview or formal defence statement).

2.29 **Further information:** The forensic scientist in *DNA Tights* adopts a proactive CAI approach, and discovers the following information from follow-up inquiries with the police. The suspect denies the offence. He also claims to have been in the same pub as W during the evening, and recalls that his hand lightly brushed the legs of a woman several times, as he passed her on the stairs on the way to the bar. He cannot give a detailed description of the woman or of her clothing, but definitely remembers that she was wearing tights. W confirms that a man had lightly touched her legs in the pub that night. She is also certain that the intruder who tied her up was not wearing gloves. However, she cannot positively identify the suspect either as one of the intruders in her home or as the man who touched her legs in the pub.

This new information enables the forensic scientist to form a more complete picture of the case as a whole, which in turn allows her to structure her inquiries to produce more useful and balanced findings.

2.30 In the light of the new information, the relevant issue can be reformulated as a question posed at *activity* level:

Is S the intruder who tied W’s hands with the tights?

Activity level issues concern conduct, chiefly acts and omissions. They are distinguishable from *offence level* issues, in that mere or ‘bare’ activities typically only satisfy offence definitions through some additional normative classification – in this case, transforming a ‘tying’ (activity level) into an ‘assault’ or ‘robbery’ (offence level). This distinction is related

to the important conceptual dichotomy between evidential facts (going to prove what the accused did, e.g. ‘put his hand through a window’) and constitutive facts (those which constitute offences, partly through acts of normative classification: e.g. ‘made a burglarious entry’: see further, Roberts and Zuckerman 2010: 130-7).

Having identified an activity level *issue*, it is a relatively straightforward procedure to generate a pair of activity level *propositions*, representing the contending assertions of the parties on this issue. The pair of propositions arising from the issue in *DNA Tights* may be formulated as:

Prosecution Proposition: *S is the intruder who tied W’s hands with the tights.*

Defence (or Alternative) Proposition: *S is not the intruder who tied W’s hands with the tights.*

In this illustration, the Defence Proposition is the negation of the Prosecution Proposition.²⁹ It is irrelevant whether the defence intends to make that argument explicitly at trial, or in due course does so or not. We are concerned with the logic of the prosecution’s argument, not with forensic strategy. The negation of the prosecution’s assertion is already logically implicit as a possibility, regardless of whether anybody actually articulates it in the case. If reference to ‘Defence Proposition’ is distracting it can always be styled the ‘Alternative Proposition’ instead. But it makes sense within an adversarial trial system, and is convenient shorthand, to call the negation of the Prosecution Proposition the ‘Defence’ Proposition, understood in this technical sense.

In this case, we can clarify the Defence Proposition by restating it in the affirmative:

Defence Proposition: *Some intruder other than S tied W’s hands with the tights.*

2.31 The probative value of the evidence can be conceptualised in terms of the extent to which it discriminates between the comparative probabilities of these two opposed events. More

²⁹ Remember that proposition pairs, for these purposes, must be mutually exclusive but need not necessarily be exhaustive.

simply put, does the evidence tend to support the prosecution proposition over the defence proposition, or vice versa; and in either event, by how much?

In our hypothetical, there is one key piece of conditioning information that refines the activity level issue: the suspect may have touched W's tights-clad legs on the evening in question. This contingency produces a modified pair of propositions, thus:

Prosecution Proposition: *S is the intruder who tied W's hands with the tights, assuming that S may have touched the tights innocently a short time before.*

Defence (or Alternative) Proposition: *Some intruder other than S tied W's hands with the tights, assuming S may have touched the tights innocently a short time before.*

Notice that, conditioned on this revised proposition-pair, evidence that DNA recovered from the tights matches S does not, without more information, discriminate between the two competing propositions. It will now be necessary to consider, in addition, factors such as concentrations of DNA and their distribution and persistence. For example, one might expect to find greater quantities of DNA, or to find DNA in different places on the tights, if S had gripped the tights firmly for several minutes to use them as improvised restraints during a robbery as opposed to merely brushing them accidentally and momentarily with his fingertips in a crowded bar.

- 2.32 Given this updated factual scenario and its evidential implications, a forensic scientist employing the CAI model might judge it insufficient simply to test a random sample of fabric taken from the tights. She might want to inspect the entire garment and undertake a series of tests following a more systematic sampling procedure. Those areas that would have been pulled tightly by the offender (whether S or somebody else) when the tights were tied around W's wrists would be of primary forensic interest. A comprehensive scientific evaluation would require consideration of the probabilities of transfer, persistence, detection, distribution, and mixing of DNA-profile components, in light of the conditional assumptions built into the respective prosecution and defence propositions.

Taking account of further evidential contingencies, such as the possibility of DNA transfer through innocent contact, may well affect the probative value of the evidence. In some

situations, this would produce smaller – possibly *dramatically* smaller - likelihood ratios. This places the forensic scientist in something of a quandary. If her findings are reported at source or sub-source level the evidence may *appear* far more probative than if the scientific findings were reported as activity level. There is no objective basis for arbitrating between these two options (unless there is further relevant evidential information which has not yet been taken into account). An even-handed way to proceed might be to report *both* sets of findings, giving quantified probabilities (where appropriate) for issues at the activity *and* source/sub-source levels in the hierarchy of issues. Proceeding in this fashion might require the scientist to make professionally-informed, but still somewhat speculative, frequency estimations, in the absence of data indicating the likelihood of relevant contingencies (e.g: Just how often is DNA innocently transferred in crowded pubs? How much is typically transferred? How long does innocently transferred material persist? And so on). The results would need careful explanation for their intended audience(s). Findings reported at activity level might still favour the prosecution's proposition, but with a smaller likelihood ratio than would be reported at source or sub-source levels.

In a case like *DNA Tights*, the prosecution might legitimately insist that S's claim to have touched W's tights is a false concoction to try to account for otherwise very incriminating DNA profiling evidence. Perhaps the prosecution thinks it can, after all, identify the real pub leg-brusher as somebody other than S; or maybe the circumstances in which S first raised the possibility of innocent contact with W's tights already tend to undercut the credibility of that claim. So the prosecution may wish to present the factfinder with a random match probability or likelihood ratio for the original prosecution proposition at sub-source level, unconditioned by the possibility of innocent contact. But the forensic scientist has no basis, in logic, science or law, for dismissing the suspect's contentions out of hand. The scientist might well judge, in this kind of (forensically unexceptional) scenario, that a more balanced and realistic expression of the probative value of the evidence would be framed at activity level, taking account of all pertinent conditioning assumptions.

2.33 This is a concrete illustration of an enormously important general proposition about the probative value of scientific evidence:

The evidential significance of a likelihood ratio always depends on the level of issue (within the hierarchy of issues) to which a pair of competing propositions is addressed.

To the extent that forensic science evidence adduced by the prosecution is routinely framed in terms of *sub-source* or *source* level issues, and in the absence of balancing information at *activity* level, its probative value could easily be overestimated by fact-finders (Evetts et al 2002; Jackson 2013). Increasing commercialisation of forensic science provision in the UK is anticipated to exacerbate this problem, since reports of scientific findings pitched at the source or sub-source levels are easier and cheaper to produce than properly contextualised scientific findings addressing activity level issues (also necessarily incorporating source or sub-source assessments).

2.34 ***Combining Evidence through the Hierarchy of Issues***

Further complexities are introduced into evidential evaluation when different pieces of information must be combined together, taking proper account of their mutual conditional dependencies. A forensic scientist following the CAI approach might be able to provide material assistance to criminal justice professionals in evaluating the evidential significance of combinations of multiple evidential items – provided that ‘consumers’ are willing and able to pay the market rate.

2.35 The following detailed hypothetical illustrates how a forensic scientist, adopting the CAI method, could best assist investigators and prosecutors in a case involving combinations of multiple pieces of evidence.

Head-stamping – case facts: V was the victim of a vicious attack outside a nightclub. Several bystander-witnesses report seeing the perpetrator repeatedly stamping on V’s head, whilst V was lying unconscious on the pavement outside the club. V sustained numerous cuts that bled heavily. Medical staff observed marks on V’s face, in the area of the wounds, which resembled the sole pattern of footwear. The marks were photographed, with a view to subsequent forensic comparison. A suspect, S, was apprehended as a result of an anonymous tip-off to the police. S denied being anywhere near the club at the time of the incident, but failed to offer any explanation for the appearance of what appeared to be fresh blood on his shoes.

On the face of it, this scenario presents two sources of evidence potentially incriminating, or possibly exonerating, S. First, a DNA profile might be generated from the blood on S's shoes and assessed for a match with V. Secondly, the marks on V's face might be compared with the sole pattern on S's bloody shoes. Both types of evidence, of course, reflect the hypothesis that S got V's blood on his shoes when he stamped on her head. More specifically, the two items of evidence address particularised source level issues. In relation to the blood evidence: 'Did the blood on S's shoes come from V?'. In relation to the facial marks: 'Was the footwear mark on V's face made by S's shoe?'. The proposition-pairs of interest at the two source levels would be:

(i) Blood evidence

Prosecution Proposition (PP): The blood on the shoe came from V.

Defence Proposition (DP): The blood on the shoe came from some unknown person (unrelated to V).

(ii) Facial marks

Prosecution Proposition (PP): The footwear mark on V's face was made by S's shoe.

Defence Proposition (DP): The footwear mark was made by some other (unknown) shoe.

These two pieces of evidence, taken together, might appear to invite a relatively straightforward progression to consider, at a activity level, the identity of the head-stamper in this case.

Head-stamping – forensic examinations: A forensic biologist examined S's shoes, observed blood-staining, and noted its distribution and pattern. A partial DNA-profile matching V was obtained from this blood, with a random match probability of approximately 1 in 1000. On evaluation, this gave a likelihood ratio for the blood evidence at source level of approximately 1000.³⁰ A footwear specialist then examined scaled photographs of the marks on V's face and observed a pattern that

³⁰ Because $p(\text{Evidence} \mid \text{V is blood donor}) / p(\text{Evidence} \mid \text{V is not blood donor}) = 1 / \{1/1000\} = 1 \times 1000 = 1000$. The logic of these – fairly elementary – calculations was explained in *Practitioner Guide* No 1.

corresponded with the sole of S's shoe. The pattern had a relative frequency of occurrence in a relevant footwear database of approximately 1 in 50, giving a likelihood ratio of approximately 50 for a *source* level issue in relation to the footwear pattern.

Thus, the prosecution has two separate pieces of scientific evidence, each with its own measure of probative value expressed as a likelihood ratio. How can these two assessments legitimately be combined together to produce a composite evaluation of probative value?

2.36 One plausible approach would be to leave the argument on the combined value to lawyers, who would presumably choose whichever interpretation of the evidence best supported their own case. After all, lawyers are responsible for the conduct of litigation and the forensic scientist is not in a position to dictate trial strategy. However, this passive approach would appear to leave much to chance. It assumes that both sets of lawyers will be competent to reinterpret the scientific findings in a way that best supports their respective contentions, and that this adversarial dispute can be presented to the jury in a fashion that enables jurors to arbitrate between contested claims on a rational evidential basis.

Alternatively, the forensic scientist might adopt a more proactive approach, taking greater professional responsibility for communicating scientific findings in a form calculated to make them more intelligible, and useful, to lawyers, courts and fact-finders. In *Head-stamping*, there is another, potentially significant piece of evidence that a forensic scientist might consider, namely, *the pattern of bloodstaining observed on the shoes*. A practical strategy for extracting the maximum probative value out of *all* the available information, taking proper account both of each individual item of evidence *and* their cumulative evidential significance, would be for the scientist to offer an evaluation at *activity* level, in the form of a composite likelihood ratio combing all three discrete pieces of evidence in a logical, coherent way.

2.37 In *Head-stamping*, for example, the *activity* level issue might be encapsulated in the following question:

Is this the shoe that was worn during the repeated stamping of V's head?

The proposition pair at *activity* level flowing from this issue could then be formulated as:

Prosecution Proposition (PP): *The shoe from the suspect is the shoe that was worn during the repeated stamping of V's head.*

Defence Proposition (DP): *Some other, unknown shoe was worn during the repeated stamping of V's head.*

An appropriately competent scientist could now provide an evaluation of the composite likelihood ratio for the combination of all three evidential items: (1) the matching DNA-profiles; (2) the footwear mark matching S's shoe; and (3) the pattern of bloodstaining .

2.38 At *activity* level, the observed pattern of bloodstaining is the additional piece of information that needs to be taken into account, in addition to the matching DNA profiles and footwear marks, in the evaluation of an LR. In this particular case, the scientist needs to ask herself two questions:

1. What is the probability that I would observe this pattern of bloodstaining if this were the shoe worn during the stamping?
2. What is the probability that I would observe this pattern of bloodstaining if some other shoe was worn during the stamping?

In addressing the first question, the scientist may consider, based on her experience of similar cases, that there is a high probability that this pattern would be observed if this were the shoe worn during the stamping. To answer the second question the scientist would need to draw on different experience, knowledge or data, in an attempt to assess whether the pattern of bloodstaining observed on S's shoe was common or rare within some relevant population of footwear (e.g. training shoes worn by young males). If the scientist had access to a reliable and appropriately representative forensic database or extensive previous casework experience, she might be in a position to conclude that bloodstaining in this pattern is very rarely seen on shoes in the relevant suspect population. Therefore, she could assign a very low probability in answering question two – let's say 1/500 or 0.002.

The ratio of the two probabilities given in answer to questions 1 and 2 is the likelihood ratio provided by the observed pattern of bloodstaining (in other words, a measure of the probative value of the bloodstaining pattern evidence). The scientist could now calculate an LR for the bloodstain pattern itself. Even if the probability of observing the pattern if the shoe made the mark is less than 1 (certainty) – say, 0.75 – the likelihood ratio for the bloodstain pattern evidence will increase the probative force (as measured by an LR) of the observations over and above the probative value of a matching DNA profile at *source* level. Using our hypothetical figures, the LR for the blood stain pattern would be $0.75/.002 = 375$. On the plausible assumption that the pattern of blood staining on S's shoe is completely independent of the DNA profiling evidence relating to V,³¹ the bloodstain pattern evidence (LR = 375) can be combined with the DNA profiling evidence (LR = 1000) to produce a combined LR for the 'blood evidence' of 375,000. Furthermore, plausibly assuming independence again, the combined value of the blood evidence (LR = 375,000) can be multiplied by the probative value of the footwear mark evidence (LR = 50) to produce a composite, three-factor LR of approximately 18.8 million. This is a significantly higher LR than could be achieved by simply combining the two pieces of evidence at source level, i.e. the matching DNA profile (LR = 1000) and the footwear mark evidence (LR = 50), producing a combined LR of 50,000. The additional probative value that can be wrung out of this evidence is attributable to the forensic scientist's knowledge and expertise in addressing *activity* level issues (in this case, expressly considering whether S's shoe was the shoe worn during the activity of stamping on V).

- 2.39 This general reasoning procedure for thinking logically about the probative significance of scientific findings is flexible and adaptable to further developments in the investigation or proceedings. For example, relevant new information may subsequently be supplied to the forensic scientist. Suppose that the suspect in our hypothetical stamping case changes his account following disclosure of the initial scientific findings. He now concedes that he may have accidentally kicked the victim's head once, as he stumbled over the body when running to flag down a taxicab. This new scenario is readily accommodated by rephrasing the alternative (defence) proposition at activity level to read:

³¹ So that the product rule for calculating the combined probability of independent events can be applied, as explained in *Practitioner Guide* No 1.

DP₂: *This shoe was worn by the suspect whilst accidentally kicking V's head once.*

The scientist would then consider the probability of obtaining the observed pattern of bloodstaining, the matching DNA-profile and the matching footwear pattern if this alternative proposition were true, and compare the probability under DP₂ with the probability of the same observations (E) if the prosecution proposition were true. It may well be that the scientist would still consider E more likely under PP than under DP₂, i.e. more likely if the prosecution proposition were true rather than if the (modified) defence proposition were true. The scientific evidence would still, on that basis, provide support for the prosecution proposition, albeit the magnitude of the LR may be much reduced under DP₂ – a result that would fit well with common-sense intuitions about the evidence. As a common sense matter, it seems more plausible that S would have bloodstaining on his shoe if – as he claims – his foot had accidentally kicked V's head than if he was claiming never to have come into contact with V at all (in that case, how did the blood get on his shoe?). But a fact-finder might well consider that the *more* plausible explanation of *a particular pattern* of bloodstaining, suggesting repeated deliberate contact, remains the initial prosecution hypothesis, that S had repeatedly stamped on V's head. In the absence of any persuasive evidence to the contrary, the fact-finder might well draw that inference.

2.40 Although this is a rather simplistic illustration compared to the multiple contingencies and complexities of real-life case work, it serves to demonstrate a profound conceptual truth with direct implications for criminal litigation. The probative value of scientific findings depends on the propositions that they are taken to be addressing. If scientists were always to restrict their interpretations to *source* level issues and propositions they would effectively be trusting other criminal justice professionals, or fact-finders themselves, to contextualise the scientific findings and interpret them correctly. But as our head stamping case shows, the probative value of scientific findings addressed to source level propositions may *appear* considerably weaker (or in a different type of case, exemplified by *DNA Tights*, stronger) than if the same findings are presented in terms of activity level propositions. It must at least be seriously open to question whether lawyers and courts are currently sufficiently well-informed about the relational nature of scientific evidence or calculations of LRs to perceive these evidential subtleties, and fully to appreciate their forensic significance, without expert assistance.

The core idea that the meaning and probative value of scientific findings are relative to propositions that can be expressed at different levels is of immense importance for the fair and effective administration of criminal justice. Its implications need to be understood by everybody professionally concerned with the production, interpretation and assessment of scientific evidence. It should be stressed that *all* scientific evidence – not merely that which is already expressed in probabilistic terms, such as the random match probabilities attached to DNA profiling evidence – share this relational character and can profitably be evaluated using a likelihood ratio approach. Refining and extending the CAI model, evaluation at *activity* level enables a coherent assessment of the combined probative force of separate pieces of scientific evidence which would generally be presented separately at *source* or *sub-source* levels. The effect may be to reduce (or increase) the combined probative value of the evidence, or even to invert its polarity, so that evidence that may appear to support the prosecution’s case at source level may be reinterpreted as supporting the defence case at activity level.

2.41 ***Activity Level Propositions in R v Barry George***

Our final illustration in this Part of activity level propositions, and their significance, is taken from a well-known case in which a conviction partly based on scientific evidence was, eventually, quashed on appeal.

Barry George was convicted in June 2001 of murdering Jill Dando, a well-known TV personality, who was shot to death on 26 April 1999. The prosecution’s case at trial comprised three main strands, one of which was the presence of a single particle of firearms discharge residue (FDR) recovered, almost one year after the murder, from the pocket of one of George’s coats.³² A first appeal against conviction, raising various points including the possibility that the FDR evidence might have been compromised by contamination, was dismissed.³³ The Criminal Cases Review Commission (CCRC) subsequently referred

³² As recounted in the report of the second (successful) appeal: *R v Barry George* [2007] EWCA Crim 2722, [3].

³³ *R v George (Barry Michael)* [2002] EWCA Crim 1923. The jury heard from the defence scientist that ‘the particle was so small that to rely on it, one year after the killing, was “incredible”. Its size “cast doubts on where it came from”— it could be the result of casual contamination’: [95]. But the

George's conviction back to the Court of Appeal, having commissioned a further report on the FDR evidence which questioned its probative value. Quashing George's conviction, the Court of Appeal concluded that:

The summing up of the judge showed that he understood that... the finding of the single particle of FDR provided cogent support for the other evidence relied upon by the prosecution as establishing that the appellant was the murderer of Miss Dando. Early in his summing-up he described the expert evidence in relation to the FDR as being an 'important part' of the prosecution case and later as one of the 'three main strands' of that case.... [W]hen considered objectively that evidence conveyed the impression that the Crown's scientists considered that innocent contamination was unlikely and that, effectively in consequence, it was likely that the source of the single particle was the gun that killed Miss Dando. In that respect their evidence at the trial was in marked conflict with the evidence that they have given to this court with the result that the jury did not have the benefit of a direction that the possibility that the FDR had come from the gun that killed Miss Dando was equally as remote as all other possibilities and thus, on its own, entirely inconclusive. In the light of the way in which Mr Keeley now puts the matter, we have no doubt that the jury were misled upon this issue.... It is impossible to know what weight, if any, the jury attached to the FDR evidence. It is equally impossible to know what verdict they would have reached had they been told as we were told, by the witnesses who gave evidence before us, that it was just as likely that the single particle of FDR came from some extraneous source as it was that it came from a gun fired by the appellant. The verdict is unsafe. The conviction will be quashed.³⁴

2.42 The FDR evidence adduced at trial was presented (in our terminology) predominantly in the form of source level explanations; that is to say, the principal question was whether the FDR particle recovered from George's clothing originated from the gun used to murder Jill Dando. The jury was told that discovering only one FDR particle a year after the shooting had taken place was 'not an unusual situation', that residue 'would not be found on ordinary members of the public unless they had been associated with firearms' and that 'the micron particle was consistent with having come from the cartridge used in the killing'.³⁵ There is no mention in the judgment of the FDR findings being interpreted or presented to the jury in terms of

Court of Appeal concluded that evaluation of the FDR evidence 'was entirely for the jury having been properly directed': [106].

³⁴ *R v George* [2007] EWCA Crim 2722, [46], [51], [54].

³⁵ As summarised from the trial transcript by the Court of Appeal in *R v George (Barry Michael)* [2002] EWCA Crim 1923, [94].

likelihood ratios during the trial proceedings. However, it appears that the FSS firearms expert called by the prosecution did say at one point during his examination-in-chief that ‘What [the FDR particle] may not tell us is exactly how it got there’.³⁶

Interpretation, communication and evaluation of the FDR evidence became the central issue for the second appeal. Addressing the nature of the fresh evidence not considered in the first appeal proceedings, the Court of Appeal remarked that ‘[t]he most significant evidence that we have received owes its origin to the initiative of Dr Ian Evett, who from 1996 to 2002 worked for the Forensic Science Service’.³⁷ The Court then proceeded to summarise the genesis of CAI as a method ‘assisting scientists to present evidence to the court in a manner that was logical, transparent and robust’:

Part of the technique... involved analysing the evidence by reference to the extent to which it supported one or other of two rival propositions. Typically one proposition would be that an event that formed part of the prosecution case occurred and the rival proposition would be that this event did not occur.... The object of this is to clarify *before* evidence is examined and analysed the likelihood of the examination achieving particular results on two different hypotheses or propositions. This technique facilitates the drawing of appropriate conclusions from the results actually obtained on the examination.³⁸

It will be observed that the Court of Appeal is here describing a likelihood ratio approach at activity level (the Court refers to an ‘event’ rather than an inferential relationship at source level), and with apparent acceptance if not positive endorsement.

2.43 Dr Evett, the Court of Appeal recounted, had met with the FSS firearms expert who had testified at George’s trial to reconsider the nature of the FDR evidence utilising the CAI method. They had considered the following pair of propositions (or *hypotheses*, hence ‘H’), formulated at *activity* level, in exploring how to answer the question: who shot Jill Dando?

H_p - *the appellant was the man who shot Jill Dando* (the prosecution proposition)

H_d - *the appellant was not the man who shot Jill Dando* (the defence proposition)

³⁶ *R v George* [2007] EWCA Crim 2722, [35] (quoting directly from the trial transcript).

³⁷ *ibid* [14].

³⁸ *ibid* [14], [16] (original emphasis).

Drawing on expertise gained through previous casework experience involving firearms residue and taking account of relevant surveys and studies, the FDR expert judged that finding a single FDR particle would be very unlikely under *either* hypothesis. He estimated the likelihood of the FDR findings conditioned on either scenario as 1 in 100, and therefore concluded that the evidence was ‘neutral’ with respect to the activity level issue: who shot Jill Dando?

The meaning of this evaluation is that the FDR evidence is logically incapable of discriminating between the two rival hypotheses H_p and H_d . In legal terms, the evidence would be irrelevant and inadmissible if its only purported purpose were to shed light on this activity level issue.³⁹ In likelihood ratio terms, a numerator of 1/100 and a denominator of 1/100 produces an LR of 1, meaning that the FDR evidence would have no logical bearing on any prior of assessment of the other evidence in the case before the FDR evidence was taken into account (because any value multiplied by 1 remains unchanged). In the pithy conclusion of the CCRC-commissioned expert re-evaluation, ‘[t]he FDR evidence is thus inconclusive. In our opinion it provides no assistance to anyone asked to judge which proposition is true’.⁴⁰

2.44 The Court of Appeal quashed George’s conviction because it was felt that the way in which the FDR evidence had been presented at the original trial may have misled the jury in assessing its meaning and probative value.

In reality, when considered objectively that evidence conveyed the impression that the Crown’s scientists considered that innocent contamination was unlikely and that, effectively in consequence, it was likely that the source of the single particle was the gun that killed Miss Dando. In that respect their evidence at the trial was in marked conflict with the evidence that they have given to this court with the result that the jury did not have the benefit of a direction that the possibility that the FDR had come from the gun that killed Miss Dando was equally as remote as all other possibilities

³⁹ This is *not* to say that the FDR evidence of a similar type would be irrelevant in every case. It could be relevant, for example, if a defendant claimed to have never fired a gun (whereas Barry George was a self-confessed gun enthusiast).

⁴⁰ Quoted by the Court of Appeal in *R v George* [2007] EWCA Crim 2722, [23].

and thus, on its own, entirely inconclusive. In the light of the way in which Mr Keeley now puts the matter, we have no doubt that the jury were misled upon this issue.⁴¹

It was now accepted that the jury should have been told, in so many words, that the finding of a single particle of FDR was completely inconclusive on the issue of the identity of the murderer.⁴²

The Court of Appeal expressly determined the appeal on the basis of the manner in which the FDR evidence was presented to the jury, without reconsidering the preliminary question of its admissibility.⁴³ The Court noted, in passing, the contextual nature of judgements of relevance, and perhaps thought it would be unwise or unhelpful to attempt to make categorical statements about matters which must inevitably be determined at first instance in the light of case-specific considerations. The appeal judgment does, however, rather beg the question as to what the relevance of the FDR evidence could have been in the context of Barry George's trial and what was already otherwise known to the jury about him. When the case was retried, the prosecution again attempted to adduce the FDR evidence, but this time the trial judge ruled it inadmissible; and the jury in due course returned a verdict of acquittal.

2.45 Barry George's case offers a striking real-world illustration of the value of a methodical CAI approach in forming balanced interpretations of forensic science evidence calculated to provide maximum assistance to the courts. Scientific findings may have significantly different implications for source level and activity level issues. In this case, the real possibility that the jury may have failed to notice this distinction, or possibly conflated the two levels of analysis in treating the FDR evidence as probative of the killer's identity, produced a wrongful conviction of murder that – at the second time of asking – was quashed on appeal.

⁴¹ *R v George* [2007] EWCA Crim 2722, [51]. Is the Court of Appeal guilty of transposing the conditional in this paragraph? That all depends on whether the 'and that, effectively in consequence' clause is interpreted as summarising the forensic scientist's evidence, or as expounding the jury's ultimate factual inference.

⁴² *ibid* [26] – [27].

⁴³ *ibid* [31].

R v George also implicitly teaches an important lesson about the value of CAI for initial case assessment and the development of a sensible forensic strategy. It was only after the fact, when the expert was invited to consider the *ex ante* likelihood of locating FDR particles, that the small probability of finding such evidence on either prosecution or defence hypotheses was consciously articulated. Had the expert pre-assessed an examination of the coat for FDR particles, he would have predicted that the most likely outcome of the examination was no particles, and that this finding would have zero evidential significance (LR = 1). Whether the police or prosecution, given this assessment, would still have requested the examination is unknown. Perhaps the police would still have requested the examination for purely *investigative* purposes, in which case the expert would have produced an *investigative* opinion, listing all feasible *explanations* for finding a single particle. This is exactly the approach subsequently endorsed by the FSS in guidance to reporting scientists.⁴⁴

In either event, CAI demonstrates its value in clarifying the purpose of forensic inquiries, structuring the format and wording of forensic reports, and informing intelligent interpretations and uses of their contents.

⁴⁴ ‘There is not sufficient data on the environmental occurrence of FDR to give a safe interpretation of finding a single particle of residue. Consequently the FSS has adopted a cautious approach to reporting LOW levels of residue and no evidential value can be offered. From an investigative point of view LOW levels of residue may nonetheless have some value; for example, finding a LOW level on a discarded item such as a glove may give a significant lead to a police investigation. When an officer is given information on LOW levels in an investigative submission he must be made aware that in most cases it is unlikely any evidential weight can be attached to the finding’: *FSS guidelines on the assessment, interpretation and reporting of firearms chemistry cases*, para. 9.5, quoted in *R v George* [2007] EWCA Crim 2722, [21].

3. Case Assessment and Interpretation Protocol

3.1 *Seven Stages of Case Assessment and Interpretation*

The model for Case Assessment and Interpretation (CAI) can be presented as a series of stages in the processing of a case, from the initial enquiry through to delivery of a report to the criminal justice client (Cook et al 1998). Each stage has its own set of prompts to structure and guide the forensic scientist's decision-making and encourage a systematic approach to casework. The model has been developed over time in the light of experience, and there are now several variants in operation in the UK. Some forensic science providers have developed their own bespoke CAI *proforma* for inclusion on individual case-files. The *proforma* both serves to guide case-reporting officers' decision-making and simultaneously constitutes an auditable record of the rationale underpinning a forensic scientist's examination strategy. CAI has been adopted as the basis for standards for providing expert opinions published by the Association of Forensic Science Providers (2009). In litigation, it supports the objectives of the Criminal Procedure Rules in encouraging early identification of the issues and agreement of non-contentious expert evidence, insofar as this is practically feasible.

The CAI model comprises the following seven principal stages (and see Jackson and Jones 2009):

- (1) Acquire information regarding all relevant circumstances of the case
- (2) Clarify and define the client's requirements
- (3) Develop a proposal for a cost-effective examination strategy
- (4) Consult client, explain potential outcomes, costs and timescales, and agree an examination strategy
- (5) Conduct the examination(s)
- (6) Interpret the scientific findings
- (7) Communicate findings and explain their interpretation to the client.

This is a stylised model. In reality, the process may be iterative rather than linear. For example, it may be necessary for forensic scientists to retrace their steps, and sometimes to revise their initial examination strategies, if new information comes to light or if unexpected

findings are obtained. Conversely, some of the stages can be omitted when undertaking relatively simple routine examinations (e.g. a straightforward service in the chemical identification and quantification of drugs). The CAI model is a flexible tool that can easily be adapted to accommodate a range of scenarios. It articulates, in a comprehensible, transparent way, the logical steps that an examiner should follow in conducting forensic examinations in every case. In truth, the CAI model is really just a formalised distillation of practical wisdom. It reflects what the best forensic examiners have been doing, unprompted and unselfconsciously, for many years.

We will now examine each of the model's seven stages more closely. Appendix A to this *Guide* reproduces an aide-memoire (adapted from Jackson and Jones 2009), summarising all seven stages, to assist forensic practitioners in applying the CAI approach to their own casework.

3.2 *Stage 1: Acquire information regarding all relevant circumstances of the case*

In order for the forensic provider to understand the client's requirements and assess the demands of the investigation, it is necessary for forensic scientists to know and consider all *relevant* parts of the case circumstances. Concern has sometimes been expressed about the potential for 'extraneous information' to compromise the objectivity of scientific examinations and findings, for example through the operation of – largely unconscious – expectation effects. However, it is important to be clear about precisely what the forensic scientist needs to know under the CAI model. Not *all* case information (let alone broader police intelligence and investigators' gossip) is *scientifically relevant* data of legitimate concern to forensic scientists.

Specifically, the forensic scientist needs to be appraised of all and only those pieces of information bearing on:

- the issues to be addressed by the scientist;
- choices of forensic strategy (e.g. which items to examine or which techniques to utilise); and
- the probability of obtaining particular results or findings.

Such information would generally include *relevant* parts of witness statements, chronological information about the timings of relevant events or actions, the proposed or actual charges in the case, and any pertinent disclosures or contentions by a suspect or defendant; as well as procedural target dates.

- 3.3 By contrast, examples of information that forensic scientists generally do *not* need to know (always ultimately depending on particular case circumstances) include matters pertaining to a suspect's or the accused's character (including previous criminal convictions), domestic circumstances, movements prior to the time of the offence, or formal or informal witness identifications. Instructions to experts along the lines of 'We have reliable eyewitness identification placing this suspect at the scene. Please confirm that this is his DNA/hair/fingerprint etc' gratuitously imperils the actual or perceived objectivity of the scientist's examinations and potentially undermines the probative value of the evidence.

Whatever information is relayed to the scientist, there should be procedures in place to minimise the risk of contextual biasing of an expert's expectations and subsequent opinions (Forensic Science Regulator, 2014). Protocols for 'sequential unmasking' of case information might appropriately be adopted for certain kinds of forensic examination. Accurate records of the information provided to forensic scientists should be maintained by all parties, and would generally be disclosable to the defence.

3.4 *Stage 2: Clarify and define the client's requirements*

Even under a 'market' provision model of forensic science services, the role of forensic practitioners is not reduced to merely passively fulfilling consumer demands. This second stage of the CAI model directs forensic providers to interpret clients' requests in the light of an informed appraisal of what particular clients might actually need. Whilst clients may have a general idea and entertain certain expectations of how forensic providers might assist their case, once the client's needs have been clarified the provider may be able to suggest a superior, more cost-effective approach. It may be that different issues should be addressed or an alternative case examination strategy adopted. For example, a police officer investigating a burglary might be interested in commissioning forensic analysis of glass fragments from the window broken by the burglar to gain access to the premises. However, if the case circumstances were to be explored more thoroughly, it might be evident to an experienced forensic scientist that the probability of recovering any significant matching glass from any

suspect (even if he *was* the offender) would be very low in any event. In these circumstances, glass analysis would largely be a waste of time and money. The police officer would be well advised that examination strategies with better prospects of bearing significant evidential fruit, such as searching for fibres or footwear marks, should be pursued instead.

3.5 In practical terms, the forensic provider would review the information received at stage (1) in order to address the following questions:

- What are the real uncertainties (the issues in the case) with which the client requires assistance?
- Are these uncertainties properly characterised as posing *investigative* or *evaluative* questions?
- To what level in the hierarchy of issues does each question relate?
- What type of scientific conclusion or expert opinion would be appropriate in relation to each issue? (What kind of answer would be appropriate for each investigative or evaluative question?)

Investigative issues or questions typically concern the circumstances of a crime or the type of offender who might have committed it. These questions may be addressed whether or not the police have yet identified or apprehended a particular suspect. For example, the scientist may be able to advise on the number of attack sites at a crime scene, the type of footwear worn by the assailant, the nature of the explosive device planted on the bus, the cause and seat of a house fire, and so forth.. For such investigative issues, the appropriate type of expert opinion (as explained in Part 2) would be in the form of ***explanations*** for the evidence; or where explicit quantification of uncertainty is possible and desirable, in the form of ***posterior probabilities*** for the explanations.

Evaluative issues or questions, by contrast, tend to centre on specific suspects or defendants (though they can also relate to complainants or other issues). Examples include: whether the defendant was the person who wielded the hammer; whether the blood at the scene came from the defendant; whether the image of a left hand in the paedophile video portrays the defendant's hand; whether it was the defendant who poured petrol through the letter box. For these evaluative issues, the scientist's opinion would be in the form of a ***likelihood ratio*** for

the evidence conditioned on mutually exclusive (but not necessarily exhaustive) prosecution and defence propositions..

3.6 *Stage 3: Develop a proposal for a cost-effective examination strategy*

At the third stage of the CAI model the forensic provider must consider the range of potential results that might be generated by the various techniques at his or her disposal.

When operating in *investigative* mode, the scientist would endeavour to assess how particular techniques might produce useful information capable of reducing uncertainty about material events and helping to direct the investigation.

In *evaluative* mode, the aim would be to consider, for each technique, the probability of obtaining particularised observations and analytical results, predicated on the truth of competing adversarial propositions. A set of calculated LR_s would then underpin informed assessment of the potential probative value of the evidence that might be obtained by adopting alternative forensic strategies.

3.7 This is a particularly important stage in case assessment when the scientist is operating in *evaluative* mode. *A priori* consideration of likelihood ratios, *before* any scientific tests are conducted, helps to guard against *post hoc* rationalisation whenever the scientist comes to interpret whatever observations or analytical result are in due course obtained.

Without the precaution of a systematic logical approach, a scientist who observes the absence of matching glass fragments may be tempted to rationalise: 'I didn't find any glass on his clothing, so maybe the suspect was standing further from the window when he broke it'. This may be acceptable as an attempt by the scientist to provide an explanation for the lack of glass, but it could also be a manifestation of pro-prosecution bias. A scientist who performs an effective case assessment, and remains faithful to its logic whatever results are subsequently obtained, would be prompted to reconsider the logic of the inquiry. A balanced interpretation of the absence of glass may be that it supports the suspect's story (that he didn't have anything to do with the incident), rather than merely qualifying the prosecution's allegation that the suspect broke the window. The scientist is by no means precluded from revisiting her case assessment in the light of further findings. CAI is an iterative process; and reconsideration of the probabilities originally assigned during case assessment may be

appropriate, as where the scientist discovers, on subsequent examination, that the glass is a very rare type rather than a common or garden variety as had been assumed in the initial assessment.

3.8 *Stage 4: Consult client, explain potential outcomes, costs and timescales, and agree an examination strategy*

If there are choices to be made on which items to examine or which techniques to employ, this is the stage at which the client must be contacted. The provider should advise the client of the potential value and forensic benefits of alternative examination strategies, as well as of any limitations, respective costs and timescales. Having heard the provider's advice and any recommendations, the final decision on the strategy to be adopted is the client's..

3.9 Details of the case information supplied to the scientist, the rationale behind the forensic strategy ultimately adopted, and all discussions, agreements and directions received from clients should be recorded on the scientist's case-file. This information is generally disclosable in criminal proceedings, and would therefore be available to be scrutinised and possibly challenged if the case proceeds to trial..

3.10 *Stage 5: Conduct the examination(s)*

This is the stage corresponding to the functions traditionally performed by forensic scientists: evidence bags are opened, observations made, material is recovered, analysed and compared; full records are kept.

Many important issues arise at this stage, including preserving chain of custody, adequacy of anti-contamination procedures, quality assurance of analytical results, proper supervision of forensic practitioners, and verification of their interpretative findings, all of which could potentially have a major bearing on the magnitude of any reported likelihood ratio and, by extension, on the quality and reliability of forensic science evidence. However, since these vital practical issues are not integral to logical frameworks for assessing or interpreting scientific findings, they are beyond the scope of this *Guide* and we say no more about them.

3.11 *Stage 6: Interpret the scientific findings*

If the provider is addressing an *investigative* issue, interpretation generally takes the form of generating a list of possible *explanations* for the findings. It is imperative that the provider

makes clear to the client that any set of explanations offered may not necessarily be exhaustive: there could conceivably be other explanations for the findings, which the scientist has either deliberately discounted (e.g. failures in the chain of custody leading to compromised samples) or has simply not considered. The scientist's task is to compile a list of all pertinent and feasible *explanations* for the findings in the instant case. The scientist may then be justified in going beyond reporting just the results themselves, to offer *posterior probabilities* for particular explanations.

In doing so, however, the scientist must first assign realistic and fair *prior probabilities* for the explanations. These 'priors' should be fair in the sense of taking appropriate account of all relevant known information about the case, including information or anticipated arguments that might be advanced by the defence. *Posterior probabilities* can then be generated using the probabilities for the findings, in accordance with Bayes' rule for updating conditional probabilities in the light of new information; with the important proviso (which should be made explicit) that the analysis is confined to the explanations originally considered, which are being treated as mutually exclusive and exhaustive for these purposes. Since a range of possible explanations is being considered, the scientist's conclusions could be reported as a list of explanations for the findings, running from the most to the least probable. In whatever form the results are presented, however, it is vital that the client understands the meaning of the list and the thinking behind it, and is able to question and seek clarification on any part of it.

- 3.12 If the scientific provider is operating in *evaluative* mode, interpretation takes the form of assessing the values of the *likelihood ratios* for the observations and analytical results actually obtained. This assessment would be carried out within the broad analytical framework devised at the start of the CAI process, now refined to take account of the specific observations and results obtained.

On occasions, forensic scientists operating in either *investigative* or *evaluative* mode may be in a position to offer *categorical* exclusions or classifications in relation to some of the issues. But categorical opinions should be advanced with circumspection, for the reasons we have seen.

3.13 *Stage 7: Communicate findings and explain their interpretation to the client*

Scientific findings, and their interpretation, should be conveyed to clients in the ways best suited to the type of evidence in question, clients' particular requirements, and the legal and evidential demands of the case. Part 33 of the Criminal Procedure Rules in England and Wales specifies generic requirements of clarity and transparency for forensic reports. The courts have sometimes specified further requirements in relation to particular types of expert evidence, such as DNA profiles,⁴⁵ footwear marks⁴⁶ and facial image comparison.⁴⁷ Beyond these fairly rudimentary requirements, there is little judicial guidance on the appropriate 'general form' of forensic reports. In particular, there is no consensus as to whether expert opinions should be given in the form of *explanations*, *probabilities* for the findings or *posterior probabilities* for the facts in issue. Nor is there authoritative guidance or consistent practice on the important question whether experts should try to assist the court by addressing *activity* level issues or questions, in addition to providing evidence going to source or sub-source issues.

There remains much work to be done on developing the most effective ways to formulate and communicate expert opinions in a format that is understandable and acceptable to police investigators, lawyers and lay people, in conformity with legal requirements and the demands of scientific objectivity, impartiality, transparency and logical reasoning. The CAI model for forensic investigation easily translates into a method for structuring forensic reports (Evetts et al 2000). This reporting protocol has been widely adopted by the UK's principal forensic providers, and has received favourable attention overseas (e.g. National Research Council 2009: 6.3).

3.14 *Case Assessment and Interpretation in Report-writing*

The seven stages of the CAI model encapsulate a protocol for logical thinking and case analysis. This approach to forensic inquiry provides the basis for writing informative, transparent and balanced reports to help readers understand *what* has been done by the forensic scientists, *why* it was done, and *what* the results mean.

⁴⁵ *R v Doheny and Adams* [1997] 1 Cr App R 369, CA; *R v Reed and Reed*; *R v Garmson* [2010] 1 Cr App R 23; [2009] EWCA Crim 2698, [53]-[55].

⁴⁶ *R v T* [2011] 1 Cr App R 9, [2010] EWCA Crim 2439.

⁴⁷ *R v Atkins* [2010] 1 Cr App R 8, [2009] EWCA Crim 1876.

Forensic reports can be structured around a set of generic headings. The following paragraphs summarise key headings that might – and ideally, *should* – be included, mapping onto the sequential stages of the CAI process.

3.15 A. ‘Case circumstances’ (or ‘Background information’)

The first section of a forensic report should clearly set out the ‘case circumstances’ or relevant background information known to the scientist prior to conducting further tests or making observations. In general terms, the scientist’s central interpretational task is to assess a likelihood ratio (LR) for forensic observations and analytical results (E). It is helpful to distinguish two types of information on which such assessments are based.

The first type of information is that which enables the forensic scientist to identify the issues in the case and specify material facts to which forensic examinations will be addressed. Typically, material facts are stated or implied by the actual or contemplated charges in the case (as communicated to the forensic scientist by police or prosecutors), also taking account of any relevant assertions or counterclaims advanced, at this stage, by the defence.

The second type of information is that which influences (or more technically, ‘conditions’) the expert’s assignments of probability (Pr) for scientific findings (or ‘the evidence’, E). This would routinely include, for example, the timings of the recovery of physical items and the extent and nature of alleged contact between those items and the offender (who may or may not be the suspect).

Information of either type on which the forensic scientist has relied should be adequately summarised in the first section of a forensic report.

3.16 It is convenient to employ – fairly simple – formal notation when working with likelihood ratios. Several variants are in common use, and have already been employed in this *Guide* (e.g. PP and DP, for ‘prosecution proposition’ and ‘defence proposition’, respectively). *Practitioner Guide No 1* provided comprehensive definitions and explanations that will be more familiar to statisticians. Briefly recapitulating, the rival contentions or ‘hypotheses’

advanced by the prosecution and defence can also⁴⁸ be represented symbolically as H_P and H_D (respectively, the ‘prosecution’s hypothesis’ and the contrary ‘defence hypothesis’); and all information of the second type, on which the pair of hypotheses is conditioned, can be represented by the letter ‘I’ (meaning ‘conditioning information’).

The likelihood ratio can then be expressed symbolically as:

$$LR = \frac{\Pr [E | H_P, I]}{\Pr [E | H_D, I]}$$

where the vertical conditioning bar ‘|’ is read as ‘given that...’ or ‘assuming that...’.

- 3.17 Symbolic notation formalises a logical truth of immense significance, which can readily be grasped intuitively without resort to mathematical formalisation: *all probabilities* (Pr) *of evidence* (E) *are conditioned on particular information* (I). If the conditioning information changes, e.g. because new relevant information subsequently comes to light, the probability of the evidence must be re-evaluated to take account of the new conditioning information. The new probability could well be greater or smaller than the original value; or it might remain unchanged. This general logical truth has direct and profound implications for forensic case work.

Scientific experts should consequently take steps to ensure that any consumers of their forensic reports are left in no doubt that all assessments, interpretations and conclusions presented in forensic reports are critically dependant on the validity, accuracy and constancy of (specified) conditioning information. If relevant conditioning information changes, or if the court or fact-finder rejects conditioning facts on which the scientist has relied, the scientist’s conclusion can, in general, no longer be safely relied upon. It is for this reason that most forensic providers nowadays include in the first part of their statements a standard-form paragraph flagging up the conditionality of their findings. It is wise to say, in so many words,

⁴⁸ i.e. as direct substitutions for PP and DP, respectively. ‘PP’ and ‘DP’, referring to rival ‘propositions’ in contested litigation, may be more intuitive to non-specialists; whereas the more mathematically literate may be more comfortable with thinking in terms of rival ‘hypotheses’, designated H_P and H_D .

that the scientist should be informed of any (known) changes in the conditioning information, and that any scientific interpretation or conclusions would need to be reconsidered if (specified) factual assumptions cease to be operative.

3.18 **B. ‘Purpose’**

Having identified relevant conditioning information in the first section of a report, the forensic scientist should next explain the *purpose* (i.e. the reasons or rationale) for undertaking particular forensic examinations or conducting particular tests. No reader of a forensic report should be in any doubt about the scope and purpose of the expert’s work. This section of a forensic report should provide a succinct account explaining how particular forensic inquiries or techniques were intended to address material issues in the case. Equally importantly, readers of a forensic report should be able to identify any pertinent issues that the expert has deliberately *not* addressed.

The phraseology typically employed to describe the purpose of forensic examinations will vary, partly depending on whether the forensic scientist is operating in evaluative or investigative mode.

When operating in *evaluative* mode, a typical statement of purpose would run along the following lines:

The purpose of my examination has been to help address the issue of whether the spade recovered from the defendant’s car boot had been used to dig the site at which the body was found.

Or for an *investigative* purpose, the following pattern is common:

My purpose in examining the clothing was to suggest possible types of weapon that could have caused the damage observed to the fabric.

3.19 Some providers include in this section a brief summary of the forensic techniques that were used to achieve the stated purpose. The following illustration elaborates on the *evaluative* purpose presented in the previous paragraph (‘to help address the issue of whether the spade

recovered from the defendant's car boot had been used to dig the site at which the body was found'):

For this purpose, I have examined the spade for the presence and distribution of deposits of soil and have utilised a variety of physical, chemical and biological analytical techniques to compare any soil recovered with the reference samples from the burial site.

3.20 C. 'Interpretation'

The content of the 'Interpretation' section of a forensic report is structured by the specified purpose of any given scientific inquiry, and turns significantly on whether the scientist is operating in investigative or evaluative mode.

When communicating *investigative* opinions, the expert should list possible *explanations* for observations or findings. The list of theoretical possibilities is never *literally* exhaustive,⁴⁹ and might not even cover all contextually plausible explanations in the case at hand – notwithstanding the forensic scientist's conscientious best efforts to be comprehensive, rigorous and fair in undertaking the evaluation. Explanations in the form of *posterior probabilities* can only be generated if the specified list of explanations is taken to be exhaustive (meaning their combined probabilities sum to 1) for the purposes of the analysis. On this assumption, the sum of the posterior probabilities for the explanations should also equal 1. Should further explanations subsequently need to be considered (e.g. because new information comes to light), the calculations must be performed again with updated prior, and therefore also posterior, probabilities. It is evident that undertaking such analyses, necessarily and unavoidably, rests on a particular expert's contextually formulated assumptions. If opinions in the form of *posterior probabilities* are provided by forensic scientists, it is therefore essential that the basis on which prior probabilities were assigned by the expert should be spelt out, so that the target audience(s) can assess for themselves, and for their own purposes, whether these assigned probabilities were realistic, apt and fair. Posterior

⁴⁹ If only because epistemically 'wild' possibilities like alien invasion, divine intervention or a Cartesian Evil Deceiver cannot be refuted empirically. Forensically less outlandish possibilities, such as deliberate evidence tampering or planting or negligent cross-contamination, could also conceivably expand the list of eligible explanations for physical traces.

probabilities can be calculated only in relation to specified explanations, taken – for these purposes – to be exhaustive.

Reports of *evaluative* conclusions should take the general form of discussing the probability of scientific observations ('the probability of the evidence'), assuming the truth of the prosecution's and defence's competing hypotheses (H_P and H_D), as specified by the case circumstances and the issues being addressed. The expert should lead the reader through the process of assigning probabilities for the scientific findings, describing and explaining whatever data have been relied on from (specified) other sources, as well as any information derived from the scientist's own personal experience. It should be possible for an informed reader to retrace the logic of the scientist's approach, recheck calculations, or reconsider the scientist's findings in the light of changed assumptions. It should be crystal clear to anybody relying on the report how the expert arrived at her opinion, and what assumptions have been made.

3.21 **D. 'Conclusion'**

The concluding section of a forensic report should briefly revisit the issue specified in the 'Purpose' section, and summarise how – if at all – the scientific findings have helped to address it. The report should convey the magnitude of the likelihood ratio favouring, or disfavouring, the main hypothesis of interest (usually, the prosecution's allegation) over its principal rival or rivals.

3.22 ***Appraising CAI: Practical Benefits, Preconditions and Challenges***

CAI is intended to benefit forensic scientists, purchasers of forensic science services and the broader public interest in the fair and effective administration of criminal justice. These benefits are not especially contingent on structural models of provision. Whilst CAI might appear naturally suited to client-supplier relationships in a commercial marketplace for forensic science services, similar benefits should be obtained in state-run systems of public provision.

If the CAI model is implemented effectively, forensic examinations commissioned by the client will be those, and only those, best calculated to resolve the key issues in the case, minimising costs for the client and promoting efficient use of resources. For the provider, applying the CAI protocol pinpoints those aspects of the case requiring more scientific data

and knowledge to underpin the provision of sound expert opinion. It also helps providers deploy resources efficiently in the pursuit of service improvement.

The criminal justice system overall will naturally benefit from expert evidence that is explicitly focused on the issues in the case and self-consciously addresses itself to the questions that need answering in the context of particular criminal investigations, prosecutions or trials. CAI opinions aspire to be balanced and justifiable. The process by which expert opinions are formulated should be transparent and readily testable. With its defined concepts and systematic procedures, rooted in the probability calculus and the logic of inferential reasoning, CAI supplies a common language to facilitate more effective communication between forensic scientists. If CAI principles were adopted, and understood, more widely by those who commission and use forensic reports, this should improve the quality and probative value of scientific evidence adduced in court and might help to dispel some of the widespread confusion and predictable misunderstandings that have dogged forensic science and expert witness testimony in the modern era.⁵⁰

- 3.23 To gain maximum benefit from the CAI model, providers and clients need to be closely engaged, with good communication and mutual understanding of each other's needs. Both parties need to understand how they can assist each other to enable forensic providers to deliver the best service, appropriately calibrated to clients' needs. Any misunderstanding between the parties, or misconceptions about their respective roles, is liable to diminish the efficiency of forensic science, and could have broader negative repercussions for the fair and effective conduct of criminal proceedings.

In a free-market model, commercial pressures and tight budgets predictably drive requests for the cheapest examinations. But price is not necessarily a reliable proxy even for cost-effectiveness, let alone for the broader aspirations of criminal justice (Roberts 1996; Lawless 2011). The cheapest forensic examinations often address only those questions at the lower levels of the issue hierarchy – *source* and *sub-source* issues – which typically require limited

⁵⁰ Systematic misunderstanding and communication breakdown between lawyers and scientists have been prominent themes in official reports and academic commentary on forensic science evidence and expert witness testimony for many decades. See e.g. Roberts (2013); Haack (2009); JUSTICE (1991: paras. 9.7-9.8); Wonder (1989); Jones (1986). For broader context, see Roberts (2014).

interpretational efforts on the part of the forensic scientist. This restricted approach carries significant forensic risks (Jackson 2013). When more expensive interpretational work has not been commissioned, and is therefore not undertaken, the probative value of the evidence is never fully considered by the forensic scientist and clients received limited advice on the meaning of the evidence. In the absence of a properly contextualised expert assessment of scientific findings, there is evident potential for the true probative value of scientific evidence to be misconstrued – either over- or under-estimated – by lawyers, courts or lay triers of fact with limited understanding of scientific method.

Furthermore, the cheapest examinations often involve requests for analysis of material collected by the police (or other clients) themselves. Such material has been ‘pre-sampled’ from a potentially far greater range and number of items, often without the forensic scientist’s input. Even where proper collection protocols have been observed and the integrity of physical evidence is assured, limiting the role of the forensic scientist in this way necessarily confines the analysis to *source* or *sub-source* issues. The scientist gains a very limited view of the wider significance of physical traces and is therefore unable to offer clients, or courts, any appraisal of the evidential value of scientific findings at *activity* level.

3.24 The basic concept of CAI may seem alien to some clients, and even to some providers. There may still be a rather stereotypical assumption that forensic scientists and other experts are capable of providing, and should provide, authoritative *categorical* answers to questions falling within their respective fields of expertise. When experts are unable to provide unequivocally dispositive answers, the next-best type of opinion, in the view of some clients, would be for the expert to provide an assessment of the probability of the contested event (i.e. a *posterior probability* for the proposition in light of the evidence). However, while *categorical* opinions and opinions expressing *posterior probabilities* may be valid and justifiable for *investigative* purposes, they are problematic when proffered as evidence in judicial proceedings because they involve making assumptions or drawing factual inferences on possibly contested issues which experts are not entitled to decide in criminal litigation. These are paradigmatically jury questions in our legal systems. The CAI model directs experts to avoid categorical opinions and posterior probabilities when operating in *evaluative* mode. Expert opinions should instead be expressed as *likelihood ratios*, encapsulating comparative probabilities for the *evidence*, rather than advancing any direct assertion about the probability of disputed events (AFSP 2009).

Forensic scientists' diverse contributions to the administration of criminal justice are currently not well understood, either by frequent consumers of expert witness testimony, including police, lawyers and courts, or even by many scientific experts themselves. Nor is there widespread appreciation of the different kinds of expert opinion and their respective strengths and weakness and proper roles in criminal investigations, prosecutions and trials. CAI has made significant progress in clarifying these distinctions. The CAI model for forensic case assessment and interpretation of results assists practitioners to devise optimal forensic strategies and provide relevant opinions in the most appropriate form for the investigative (policing) or evaluative (judicial) task in hand. Whilst progress to-date is encouraging, significant work still lies ahead in explaining CAI's philosophical underpinnings to the forensic and legal communities, exploring its practical implications with legal professionals and other consumers of forensic science, fostering mutual understanding between scientific experts and lawyers, and spurring further, ideally collaborative, refinements to the model.

4. CAI in Forensic Case Work: Three Detailed illustrations

4.1 This Part presents three detailed examples to further explore, and illuminate, practical applications of the Case Assessment and Interpretation method in forensic case work. The first example concerns a type of expert evidence, involving sniffer dogs, that may be somewhat esoteric for many practitioners, but which usefully offers a relatively straightforward illustration of the general CAI method and pattern of reasoning that can be grasped intuitively and is easily extrapolated to other, more prosaic types of expert testimony. Example #1 demonstrates how probabilities for observations may be assigned, and explains how these assignments in turn inform commissioning decisions, the development of effective forensic strategies, and robust, logical interpretation of the evidential value of scientific findings.

Example #2 focuses on two more familiar, routinely employed forensic techniques, DNA profiling and analysis of clothing fibres. This example highlights CAI's contribution in helping forensic scientists to develop effective investigative strategies and to decide which techniques to utilise when confronted with choices between viable alternatives.

Finally, Example #3, involving various types of trace evidence, explores the complexities of the scientist's role through the lens of CAI. It emphasises the distinction between the initial investigative phase of a forensic enquiry, when the forensic scientist is operating in 'investigative mode', and the different kind of work involved in providing expert evidence in the form of written reports or testimony for the court, when the scientist is operating in 'evaluative mode'. This example covers both types of forensic reasoning and analysis, describing the transition from one to the other in a fairly typical concrete case scenario.

4.2 ***Example #1: 'Sniffing for Death' – How to Evaluate the Response of a 'Cadaver Dog'***

Human remains detection (HRD) dogs, popularly known as 'cadaver dogs', are trained to give a signal when they detect the scent of human remains. These animals have been used extensively for searching disaster scenes and, to a lesser extent, in forensic investigations when searching for murder victims.

Like any other forensic detection technique, cadaver dogs are not infallible or 100% effective. On occasions, they will fail to give the expected signal in the presence of human remains (false negative response), whilst on other occasions the signal will be given despite the absence of any body (false positive response). It is consequently very important that investigators employing HRD dogs know and understand the meaning of the dogs' success and failure rates, and that they, or anybody else interpreting the evidence, are able to assess the true probative value of a signal, or its absence, in inferring the presence (or absence) of human remains in a particular location.

- 4.3 In order to make our illustration a bit more realistic and tangible, we will adopt the following facts from a German case reported by Oesterhelweg et al (2008):

A married couple went on a sailing trip. During this trip, the wife disappeared and was reported missing by her husband. A criminal investigation was initiated. The husband was immediately under suspicion of having murdered his wife. The State Police of Hamburg employed a cadaver dog to search the yacht, and the dog gave a signal on sniffing a mattress in the yacht's cabin.

The prosecution wanted to adduce evidence of the cadaver dog's signal in proceedings against the husband. However, little was known about the reliability of cadaver dogs, so new scientific trials were conducted. These trials involved three trained dogs and their handlers being presented with a line of six glass jars containing cut squares of carpet. Some of the squares had been impregnated, for varying periods, with the scent of fresh cadavers. In most trials, only one of the squares was impregnated with cadaver scent; in some, no squares were impregnated. The contents of the jars were unknown to the handlers – so it was a 'blind trial' in that sense.

- 4.4 The three dogs and their handlers were tested intermittently over a two month period, until 354 trials with the 'six jar line-up' had been completed. Table 4.1 summarises the results of these trials:

Table 4.1 Cadaver dog signals in the presence or absence of cadaver scent

Dog's response	Scent present	Scent absent
1. Positive signal	224	4
2. No signal	11	115
Total	235	119

It can be seen from Table 4.1 that the scent was present in 235 trials, and a dog detected it and gave the positive signal in 224 trials. There were 11 ‘false negatives’ where the dog failed to give the signal in the presence of the scent. It can likewise be seen from Table 4.1 that the scent was actually absent in 119 trials, of which the dog gave a ‘false positive’ signal in just 4 trials.

It is a simple matter to convert these raw data into relative frequencies indicating the sensitivity of the dog test:

$$\text{Proportion of signal/scent} = 224/235 = 0.953$$

$$\text{Proportion of no signal/scent} = 11/235 = 0.047$$

$$\text{Proportion of no signal/no scent} = 115/119 = 0.966$$

$$\text{Proportion of signal/no scent} = 4/119 = 0.034$$

4.5 These data were generated from conducted trials. In the ‘Assessment’ part of the CAI model, however, the aim is to assess the potential benefit of carrying out an examination *before* it is actually commissioned. In making such an assessment in this particular case, it would be useful for the investigator to have some idea of the reliability of the dog’s signal as an indicator of the presence of human remains. Reliability might usefully be expressed in terms of the probabilities of the dog’s giving positive or negative responses under specified conditions (including the presence or absence of cadaver scent).

As a general matter, historic data of past performance, suitably adapted to meet the exigencies of the instant case, can inform rational predictions of future performance. So it is possible to rely on empirical data and associated statistical frequencies, such as those produced by the German authorities’ cadaver dog trials, to assign probabilities and calculate

likelihood ratios for a cadaver dog's responses. These assignments and calculations are presented in Table 4.2:

Table 4.2 Conditional probabilities, and associated likelihood ratios, for cadaver dog signals

Dog's response (E)	Probability of the response, given scent is present $\Pr[E H]$	Probability of the response, given scent is absent $\Pr[E \bar{H}]$	Likelihood ratio $\frac{\Pr[E H]}{\Pr[E \bar{H}]}$
1. Positive signal	0.953	0.034	28
2. No signal	0.047	0.966	1/21 (0.048)
Total	1	1	

4.6 Table 4.2 puts the cadaver dog trial data into the familiar symbolic format introduced in *Practitioner Guide* No 1, and briefly summarised by Part 3 of the present *Guide*. The evidence, E, is the dog's signal or its absence. The proposition or hypothesis, H, is the presence of the scent. The negation of the hypothesis, \bar{H} , is the absence of the scent. (Notice that both variables are logical disjunctives – either the dog gives the signal, or does not give the signal: it cannot be both or neither. Likewise, either the cadaver scent is present, or it is not: it cannot be both or neither. The probabilities for the two responses, conditioned on the stated hypotheses, are complementary and therefore sum to 1, as shown in the 'totals' for the two columns. The probabilities across the rows are unrelated. They need not sum to 1.

As we have previously explained, the likelihood ratio (LR) is calculated by dividing the probability of the evidence (E) assuming (*conditioned on*) the hypothesis (H), by the probability of the evidence (E) assuming (*conditioned on*) its negation, \bar{H} . From here, it is a simple matter of factoring the relative frequencies derived from Table 4.1 into the new format, to produce the probabilities and likelihood ratios reported in Table 4.2

4.7 The type of information summarised by Table 4.2 is useful to the forensic scientist operating in investigative mode in making two kinds of assessment:

- (a) assessing the potential evidential value, or impact on the investigation, of the dog's anticipated responses; and

- (b) assessing the probability of obtaining those responses if cadaver scent were, or alternatively were not, present in fact.

4.8 (a) *Assessing the potential evidential value or investigative impact of the dog's responses*

If a cadaver dog were to be employed, there are two – and only two – logically possible outcomes. Either the dog will give a signal when exposed to a particular locus or material, or it will not.

For any positive signal obtained, the LR for this evidence is 28 (from the fourth column, first row of Table 4.2). Applying the 'odds' version of Bayes Theorem,⁵¹ whereby prior odds for the proposition are multiplied by the LR to produce posterior odds for the proposition in light of the new evidence, the prior odds of cadaver scent being present would be multiplied by a factor of 28 if a positive signal were obtained. Ideally, case specific prior odds should be assigned by the persons(s) best placed to take account of all relevant information known at the material time; which in the earlier phases of criminal process would logically be a police detective, aided by a forensic scientist with relevant experience and expertise. However, this ideal model of scientific input raises institutional, organisational and operational issues regarding the conduct of criminal investigations and the generation of judicial evidence which are somewhat jurisdiction-specific, and in any event beyond the scope of this *Guide*. Purely for illustrative purposes, we will proceed on the basis that, in the following hypothetical case, the scientist makes the assignment.

- 4.9 Suppose that our forensic investigator rated the prior odds of cadaver scent being present at the specified location as 100:1 against, or 1:100. If the dog were to give a positive signal, multiplying those prior odds by an LR of 28 would give posterior odds for the presence of cadaver scent of 100:28 – or approximately 7:2 – against. Thus, the odds would shorten significantly, but the odds would still be *against* cadaver scent being present (even, to repeat, after the dog had given a positive signal).

The LR for the alternative scenario, in which the dog gives no signal, is stated in the final column, second row of Table 4.2, to be 1/21 or 0.048, or (column 4, row 2). Notice that this LR is smaller than 1; confirming the untutored expectation that the dog has particular

⁵¹ As set out in Part 2 of *Practitioner Guide* No 1.

abilities, such that its failure to give a signal supports the proposition that no cadaver scent is actually present. For prior odds that scent is present of 1:100, the posterior odds for the presence of scent *after the dog has failed to signal* is calculated as $1/100 \times 1/21 = 1/2100$ (in words, “odds of 2100 to 1 against”). The posterior odds are smaller than the prior odds by a factor of 21, indicating that the dog’s failure to give the designated signal is evidence providing ‘moderate support’⁵² for the proposition that there is no cadaver scent present. And in the absence of cadaver scent, it may be reasonable to draw the further inference that no cadaver had recently⁵³ been in contact with that location.

- 4.10 Calculations employing Bayes Theorem refer back – and may be highly sensitive to – the designation of prior probabilities. If the forensic scientist had rated the prior odds of detecting cadaver scent as 50:50 (i.e. evens), instead of 1:100, a positive signal from the dog would convert the posterior odds of the scent’s being present into 28:1 on; an impressive probability of $28/29 = 0.966$ (or approximately 97%).

Of course, in none of these modelled scenarios does a positive signal equate to a 100% guarantee of cadaver scent, nor does the absence of a signal ever equate to 100% certainty that no cadaver scent is present. The dog is inherently fallible; and all empirical data are probabilistic. Bayes’ Theorem, and calculations of LR’s in particular, assist forensic scientists (and anybody else committed to clear thinking) to approach uncertainty in a systematic and logical fashion, enabling prior probabilities to be updated coherently in the light of whatever relevant new information is obtained.

- 4.11 (b) *The probability of obtaining the responses*

The second insight to flow from the LR table is the probability of a successful outcome if the technique were to be deployed. From the second column of Table 4.2, it can be seen that, if cadaver scent were actually present, the probability that the dog would respond in the ‘right’ way, with a positive signal, is approximately 0.95 or 95%. Correspondingly, the probability that the dog will erroneously fail to signal in the presence of cadaver scent (i.e. give a false

⁵² Employing a well-known scale of verbal expressions for likelihood ratios, reproduced in *R v T* [2011] 1 Cr App R 9, [2010] EWCA Crim 2439, [30]-[31]; discussed by Roberts et al (2011).

⁵³ Criteria of recency would need to be defined, in terms of an appropriate time interval calibrated to the test data. We set aside this complication here for the sake of simplifying the illustration.

negative response) is approximately 5%. If cadaver scent is not present in fact, the probabilities (as recorded in the third column of Table 4.2) are 97% for the correct response (no signal) and 3% for the wrong, false positive response (signal given in the absence of cadaver scent).

Data indicating the potential performance of particular tests or technique provide forensic investigators with valuable information that should inform decisions on the development of investigative strategies. The decision to employ a particular test or technique will invariably have time, resource or evidential implications – and quite possibly all three. Such decision-making ought to proceed on an informed basis. CAI encourages a systematic approach. For example, both the sensitivity of a test and its specificity need to be considered. An exceptionally sensitive test may be almost worthless if it produces too many false positives, whereas a moderately sensitive test (some false negatives) may be useful and cost-effective if it is also highly specific. These cost-benefit calculations are always, ultimately, subject to the contextual demands of particular investigations or legal proceedings.

- 4.12 It is always possible to pose methodological challenges to the validity of experimental data. In our example, the dog scent trials were conducted under (contrived) experimental conditions using carpet squares impregnated with the scent from recently-deceased bodies. The probabilities in LR tables should be derived from experimental data that are as representative as possible of the circumstances of the case in hand. For example, it would be best if cadaver dog scent trials replicated natural search conditions; and ideally, the actual dog used in the search of the yacht might have been tested in the German case. In reality, experimental test conditions are always (more or less robust) proxies for natural environmental conditions. There are difficult methodological questions of sampling size, representativeness, replicability of results and so forth (some of which can be subjected to advanced statistical modelling) which we do not need to pursue here.

In the final analysis, every individual case has its own unique facts and circumstances. Experimental data produce generalisations and approximations of reality, and even evidently flawed data may be superior to pure guesswork. Forensic scientists need to make intelligent use of whatever pertinent data might be available to help them tame the uncertainties and deal with the investigative and evidential demands of the instant case. Experienced and knowledgeable forensic experts will adjust the available data, taking account of its

methodological strengths and limitations, better to reflect what they judge to be the probabilities of obtaining particular scientific findings in the circumstances of particular investigations or proceedings.

- 4.13 To say that forensic scientists adjust data to particular case circumstances in light of their knowledge and experience is merely to describe what expert witnesses have always done. Indeed, the ability to identify, organise and interpret specialist information drawing on accumulated knowledge and utilising refined skills and judgement might be regarded as central to the very definition of an ‘expert’ in any disciplinary field.

The CAI model merely formalises and, importantly, makes transparent the explicit assumptions and contingencies of longstanding forensic good practice. The method of assigning and appraising likelihood ratios can be applied by experts in any traditional area of forensic practice. The discipline imposed by the method helps experts to become aware of implicit assumptions, which may not be fully articulated or substantiated, or aspects of the analysis lacking the underpinning of good data or knowledge. It is precisely in these equivocal areas where experts need to be especially wary of advancing premature or overconfident interpretations of their findings, unless and until relevant data are actually obtained.

- 4.14 ***Example #2: ‘The Robber’s Mask’ – How to Decide Which Scientific Techniques to Employ***
Forensic scientists are routinely confronted in their casework with circumstances in which alternative approaches are available, and choices have to be made in selecting which techniques to employ. The following hypothetical scenario is both a realistic casework example, and an exemplar of a broad fact pattern frequently encountered in relation to a range of physical traces and objects submitted for forensic examination.

Robber’s Mask: The information received by the scientist was that a balaclava was discarded by a robber at the scene of a bank robbery. The balaclava was quickly retrieved by a witness and handed to the police. A suspect was arrested and combings were taken from his head hair. The police submitted the balaclava and hair combings to a forensic science provider and requested an examination ‘to link the suspect with the balaclava’.

Let us now consider, step by step, how a scientist adopting the CAI model might proceed on these instructions.

4.15 *Step 1: Acquire information regarding all relevant circumstances of the case*

It is not unusual for there to be gaps in the initial information accompanying submissions to forensic science providers. Efforts should always be made to fill in significant gaps in background information, to enable the provider to understand fully the issues in the case and to inform *ex ante* assessments of the probability of finding particular evidential material.

Given the limited information initially received in *Robber's Mask*, the forensic scientist would request more background details before attempting to formulate an effective examination strategy. Suppose that, on contacting the police, the following additional information was obtained regarding (i) the offence and the offender, (ii) the subsequent investigation, and (iii) the client's requirements:

(i) Additional information about the offence/offender

- witnesses said that the robber held them hostage in the bank for an hour before he made his escape;
- the robber wore the balaclava all the time that he was in the bank;
- the robber talked at length and shouted frequently during the raid.

(ii) Additional information about the investigation

- the balaclava was handled and bagged by an officer who had no subsequent contact with the suspect;
- the suspect was arrested approximately 6 hours after the incident;
- hair combings were taken from the suspect 1 hour after arrest by an officer who had had no contact with the balaclava;
- in police interview, the suspect flatly denied any involvement in the robbery. However he admitted the balaclava was his. He explained that he used to wear it regularly, but that he hadn't worn it – or any other headgear – for over two months, when it mysteriously went missing and not been seen again. The suspect said that he assumed he had inadvertently left it somewhere, or perhaps someone had taken it;

- the suspect has not yet been charged with any offence.
- (iii) Additional information about the client's requirements
- forensic results are required for the suspect's next bail date in six weeks' time;
 - the suspect will be charged and prosecuted only if the scientific evidence is sufficiently strong to satisfy the Code for Crown Prosecutors.

Apprised of this further information, the forensic scientist is now in a position to start developing an appropriate and effective strategy. A time-frame has been ascertained; and the probative value of the scientific findings must be high for the case to proceed. If the forensic scientist judges that, put at their highest, scientific results could only provide limited support for the prosecution's case, it will be appropriate to advise the police or CPS accordingly.

The pre-trial role of the forensic scientist is conceptualised, under CAI, as providing maximum scientific assistance to investigators, lawyers, and prosecutors in the development of investigations and the conduct of criminal proceedings. Contextual information is often required to enable the forensic scientist to identify the salient issues in the case and perform this supporting role effectively. There are undeniable *risks* of confirmation bias in supplying contextual case-specific information to forensic scientists, but these can – and should – be minimised and managed, for example by setting up systems of sequential unmasking. Moreover, there is never good reason to risk biasing scientific investigations by exposing forensic scientists to *irrelevant* (and potentially prejudicial) information about a suspect or the case. The CAI approach contemplates scientists being provided with the contextual information they need to give well-informed advice to their clients – only that, and no more. There may be scope for considering institutional or procedural changes designed to insulate forensic scientists from the risks of contextual bias more systematically and effectively, e.g. by allocating 'assessment' and 'interpretation' functions to different personnel (where that separation is both feasible and appropriate).

4.16 *Step 2: Clarify and define the client's requirements*

The initial request from the police referred to establishing a 'link' between the suspect and the balaclava. However, linkage is a rather vague notion, scientifically speaking. The expert's

role here is to delve deeper and to try to define more precisely the evidential needs of the client so that these needs can be addressed effectively.

The forensic scientist's approach should be strictly logical rather than strategic (much less adversarial). It is not a question of 'trying to support the police hypothesis', but rather of thinking through how the evidence might be capable of doing so; and equally of how the evidence could logically contradict that hypothesis.⁵⁴ From the information received, the scientist might formulate the central question for further consideration in the following way:

Is this suspect the man who wore the balaclava during the robbery?

It is essential to formulate this framing question carefully and a precisely, ensuring that its answer logically bears on the issues in the case (i.e. that it truly 'links' the suspect to the crime). An alternative question, such as '*Has this suspect worn the balaclava?*', would be irrelevant. The suspect admits to wearing the balaclava; and his wearing it on other occasions is minimally probative on the question of his wearing it during the bank robbery.

In terms of the hierarchy of issues previously mentioned in Part 2.24 of this *Guide*, this would be an *activity* level issue - the level that may be of greater interest and value to a court than propositions addressed to the *source* or *sub-source* levels.

4.17 Having specified the material issue in terms of a central framing question, it is a relatively straightforward matter for the scientist to derive a pair of mutually exclusive propositions (one for the prosecution, and one for the defence), along the following lines:

H_P - The suspect is the man who wore the balaclava during the robbery

H_D - Some other man wore the balaclava during the robbery

When evaluating a likelihood ratio for this pair of propositions, the scientist needs to identify, and keep firmly in mind, those parts of the conditioning information (I) that should influence

⁵⁴ It is interesting to note, in this connection, that the Code of Practice, issued pursuant to s.23 of the Criminal Procedure and Investigations Act 1996, stipulates that '[i]n conducting an investigation, the investigator should pursue all reasonable lines of inquiry, whether these point towards or away from the suspect' (para.3.5).

the assignment of probabilities for her observations and analytical results. In this case, several pieces of background information constitute pertinent conditioning factors: the suspect said the balaclava was his and that he wore it two months earlier; witnesses say the offender wore the balaclava for an hour, and that he talked at length and shouted frequently whilst wearing it; hair combings were taken from the suspect approximately seven hours after the incident (one hour after his arrest).

4.18 *Step 3: Develop a proposal for a cost-effective examination strategy*

Various scientific techniques could conceivably be deployed to help address the central issue in this case, including chemical testing for saliva on the balaclava, DNA-profiling, and microscopic examinations for fibre transfer (from the balaclava to the suspect) or hair transfer (from the suspect to the balaclava). Clients are not necessarily willing to pay for all viable tests or for unconstrained pursuit of all potential avenues of inquiry. Informed decisions have to be made about which forensic techniques would provide the most cost-effective examination strategy for the case at hand. In *Robber's Mask*, the police are looking for strong evidence to determine whether, or not, the suspect was the robber, and for this to be produced within a defined time-frame and to budget.

On CAI principles, the techniques with greatest investigative potential are those with high probabilities of providing large LRs in favour of the prosecution proposition if the prosecution proposition were true, and large LRs in favour of the defence proposition if that, contrary proposition were true. These are the techniques with greatest discriminating power, where the evidence is most decisive, one way or the other. Evidence generating large likelihood ratios is especially probative, relative to the defined proposition pair (rival prosecution and defence hypotheses) and specified conditioning information.

4.19 For any chosen forensic technique, the scientist is required, at the outset, to think through all the potential observations and analytical results that might be obtained were that technique to be employed. Next, probabilities for obtaining these outcomes must be assigned, assuming the prosecution proposition H_P is true and taking account of all relevant case circumstances, i.e. the conditioning information (I). But this is only half the job. The scientist must then adopt a completely different perspective, and mind-set, in order to assign fair and realistic probabilities for all of the anticipated potential outcomes on the assumption that the alternative, defence proposition H_D is true, again taking account of the same conditioning

information (I). An important part of the conditioning information (I) in *Robber's Mask* is that the suspect admits owning the balaclava and wearing it two months earlier.

The process of assigning probabilities for a suite of potential scientific findings can be complex and challenging. The scientist will be combining knowledge and understanding from personal experience with data and findings derived from published experiments and surveys. There may well be numerous factors for the scientist to take into consideration, including: the probabilities of transfer, persistence and detection of particular trace materials; probabilities of obtaining 'matches' for different types of comparison; and probabilities of background occurrence (base rates) and variability of materials. Different experts may legitimately assign different probability values, reflecting varied personal experiences and knowledge of particular evidence types or a range of plausible adjustments of published experimental data or case-histories to accommodate idiosyncratic features of the instant case. Generally speaking, however, experts basing their evaluations solely on the same published data-set should arrive at similar assigned probabilities. These more personalised and, perhaps, speculative dimensions of experts' opinions are always susceptible to further exploration, and adversarial challenge, in criminal proceedings. It is another virtue of CAI that it makes the epistemological foundations of expert opinion more explicit and open to review.

- 4.20 A variety of investigative techniques and avenues for further forensic enquiry might be pursued in *Robber's Mask*. In order to keep the illustration at sensible length, we will confine our analysis to just two scientific examinations: (a) DNA-profiling of the mouth area of the balaclava; and (b) comparison of any fibres recovered from the suspect's hair combings with balaclava fibres. The same general approach could be applied to any number or type of forensic examinations.

There are multiple ways of capturing and presenting the data, calculations and findings produced by these examinations. We will stick with the likelihood ratio table format previously introduced by Table 4.2.

- 4.21 (a) DNA profiling

The scientist in *Robber's Mask* would begin by considering the technique of DNA profiling of the mouth area of the balaclava, and identifying the range of possible results that the use of this technique might produce. Owing to the wide range of conceivable potential results, pre-

assessment necessarily has to simplify the possibilities by consolidating them into a manageable smaller number of broad categories. Table 4.3 summarises the scientist's assignment of probabilities for this projected array of evidential findings, 'E'.

Table 4.3 Likelihood ratios for findings of DNA-profiling of the balaclava's mouth area

DNA profiling findings (E)	Probability of findings, assuming truth of prosecution proposition $\Pr[E H_P,I]$	Probability of findings, assuming truth of alternative proposition $\Pr[E H_D,I]$	Likelihood ratio $\Pr[E H_P,I]/$ $\Pr[E H_D,I]$
1) No profile obtained	0.02	0.02	1
2) Single profile matching suspect	0.87	0.35	~2.5
3) Single profile not matching suspect	0.01	0.04	0.25 (1/4)
4) Mixed profile containing components shared with the suspect	0.09	0.47	0.2 (1/5)
5) Mixed profile containing components not shared with the suspect	0.01	0.12	0.083 (1/12)
Total	1	1	

Note that this Table specifies that the probability of E is always also conditioned on I, as well as on the relevant hypothesis. This complication was omitted from Table 4.2 for the sake of simplicity, but it is an essential consideration in practice. If the conditioning information changes, the probability of the evidence assuming a given hypothesis often changes, too. This in turn may affect the likelihood ratio for the evidence. In *Robber's Mask*, it is significant that the conditioning information includes the admitted fact that the suspect wore the balaclava two months prior to the robbery.

The probability values presented in Table 4.3 are hypothetical (like the case); though in fact they do broadly resemble actual values from similar real cases. Methods for assigning these values need not concern us here. That would be a substantive topic for a training manual for forensic DNA experts.

4.22 Table 4.3 provides a basis for a range of significant inferential conclusions that would greatly assist a forensic scientist to assess the potential value of pursuing this forensic strategy.

Firstly (from row 2 of the second column), if the suspect truly were the person who wore the balaclava during the robbery (H_P), the most likely outcome (with 87% probability) is a single DNA profile matching the suspect. On the face of it, this looks like a very good chance of achieving a strongly probative result.

Probing the issue further, however, we see (from row 2 of the fourth column) that the likelihood ratio for the evidence is only about 2.5, which is not especially impressive. The reason that the LR in this case falls far short of the sort of values routinely associated with DNA profiling is that there is a significant pre-examination probability that the suspect's DNA would be found on the balaclava (assessed by the scientist as 35%, taking account of persistence over two months) *even on the assumption that H_D is true*; i.e. even if the suspect was *not* wearing the balaclava during the robbery (and by further logical inference, was not the robber, who *was* wearing it). A modest LR, as a measure of the probative value of the evidence, follows logically from a combination of the precisely specified proposition pair, H_P and H_D , and the conditioning background information, I.

A result of this nature might conventionally be reported as providing 'limited support' for the (prosecution's) proposition that it was the suspect, rather than some other man, who wore the balaclava during the robbery. An LR value greater than 1 logically supports the proposition, rather than its proposition-pair alternative; but a value of 2.5 could not fairly be described as supplying anything like 'strong' support for the prosecution's case.

4.23 A second interesting feature of Table 4.3 is that a 'No profile' finding gives an LR value of 1. In other words, finding no profile at all is equally likely on either the prosecution's or the defence's proposition, and would in no way discriminate between them. (When the LR is 1 the posterior odds remain the same as the prior odds; the evidence has failed to help us decide between rival contentions.)

On the other hand, the probability of obtaining this outcome is low – only 2% - whichever proposition, H_P or H_D , is true. Such a remote risk of producing entirely unhelpful findings might well be worth running, from the client's point of view.

- 4.24 Thirdly, a fascinating prediction from Table 4.3 is that for all three scenarios in which a non-matching or mixed DNA profile is obtained, the evidence would favour the defence proposition, H_D , over the prosecution proposition, H_P . We know this because the LR in each of rows 3, 4 and 5 is less than 1. Moreover, by summing the probabilities of each of these outcomes in the second column (assuming the truth of H_P) we can say that there is a combined probability of 11% ($0.01 + 0.09 + 0.01$) that the evidence will logically support the defence case *even if the suspect was in fact the robber*. To spell this out, even if the police have arrested the right man – the man who wore the balaclava during the robbery, and therefore was the robber – employing this technique would run an 11% risk of *weakening* the prosecution’s case.
- 4.25 Fourthly, turning now to examine Table 4.3’s third column, a parallel analysis can be applied from the defence point of view. On the assumption that H_D (‘some other man’ wore the balaclava) is true, there would be a combined probability of 63% ($0.04 + 0.47 + 0.12$) that DNA profiling results would support H_D - in those three outcomes with LRs smaller than 1. In each of these three scenarios, the magnitude of the LR favouring H_D (by factors of four, five and twelve, respectively) is greater than the magnitude of the LR (2.5) in the one scenario favouring the prosecution’s proposition – i.e. where a single matching profile is found.

Furthermore, even if H_D is true (implying that the suspect is innocent of the robbery because some other man wore the balaclava), there is a 35% probability of obtaining an outcome (single matching profile) that would increase the strength of the evidence against the suspect, represented by an LR of about 2.5. To spell this out, there is about a one-in-three chance that conducting the DNA profiling would produce evidence appearing to (mildly) incriminate a suspect *on the assumption that he is actually innocent*. (Whether, in practice, such evidence would be generated and adduced to the detriment of a factually innocent suspect would depend on many other contingencies, that the forensic scientist has neither the authority nor the wherewithal to assess.)

- 4.26 This strictly logical analysis of the potential of DNA profiling to provide probative evidence, either supporting or contradicting the prosecution’s case theories, can be conducted before any analytical scientific work is undertaken. CAI proposes that a scheme of potential

outcomes and their relative probabilities, such as that summarised in Table 4.3, should form the basis of conversations between clients and forensic providers regarding which scientific techniques would best meet the client's needs. Of course, additional practical considerations, including the price of the technique and the timescales involved, must also be factored into any decision to commission particular scientific work.

4.27 (b) Fibres in hair combings

A parallel analysis can be performed for examination of any fibres in the hair combings, employing the CAI approach. To facilitate this analysis, the scientist requires knowledge of the type of fibres forming the balaclava. The probabilities of fibre transfer between the balaclava and its wearer's head hair are partly determined by the type of fibres involved, e.g. natural versus synthetic fibres.

Assume that the scientist performs a preliminary visual examination of the balaclava. She ascertains that fibres are readily shed from the fabric, supporting a high probability of fibre transfer. The scientist can now assign probabilities for the full range of potential evidentiary outcomes, E, of a thorough examination of the hair combings from the suspect.

This scheme of probabilities would be conditioned on the same proposition pair as previously formulated when assessing the likely outcomes of DNA profiling, namely:

H_P - The suspect is the man who wore the balaclava during the robbery

H_D - Some other man wore the balaclava during the robbery

4.28 Table 4.4 summarises the scientist's probability assignments, and the associated likelihood ratios, for the range of outcomes for this technique.

Table 4.4 Pre-assessment table for fibre examination

Outcome of searching hair combings for fibres matching the balaclava (E)	Probability of outcome, assuming prosecution proposition true $\Pr[E H_P, I]$	Probability of outcome, assuming defence proposition true $\Pr[E H_D, I]$	Likelihood ratio $\frac{\Pr[E H_P, I]}{\Pr[E H_D, I]}$
1) No matching fibres	0.01	0.92	~ 0.011 (1/92)
2) Few (1-5) matching fibres	0.05	0.05	1
3) Some (6-15) matching fibres	0.24	0.02	12
4) Many (>15) matching fibres	0.70	0.01	70
Total	1	1	

For the purposes of this analysis, ‘matching’ is understood in the conventional sense of ‘corresponding with the constituent fibres of the balaclava in all aspects tested’.

Categorising fibre findings according to a tripartite division of ‘few’, ‘some’ and ‘many’ is arbitrary, in the sense that more or fewer categories could conceivably be employed. The numerical intervals differentiating one category from the next are similarly ‘arbitrary’. One could have designated the ‘few’ category as 1-10, ‘some’ as 11-25, and so on. These selections are made on the basis of the forensic scientist’s relevant knowledge and experience with handling similar evidence in the past (and consequently they are, in principle, no more ‘arbitrary’ than any other experience-based clinical judgement). The categories and intervals summarised in Table 4.4 in fact broadly reflect what might be expected in such cases, though it is not necessary for our purposes of illustrating the CAI analytical technique that the specified probabilities should be realistic. The numbers of fibres in each interval could easily be adjusted, along with their respective probabilities, to whatever sizes are deemed appropriate in any given case.

4.29 As before when considering the likely outcomes of DNA analysis, various interesting inferential conclusions can immediately be drawn from Table 4.4 without undertaking any further scientific work.

Firstly, if the suspect truly was the person who wore the balaclava during the robbery (H_P), the most likely outcome, with 70% probability (second column, row 4), is ‘Many (>15) matching fibres’. The associated likelihood ratio for this outcome is 70, indicating that the prior odds of the prosecution proposition (H_P) being true (*whatever* those prior odds were assessed to be) would be increased by a factor of 70 were this outcome to be obtained when the analysis was carried out.

Secondly, we can see from the fourth column of row three that the outcome of ‘Some (6-15) matching fibres’ also gives an LR greater than 1, which means that this outcome, too, would favour the prosecution proposition (H_P), albeit not to the same degree as ‘Many matching fibres’. The LRs for *both* of these outcomes notably provide greater support to the prosecution’s proposition than any anticipated results from DNA profiling. Combining these two potential outcomes for fibre examination indicates a combined probability of 94% (0.24 + 0.70) that this technique would generate evidence supporting the prosecution’s proposition (i.e. LR > 1).

4.30 Thirdly, the outcome of ‘Few (1-5) matching fibres’ gives an LR value of 1. In other words, this finding would be entirely neutral as between the prosecution’s and defence’s respective propositions and the evidence could not assist in discriminating between them. To the extent that these propositions were the only matters of interest, the evidence would be irrelevant. However, the probability of actually obtaining this outcome, conditioned on either proposition H_P or H_D , is fairly low at 5%.

Finally, even if the suspect truly was the man who wore the balaclava (H_P), there would be a 1% probability of obtaining an outcome that would advance the alternative proposition (H_D) that it was some other man. As with the previous logical pre-assessment of DNA profiling, there is a small but non-trivial possibility that the scientific work would support the defence case even on the assumption that the suspect wore the balaclava mask during the robbery (and, by further inference, is guilty of the robbery).

4.31 Turning now to the defence point of view, if the alternative proposition of ‘some other man’ (H_D) were really true, then the most likely outcome, with 92% probability (first row of the third column), would be ‘No matching fibres’. This outcome would generate an LR of 1/92; indicating that a finding of ‘No matching fibres’ would increase the prior odds on the

suspect's not being the balaclava wearer (H_D) by a factor of ninety-two. This would be, in fact, the most probative of all forensic findings considered thus far.

Conversely, there is a 3% risk ($0.01 + 0.02$) that scientific analysis of fibre transfer would increase the weight of evidence against the suspect, even on the supposition that H_D is true (meaning that the suspect was *not* the balaclava wearer, and by further inference, was not the robber). As can clearly be seen from Table 4.4, the third and fourth possible outcomes for the third column, where the probabilities are all conditioned on H_D 's being true, are associated with LRs greater than 1 – meaning that the evidence supports H_P , the prosecution's proposition that the suspect was wearing the balaclava during the robbery. Although finding many (>15) fibres was judged by our forensic scientist to have only 1% probability of occurring (put another way, the forensic scientist would anticipate this outcome in one out of every hundred relevantly similar cases), that finding – notwithstanding the suspect's innocence – would actually supply the most probative evidence of *guilt* of any of the modelled scenarios.

How can this be? In one sense, this result is merely logical confirmation of the truism that very incriminating evidence can be presented against innocent accused. People can *look* very guilty without being guilty, for the simple reason that evidence is not always what it seems and even generally very reliable evidence of guilt is occasionally misleading. The fact that the police or prosecution may be making unilateral commissioning decisions, which affect the prospects for proving innocence as well as the likelihood of generating incriminating evidence (including evidence incriminating the innocent), raises broader normative questions of fairness and due process which are beyond the scope of this *Guide*.

4.32 As before with the CAI pre-assessment of DNA profiling, conversations between provider and client in *Robber's Mask* would centre on exploring and understanding the potential benefits and limitations of undertaking fibres analysis, as well as the costs and timescales involved.

Of course, the client may wish to commission both types of examination (and possibly others, as well). In that event, the LR table for the second technique would need to be reconsidered, and possibly amended, in the light of the findings produced by the first technique. Such

reappraisal is necessary in order to accommodate any dependency there may be between the two different techniques (as explained by *Practitioner Guide No 1*, paras 3.40-3.43).

If there was no dependency, the LRs for the two sets of outcomes could be multiplied together, in accordance with the ‘product rule’ for combining the probabilities of independent events. For example, if the scientific results obtained were ‘Single profile matching the suspect’ (LR = 2.5) for technique #1, followed by ‘Many matching fibres’ (LR = 70) for technique #2, their combined LR, applying the product rule (assuming independence), would be 175 (2.5 x 70). In these circumstances, the combined effect of the two forms of scientific analysis would be to increase the prior odds of the prosecution’s proposition – that the suspect was wearing the balaclava during the robbery – by a factor of 175. Evidently, the *combined* probative value of scientific (or, indeed, of any other) evidence may be far greater than the probative value of its individual parts taken in isolation.

It bears repeating that combining LRs in this fashion must be approached with care and circumspection. Forensic scientists need to be alive to the possibility of submerged or concealed dependencies between items of evidence, which then need to be taken into account in producing composite LRs. Furthermore, only LRs relating to the same *activity* level issues and propositions, or which address the same *source* level issue conditioned on identical proposition pairs, can be combined. But whenever LRs generated from two or more lines of scientific inquiry can legitimately be brought together in a coherent and logical way, their combination may contribute significant additional information that a forensic scientist should consider in making casework pre-assessments of forensic investigative strategy and in advising clients, in accordance with the CAI model.

4.33 *Step 4: Consult client, explain potential outcomes, costs and timescales, and agree an examination strategy*

The scientist’s assessment of the potential outcomes, benefits and drawbacks of the two techniques considered would form the basis of further dialogue between provider and user. As well as communicating essential information to the client, this two-way process of dialogue helps the forensic provider to engage more directly with the exigencies and demands of criminal proceedings and to form a better appreciation of other actors’ expectations, and of the expert witness’s professional duties to the court. Overall, that routine dialogue which is integral to the CAI model should promote the delivery of more effective and efficient

forensic science services, counteracting the very well-documented deleterious effects of communication breakdowns at all stages of the process through which forensic science evidence is generated, interpreted and evaluated (Roberts 2013).

Part of the background information communicated to the forensic scientist in *Robber's Mask* was the prosecution's need for sufficient evidence to charge and prosecute. On the basis of the scientist's pre-assessment of the two techniques' potential evidential value, and with particular regard to the respective LR values in the best-case prosecution scenarios, let us say that the client decided to commission the fibre examination but concluded that DNA profiling would not be worthwhile.

4.34 *Step 5: Carry out the examination*

The fifth step in the CAI process covers the actual conduct of whatever scientific examination, technique, or analysis the client agrees to commission. This *Guide* is not concerned with the mechanics of physical comparisons or the 'wet chemistry' etc of particular forensic disciplines.

4.35 *Step 6: Interpret the findings*

Once particular forensic examination(s) have been completed and results obtained, the sixth step in the CAI model requires the scientist to revisit her initial assessments (as recorded, for example, in tables of likelihood ratios) and to refine the likelihood ratios originally assigned to the range of anticipated outcomes. A great strength of the CAI model is that anticipated probabilities are projected in advance, and these establish an articulated framework for interpreting any finding subsequently obtained. In this way, the scientist to some extent protects herself against the risks of *post hoc* rationalisation of scientific findings to fit with the client's case theories – a prophylactic all the more necessary to the extent that, as behavioural science data indicate, nuanced *ex post facto* accommodations between data and theories are typically largely the product of unconscious biases. The discipline imposed by CAI on the logical interpretation of scientific findings in this respect does for forensic scientists what Odysseus' comrades did for him in tying him to the mast of his ship, knowing that otherwise he might find it impossible to resist the Sirens' song.

Let us say that, in *Robber's Mask*, the outcome of the scientist's examination was that 20 fibres recovered from the suspect's hair matched the balaclava. This fits in with the most

likely (70%) anticipated finding if the suspect had worn the mask during the robbery, with an LR of value 70 (reading across row 4 of Table 4.4).

Twenty matching fibres is obviously a larger number than 15 fibres, the threshold criterion for the 'Many fibres' category. The scientist could conceivably try to refine the likelihood ratio specifically to accommodate 20 matching fibres, but there would be little to be gained from this minor adjustment. It is the broad magnitude of the LR that is significant in assessing the probative value of the evidence, rather than its exact value (or limited range of values).

On the other hand, if a *much* larger number of matching fibres had been found - say 600, or 1,000 - re-appraisal of the initially assigned probability would be necessary. Our forensic scientist, we supposed, had anticipated that finding 15 or more matching fibres in the circumstances might reasonably be categorised as 'many'. Six hundred matching fibres would be completely off the scale of reasonable expectations, making any calculated LR for this outcome difficult to evaluate. Such a finding might ring alarm bells with the forensic scientist, prompting a switch to *investigative* mode to explore alternative *explanations* for such unanticipated results. Perhaps the physical evidence has been innocently contaminated or tampered with in some way, or maybe the case circumstances are significantly different from those communicated to the scientist. Whatever the true explanation in this case, further enquiries with the client would be an obvious, and perhaps necessary, course of action.

4.36 *Step 7: Communicate the findings and interpretative opinion*

Forensic scientists and other expert witnesses typically set out their findings in a report, which informs the client in the first instance, and may subsequently be itself adduced in court (if the case goes to trial) or form the basis for the expert's live oral testimony. CAI offers a logical framework for helping to structure experts' forensic reports (also see Evett et al 2000).

When it comes to presenting and interpreting scientific findings, forensic reports should begin by explaining the reasoning behind the choice of techniques employed and specify the issues and propositions (including their level(s) within the hierarchy of issues and propositions) that particular scientific investigations were attempting to address. These choices should have already been discussed with clients earlier in the process, which – as we

have seen – features proactive dialogue between service providers and their clients. Scientific findings will then be described in appropriate detail, followed by the expert’s interpretation of those findings. Readers of a forensic report should be able to follow the logic of the thought-process employed by the expert in developing an examination strategy, conducting agreed scientific tests, and arriving at (revised) values for a likelihood ratio – or ratios – for the scientific findings, conditioned on clearly specified background information and proposition pairs.

4.37 The likelihood ratio, as an expression of the probative value of the evidence conditioned on particular proposition pairs and specified background information, is at the heart of the CAI method. It is vital that the magnitude of the LR should be communicated to clients and other criminal justice audiences in an accessible way. Currently, there are two predominant reporting formats. In the first, the magnitude of the LR is expressed as a numerical ratio indicating the extent to which it supports one proposition in the proposition pair over the other. For example:

The finding of 20 matching fibres in the head hair combings of the suspect is approximately 70 times more likely to be obtained if the suspect, rather than some other man, wore the balaclava at the relevant time.

In the second reporting format, the degree to which the LR favours H_P or H_D (or is neutral between them, where $LR = 1$) is expressed as a verbal generalisation, as in the following example:

*The finding of 20 matching fibres in the head hair combings of the suspect provides **moderate support** for the proposition that it was the suspect, rather than some other man, who wore the balaclava at the relevant time.*

Most forensic science providers in the UK have adopted scales of verbal equivalents for likelihood ratios, running from ‘limited’ or ‘weak’ support, through ‘moderate’ to ‘strong’ and ‘very strong’ support, or slight variations on a similar theme. Verbal equivalents are probably more immediately accessible to non-experts than numerical LRs, but whether they effectively convey gradations of probative value or not is a controversial issue. The degree of compatibility between scales of verbal equivalents employed by different forensic providers,

or across disciplinary divides, is also a matter for further critical examination and policy development. In the meantime, *transparency* should be the forensic scientist's watchword: schemes and numerical ranges of LR_s, and their corresponding verbal equivalents (if used), should be clearly set out in the forensic report.

4.38 *A plausible factual variation – and its dramatic impact on the likelihood ratio for the evidence*

It is instructive to consider what might have happened if the client in *Robber's Mask* had submitted the case in a different form and had specified a different type of service.

Imagine that the police client had sampled the mouth area of the balaclava for saliva residue or skin cells and had submitted only this sample and a reference DNA sample from the suspect. The request was solely for a DNA analysis, with minimal further case information supplied. In that situation, and without further discussion with the client, the scientist can only address a *sub-source* issue, i.e. whether the crime scene DNA came from the suspect. The proposition pair then to be considered would necessarily take the following form:

H_P - The crime scene DNA came from the suspect

H_D - The crime scene DNA came from an unknown person, [genetically] unrelated to the suspect

Assume that, on analysis, a full DNA profile matching the suspect was obtained. The probability of obtaining this match, conditioned on the prosecution proposition (*H_P*), is close to 1, i.e. it is practically certain (absent lab contamination, foul play, etc – of which, there is absolutely no indication in this case) that a match would be obtained. The probability of obtaining a match, assuming the truth of *H_D*, is generally accepted to be of the order of 1 in a billion (this is the standard random match probability for a full SGM Plus™ DNA profile). This produces a likelihood ratio of the order of 1 billion for the evidence (because 1 divided by 1/1 billion = 1 billion).

4.39 The difference between the LR_s in the original *Robber's Mask* and this plausible and realistic variation is staggering. We said before that considering the DNA profiling results at activity level (wearing the mask during the robbery) might produce, on a prosecution best-case

scenario, a likelihood ratio of 2.5. A likelihood ratio more favourable to the prosecution (LR = 70) could reasonably be anticipated if fibre analysis were undertaken, and – having verified that the two types of evidence were genuinely independent – a combined LR in favour of the prosecution’s (activity-level) proposition (that the suspect was the balaclava-wearer) as high as 175 might be obtained if both tests were done. Yet an LR of one billion for the prosecution’s (different) proposition at sub-source level (the suspect is the donor of the crime scene DNA) utterly eclipses any of the other scientific findings, and could plausibly be described as providing extremely strong support for the prosecution’s case.

So what has gone wrong? The answer is that, *nothing has gone wrong*, from a strictly logical point of view. All the analyses described are sound and their results valid, given the specified propositions and assumptions. The same evidence can be made to appear highly or only moderately probative, depending on the particular proposition pairs it is taken to address. In this case, the vital distinction is between evidence of contact⁵⁵ between the suspect and the balaclava (which, recalling the extended facts of the illustrations, the suspect is admitting anyway) and evidence that the suspect wore the balaclava *during the robbery*. In reality, proof of the latter does more to support the prosecution’s overall case than evidence of the former, irrespective of the different magnitudes of likelihood ratios. It is to be hoped that these distinctions would be clearly brought out for the benefit of the fact-finder at any trial of the issue, though there is no guarantee that a jury would not be distracted by large likelihood ratios and expert testimony asserting ‘extremely strong support’ for (one of) the prosecution’s contentions (i.e. association between the suspect and the balaclava). The prospects for pre-empting confusion and successfully communicating a logical assessment of the evidence to the fact-finder should be considerably enhanced if forensic scientists, lawyers and courts properly understand the rudiments of likelihood ratios as a measure of probative value, within the framework of the CAI model.

⁵⁵ Note that, strictly speaking, forensic science evidence warrants the inference that the suspect is the source of the DNA on the balaclava, from which the obvious *further* inference is that the suspect was in contact with the balaclava. But it is certainly *possible* that the suspect’s DNA came to be on the balaclava, e.g. through secondary transfer, without the suspect ever having come into contact with the balaclava.

4.40 At all events, it is readily apparent from Example #2, and its perfectly realistic factual variation, that potentially misleading expert opinion evidence could be given if a *sub-source* evaluation were to be presented instead of an *activity* level evaluation. Furthermore, by not requesting fibres examination, the client has overlooked the opportunity to secure additional probative evidence, not merely of physical association, but of guilt.

4.41 ***Example #3: ‘The Case of the Older Offender’ - Moving from Investigative to Evaluative Opinion***

This third illustration of the CAI approach to forensic case work is based on a real case, with some incidental modifications. It highlights, in particular, a commonplace transition in forensic examination from *investigative* to *evaluative* mode.

The case circumstances, as communicated to the forensic scientist, were as follows:

Intruder Murder: Police requested the attendance of a forensic scientist to help examine the scene of a murder. The body of an elderly woman had been discovered by her neighbour in the living room at the front of the deceased’s home. The victim had sustained numerous injuries to her head that had bled and she had extensive bruising to her body. She had been tied up with a length of what appeared to be washing line. Her clothing had been pulled up, exposing her breasts and upper legs. There was extensive bloodstaining in the living room, where items of furniture appeared to have disturbed. Further bloodstaining was present in the hallway. The point-of-entry to the house was probably a broken, stained-glass window in the dining room at the rear of the property. From the disturbance of the deceased’s clothing, the attending police officer suspected there was a sexual element to the assault. The police had no identified suspect at this time.

We will now elucidate how a forensic scientist, operating within the framework of CAI, might approach this case, beginning with (a) the investigative phase of the inquiry, and then turning to (b) the evaluative phase.

4.42 (a) *The investigative phase of the forensic examination*

The first four steps of the CAI approach, as previously outlined in paras 3.2–3.9, are:

- Step 1: Acquire all relevant circumstances of the case;
- Step 2: Clarify and define the client's requirement;
- Step 3: Develop a proposal for a cost-effective examination strategy; and
- Step 4: Consult client, explain potential outcomes, costs and timescales, and agree an examination strategy.

Having outlined each of these four steps in previous examples, there is no need to repeat the basics. We can consider the forensic scientist's initial assessments and actions holistically, grouping these four steps together.

Given that the police have requested attendance at the scene, and in the absence of an identified suspect, the forensic scientist is obviously being asked for *investigative* assistance. For these purposes, the scientist requires very little additional background information about the circumstances of the case (in sharp contrast to what is required when operating in *evaluative* mode, as we have seen).

4.43 Interactions with clients in *investigative* scenarios are typically more direct and immediate than laboratory-based work. The scientist normally has face-to-face conversations with investigating police officers before, and whilst, attending the scene. The earlier stages of the CAI model, when operating in *investigative* mode, should structure the forensic scientist's approach to understanding and agreeing the client's investigative priorities during these exchanges. Crime scene examination is necessarily a dynamic phase of criminal investigation. Observations made by the forensic scientist may open up new lines of inquiry or prompt reconsideration of existing case theories or hypotheses. A rigid, formulaic model would not be appropriate for this fluid, evolutionary stage of the investigative process, which perforce demands inventive thinking on the part of detectives and forensic scientists alike.⁵⁶ Nonetheless, the underlying inferential logic of the CAI model remains the sheet anchor of a rational, systematic approach to forensic examination.

When operating in *investigative* mode, the forensic scientist generates provisional hypotheses and assigns plausible prior probabilities to them, drawing on initial information supplied by

⁵⁶ The early parts of investigations are characterised by imaginative, 'abductive' reasoning as explained in *Practitioner Guide No 3*.

police investigators and her own knowledge and previous experience. For each subsequent observation made at the crime scene, existing hypotheses can be re-ranked to take account of the probability of particular observations under each hypothesis. The process is reiterated for every further observation that is made. At the end of the examination, the scientist should be in a position to offer investigators a list of explanatory hypotheses for the state of the crime scene, ranked by order of decreasing probability (*posterior probabilities*). The logic of the scenario is that detailed observations of the crime scene have now been made, and the question is: what events could have produced a crime scene that looks like *this*? The discipline of the CAI approach, requiring hypotheses to be reconsidered and updated following the discovery of each new pertinent item of information, helps the scientist to guard against *ex post* rationalizations to accommodate observations to ‘pet theories’ or premature commitment to the most initially appealing hypothesis. If the scientist is unable or unwilling to commit herself to assessing *posterior probabilities*, then she may simply list all feasible *explanations* without attaching any probabilities to them.

In *Intruder Murder*, the police require any information that would help them understand the sequence of events at the scene, and that might help them to identify potential suspects. Put in more general terms, the police are seeking information that would reduce the range and degree (and perplexity) of the uncertainties confronting their investigation into what happened, and who is responsible.

4.44 Steps five and six of the CAI model are (5) Carry out the forensic examination(s); and (6) Interpret the findings.

On examination of the scene, the scientist in *Intruder Murder* draws a number of inferential conclusions (or expert ‘opinions’) based on her observations of such things as the pattern of bloodstaining, disruption of living room furniture, and damage to property. For illustrative purposes, each observation or inferential conclusion in the following list has been coded ‘E’ for *explanation*; ‘PostP’ for *posterior probability* for an explanation; or ‘C’ for *categorical opinion*.

- The washing line in the rear garden had been cut (C), and may have been the source of the length used to tie up the deceased (E).

- A fresh-looking footwear mark in soil below the point of entry window was very likely to be that of the intruder (PostP).
- Initial confrontation between the intruder(s) and the woman, and the first injuries that bled, very likely occurred in the hallway (PostP).
- Further violent assaults occurred in the living room (C).
- It was possible that sexual assault (involving penile penetration)⁵⁷ took place (E).
- A blood-stained whisky bottle in the dining room was probably handled and drunk from by the intruder (PostP).

Measurements and photographs were taken of the scene, and relevant evidential items were bagged, labelled and taken back to the laboratory for further examination. Having examined these various items more closely and taking account of additional analytical information, the forensic scientist was able to draw the following further case-specific inferences and conclusions:

- The implement used to cut the washing line most likely had a sharp, single-edged blade (PostP).
- The piece of washing line used to tie the body was once part of the washing line in the garden (C).
- The footwear mark below the point of entry to the premises was made by a shoe that had a broad, plain, crepe-like sole (C).
- That shoe was probably a man's size 7 or 8 (PostP).
- Based on the forensic scientist's expert knowledge of footwear mark comparisons, and drawing on an investigative database of shoes and their sole patterns, the shoe was probably of a suede boot style (PostP).
- Wearers of such footwear are likely to be 'older' persons (PostP).
- A large amount of semen was present on internal vaginal swabs (C).

⁵⁷ We should note, for the sake of completeness, that it is *logically possible* that the victim may have had consensual intercourse with a third party, or even with the murderer, prior to the fatal attack. This possibility appears exceedingly remote on the facts of *Intruder Murder*, but one can readily imagine counterfactual variations on the case scenario which would make this possibility more salient: e.g. the donor of semen subsequently recovered from the victim turns out to be her (completely innocent) husband.

- The semen could logically have been the result of vaginal intercourse with ejaculation *or* penetration with an object or body part carrying semen on its surface (E).
- However, drawing on the forensic scientist's previous experience and expert knowledge of sexual assault, the most probable explanation for the semen was penile penetration with ejaculation (PostP).
- A DNA-profile obtained from the neck of the whisky bottle indicated that the donor of the DNA was male (C).
- Given that (female) V herself could not be the donor of the whisky bottle DNA, and in the absence of any pertinent information about male relatives or recent male visitors, the donor of the DNA might well be the offender (PostP).

4.45 The forensic scientist's initial analyses and conclusions in *Intruder Murder* usefully put some illustrative flesh on the bones of Part 2's more schematic discussion of explanations, posterior probabilities and categorical opinions.

Recall that *explanations* are hypotheses generated to explain observations or scientific findings. They might equally be described as 'possibilities'. *Explanations* may give investigators new avenues of enquiry, as with the forensic scientist's initial conclusions that the washing line might be the source of the cord used to tie up the victim and that sexual assault may have taken place. Or the scientist's explanations may seem to reinforce existing theories (perhaps it was already evident to detective that this was a sexual assault), or possibly contradict them (perhaps the police were initially thinking, on the basis of general offender profiles, that the assailant was probably a younger man).

An obvious limitation of *explanations* is their non-exhaustive nature. It may be that other possibilities exist and have not been considered or generated by the scientist. For example, the forensic scientist's examination of the scene does not exclude the logical possibility that V was raped and murdered by the semen donor, and subsequently robbed by an entirely different person, the 'older' crepe-soled suede boot-wearer. A further limitation is the lack of any precise numerical assessment of the probability of the *explanations* being true, although the explanations could still be ranked according to more broad-brush verbal statements of probability – in which case they should be characterised as 'posterior probabilities' for the explanations.

4.46 *Posterior probabilities* provide investigators with the scientist's considered assessment of the probability of the truth of particular explanations. Such probabilities might be expressed through precise numerical formulations, or using more qualitative, 'fuzzy' linguistic expressions such as 'highly probable', 'more likely than not', 'quite unlikely' etc. A scientist adopting the full CAI protocol might even be able to offer a list of viable explanations in descending order of probability. Thus, the forensic scientist in *Intruder Murder* might be able to advise police detectives of how likely it is that the person who made the footprint under the window was wearing male size 7 or 8 shoes at the time. Investigators might plausibly conclude, if only as a working hypothesis, that the print-maker was a man with size 7 or 8 feet. A shoe print with a distinctive sole pattern might support further inferences that the print was made, say, by a suede boot, which in turn may favour the possibility of the wearer being an 'older' person rather than a trendy youngster, etc. It might be possible for a forensic scientist to assist investigators to generate quantified probabilities for at least some of these predictions,⁵⁸ and in that way to assist in the efficient allocation of investigative focus and resources.

Numerical probabilities seem like a step up from qualitative probabilities for explanations, but scientists offering posterior probabilities (when operating in investigative mode) must proceed with care and express themselves with clarity. In addition to being subject to all the same limitations as explanations for observations and findings, the validity of posterior probabilities is conditioned on the forensic scientist's assignment of fair and realistic *prior probabilities* for the explanations. Such assignments rest on expert knowledge, skill and experience; and there is plenty of room for legitimate differences of scientific opinion in many areas. Complete transparency, buttressed by the logical structure of CAI, is vital for providing reassurance to other criminal process actors (and thereby sustaining confidence in the administration of criminal justice) that the scientist has been scrupulously fair and realistic in her assignments of prior probabilities.

⁵⁸ Strictly speaking, these are *post-dictions*: probabilities for events that have already happened rather than predictions of what might happen in the future. But 'prediction' is the vernacular expression, and might intuitively be thought of in terms of a prediction as to what the police or a court will find when (in the future) they discover the truth about what happened (in the past).

Categorical opinions, as we saw in Part 2, purport to state certain conclusions, which investigators may particularly welcome in reducing investigative uncertainties, closing down unpromising avenues for inquiry and excluding competing case theories. Categorical opinions are not overtly probabilistic, as posterior probabilities for explanations *always* are. However, although categorical opinions assume the posture of logical deductions, inferences from empirical data are never truly deductive in the classical logical sense. They are always *implicitly* probabilistic. When the forensic scientist in *Intruder Murder* states categorical conclusions about the presence of semen on V's vaginal swabs or the gender of the donor of the whisky bottle DNA, these *may be* safely treated as fixed points in the investigation, but they do not establish that the probability of alternative scenarios is zero – literally *impossible*. For example, there is always a residual possibility that physical exhibits have been tampered with or become contaminated, and forensic scientists are never in position to rule out *all possibility* of such eventualities in any given case.

4.47 Step 7 in the CAI model involves communicating findings and interpretation to the client. As we have said, the scientist's inferential conclusions and opinions should be conveyed to the police (or other client) in an appropriately concise, informative and timely fashion. The relative strengths and limitations of the expert's opinions should be clearly flagged up to the client. The differences between explanations, posterior probabilities and categorical opinions, as briefly recapitulated and illustrated in the preceding paragraphs, might be amongst those material considerations of which clients should be made aware, both in general terms and, crucially, as they bear on the instant investigation or criminal proceedings.

Suppose that, as a result of further enquiries and with the benefit of reports received from our scientist and a fingerprint expert, the police in *Intruder Murder* identify a suspect, Mr. S, who is 45 years old. S is arrested two weeks after V's body was discovered. Following a largely 'no comment' interview punctuated by blanket denials, S is charged with V's murder and rape. Various items are recovered from the suspect's person and during a post-arrest search of his home. A pair of shoes and a single-bladed knife are amongst the items seized and submitted to the forensic provider for further analysis and comparison with physical material collected from the scene.

It is at this point in the case-history that the forensic scientist's approach switches from *investigative* to *evaluative*, within the overarching framework of the CAI model.

4.48 (b) *The evaluative phase of the forensic examination*

Returning to the beginning of the CAI model at the outset of this new phase of the forensic examination, the first two steps are (1) to acquire information regarding all relevant case circumstances; and (2) clarify and define the client's requirements.

In *Intruder Murder*, the forensic scientist had already attended the crime scene and subsequent investigative briefings, and so was fully apprised of the general circumstances of the crime. However, some key pieces of information were still missing. This is quite typical. The course of an investigation and prosecution may change over time, not least because new information may emerge through the preliminary phases of proceedings and during pre-trial hearings. As we have stressed throughout, in order to provide their best assistance to commissioning clients and courts, forensic scientists need a clear and reasonably comprehensive understanding of the facts in issue in the proceedings. Informed appreciation of the prosecution's allegations, and of any counter-assertions being advanced by the defendant, are especially significant. Effective briefing is facilitated by on-going communication and close engagement of the scientist with investigators, prosecutors or defence lawyers throughout the course of the proceedings as the case develops, from suspect to discontinuance, guilty plea or contested trial.

4.49 How would the forensic scientist in *Intruder Murder* approach the evaluative phase of her work, within the broad framework of CAI?

We have seen that activity level issues, questions and propositions are, generally speaking, more pertinent to the resolution of contested issues in criminal adjudication than source or sub-source level issues, questions or propositions. Given that the accused S is simply denying the charge (which doesn't help the forensic scientist to narrow down the issues, as she could, for example, if the suspect were admitting physical contact but claiming accident, self-defence or provocation), the following activity-level issues (formulated as questions) might suggest themselves to the forensic scientists as matters on which she might usefully provide further assistance:

- *Was the knife recovered from S's home used to cut V's washing line?*
- *Did the shoe recovered from S make the mark in V's garden?*

- *Is S the person who broke V's dining room window?*
- *Is S the person who entered via that window?*
- *Is S the person who struck the victim several times about the face and body?*
- *Is S the person who tied V up?*
- *Is S the person who penetrated V's vagina with his penis?*

This list of issue-specifying questions is by no means exhaustive, but it is sufficient to illustrate the type and range of facts in issue that could be addressed with the help of scientific evidence. Note, to repeat, that these are all activity-level questions, concerning actions (also potentially including omissions) that a person might or might not have done. For most questions, the specified issue is whether S did the relevant act or omission, or not. But notice that the first two questions are neutral as to identity of the actor. The facts that the knife was used to cut the washing line, and that the shoe made the mark under V's window, if proved as facts, would not necessarily show that S did the cutting or was wearing the shoe at the material time.

None of these questions is addressed to a *source* or *sub-source* issue. They do not ask about the provenance of bodily fluids, or the donor of DNA contained in bodily fluids, for example. Nor are they *offence* level questions. Thus, the final question asks whether S penetrated V sexually, not whether he *raped* her. The existence of semen on vaginal swabs, taken in isolation, cannot discriminate between non-consensual sexual penetration (rape) and consensual intercourse (sex).

4.50 The next stage in the forensic scientist's analysis is to formulate proposition pairs for each (activity-level) question of interest, reflecting the prosecution's allegation(s) and the defence's assumed or advised position (in *Intruder Murder*, blanket denial). Here are some examples:

H_P – This knife is the implement that was used to cut the washing line

H_D - Some other implement was used to cut the washing line

H_P – This is the shoe that made the mark in V's garden

H_D - Some other shoe made the mark in V's garden

H_P – *S is the person who broke the dining room window*

H_D – *S did not break the dining room window; some other person broke it*

H_P – *S penetrated V's vagina during the incident*

H_D – *S did not penetrate V's vagina; some other man did so*

Note that, for the last pair of propositions, the forensic scientist has assumed that V had vaginal intercourse with some man (S or somebody else) close to the time of her murder. This reflects an experienced-based expert judgement of *posterior probability* on the part of the forensic scientist, conditioned on known case circumstances, as explained above. But this is just a contingent feature of the example; it is not implied by the CAI method. If the forensic scientist did not want to make any such assumption, or if it were made and later challenged in the proceedings, the proposition H_D could easily be revised accordingly. By excising the second clause after the semi-colon, for example, H_D would be expanded to include (i) somebody other than S penetrated V *and* (ii) nobody penetrated V (she did not have vaginal intercourse with anybody) as negations of H_P .

- 4.51 The third and fourth steps in the CAI model are (3) Develop a proposal for a cost-effective examination strategy; and (4) Consult client, explain potential outcomes, costs and timescales, and agree an examination strategy.

As we explained in previous examples, for each pair of propositions, the scientist would consider the probability of obtaining the range of observations and analytical results for the various available techniques, conditioned on the rival assumptions that either the prosecution's proposition (H_P) or a mutually exclusive alternative (H_D) is true. From that evaluation, the scientist can provide the client with an informed indication of the potential weight of evidence that could be provided by different types of forensic examination, as well as how likely it would be to obtain analytical observations or findings with that degree of probative value in the instant case. The ultimate decisions on which techniques to commission remain the client's prerogative (and responsibility).

For illustrative purposes, we will confine further analysis to just two issues and their associated proposition pairs. The first issue concerns the identity of the shoe that made the mark under V's broken window; the second relates to V's apparent sexual violation.

4.52 (i) Footwear mark evidence

Building on the scientist's opinion, formed in the investigative phase of the examination, that a shoe (rather than some other object) had made the mark identified at the scene of the crime, appropriate proposition pair would be:

H_P – This is the shoe that made the mark in V's garden

H_D - This is not the shoe that made the mark in V's garden; some other shoe made the mark

The forensic examination would involve visual inspection, and measurements, of a plaster cast and photograph of the assumed footwear mark in the soil below V's broken window. This initial examination would later be augmented by comparisons with test marks made by the forensic scientist using S's shoe as a reference sample. The scientist's pre-assessment of potential values for LR_s for the evidence that might be obtained by forensic examination – ideally, at this *pre-assessment* stage, taking account only of information relating to the crime scene mark – is reproduced by Table 4.5.⁵⁹

⁵⁹ We should stress that this is a simplified illustration; not least because a 'match' is short-hand for the outcome of what may be a complex process of analysis to identify similarities and dissimilarities in the class characteristics and acquired features of questioned ('crime scene') marks and reference marks (from known sources). As employed here, 'matching' means 'corresponding in all respects and showing no unexplainable differences'.

Table 4.5 Pre-assessment of footwear mark examination (addressing propositions relating to S’s shoe)

Outcome of comparing scene mark with S’s shoe (E)	Probability of observations, assuming truth of prosecution’s proposition $\Pr[E H_P,I]$	Probability of observations, assuming truth of alternative proposition $\Pr[E H_D,I]$	Likelihood ratio $\Pr[E H_P,I]/\Pr[E H_D,I]$
1) No ‘match’	0	0.998	0
2) ‘Match’	1	0.002	500
Total	1	1	

4.53 Table 4.5 records the scientist’s pre-assessment that, if ‘no match’ were the result of the visual comparisons, the LR (fourth column, row 1) would be approaching zero (simplified to 0 in Table 4.5). An LR approximating zero implies a posterior probability of zero for the prosecution’s proposition, that S’s shoe made the mark. In other words, it is a practical certainty (i.e. certain for all practical purposes) that S’s shoe did not make the mark if the crime scene footwear mark and the reference marks made by the scientist using S’s shoe do not match. Conversely, on the assumption that S’s shoe did make the mark, obtaining a ‘match’ is practically certain (assigned a conventional value of 1, in row 2 of the second column). If S’s shoe made the mark it will ‘match’ in the relevant sense, even if there are explicable visual discrepancies.⁶⁰

These may seem like self-evident and mathematically trivial conclusions. The finding of ‘no match’ refutes the hypothesis of common source (just as OJ Simpson’s lawyers catechized the Californian jury, ‘if the glove don’t fit, acquit!’). But it is worthwhile to illustrate how intuitive inferential reasoning, in this instance at least, smoothly coheres with CAI’s logical framework of inductive inference.

Moreover, the forensic scientist’s pre-analysis offers a quantified assessment of the probability of observing ‘no match’ on the rival assumption that S’s shoe did *not* make the mark; some other shoe made it. This probability is high, at 99.8%, but it is notably not tantamount to 100%. Significantly, the forensic scientist assigns (or estimates) a 2%

⁶⁰ For further exploration of the complexities implicit in the idea of ‘matching’ physical traces to their source(s), see Champod, Evett and Jackson (2004).

probability of observing a match *on the assumption that some other shoe – not S's shoe – made the mark.*

- 4.54 The LR for a 'match' has the value 500 (fourth column, row 2). This is determined by the random match probability (RMP) within a defined population, which is the probability of a 'match' without a common physical origin. Unless a particular trace or test is *uniquely* identifying (or 'individualising') for a specific source, there is always an RMP, because there are at least two items in the world with relevantly matching characteristics. For example, there might be two – or fifty-two, or fifty-two thousand – similar shoes (only one of which is S's) that could have made the crime-scene mark outside V's dining room window.

The RMP for the footwear mark evidence in *Intruder Murder* would follow from the forensic scientist's assessment of the rarity of the matching footwear pattern in some imagined population of shoes that potentially could have left the crime scene mark. This assessment should be informed by whatever data may be available indicating the incidence or distribution of footwear types among potential offenders, mediated by the scientist's experience and knowledge of footwear types. If it were believed that the offender was a local man, patterns of footwear ownership in the locality would have the greatest salience for the forensic scientist's assessment. If the geographical pool of potential offenders were extended, say, to the county boundaries or to the entire country, the relevant pool of footwear should ideally be adjusted accordingly. Conversely, it is unlikely to be relevant to the scientist's assessment of a British crime scene that in many (sunnier) parts of the world many people wear sandals or go barefoot.

From all this information, the relative frequency of occurrence of the pattern in the relevant shoe population has been assessed by the forensic scientist as 1 in 500, represented mathematically as 0.002 or 0.2% (because 1 in 500 is the same as 2 in 1,000). Hence, the probability of a random 'match' within the relevant population (i.e. the RMP) is assigned a value of 0.002. A corresponding LR of 500 means that evidence of a match increases the prior odds of *this* particular tested shoe being the actual shoe that made the crime scene mark by a factor of 500. If this were *in fact* the shoe that made the mark, it is certain that this outcome would be obtained ($\Pr[E|H_p, I] = 1$; as shown in Table 4.5's row 2, second column). There is only a tiny – but still non-trivial! – 0.2% probability of obtaining that outcome if the tested shoe, though matching, did not in fact make the mark. Thus, in a population of, say,

2,000 eligible pieces of footwear, four shoes would be expected to randomly match the crime scene mark. On the assumption that the eligible footwear population also includes the actual shoe that made the mark, and given that this shoe, if tested, would match through physical causation rather than probabilistic expectation, the probability that any one particular matching shoe actually made the mark (taking account of no other information beyond the fact of an observed match) is (only) 1/5 or 20%. That is to say, the matching shoe could be any one of the four 'randoms' or the true mark-maker. Table 4.5 handily and vividly encapsulates essential analytical results.

4.55 As we have been stressing all along, it is essential to formulate proposition pairs carefully, and to take careful notice of the precise terms of propositions formulated (or implicitly assumed) by others. The hypothesised propositions for the footwear mark evidence in *Intruder Murder* specify that S's shoe did, or conversely did not, make the mark. They do not address the different (also activity-level) proposition that S himself made the mark by stepping in the soil of V's garden. In order to address this issue, the forensic scientist (or, later in court, the fact-finder) would have to take a view on the probability that S was wearing the shoe at the material time. Whilst this may be a reasonable inference from the fact that the shoe is his, that inference is not inevitable and could be blocked by other evidence in the case. S would not be the first person in the history of the world who had lent his shoes to somebody else, shared footwear with family members, or who had recently acquired second-hand shoes, etc.

4.56 LRs are always predicated on precisely specified propositions. Whenever conditioning propositions change, the corresponding LRs need to be reevaluated.

Imagine, for example, that the defendant in *Intruder Murder* subsequently conceded that it might have been his shoe that left the mark in V's garden, but claims that it must have happened whilst he was doing repairs to the window some four weeks prior to V's murder. The original proposition pair would now need to be revised, along the following lines, to accommodate the fact that the question of *when* the mark was made, rather than the identity of the shoe that made it, has become the contested issue:

H_P – *S's shoe made the mark in the garden at the material time (of the offence)*

H_D – *S' shoe made the mark in the garden four weeks earlier*

A forensic scientist confronting this scenario would be primarily concerned with the appearance of the mark in the soil rather than with the incidence or distribution in some relevant population of footwear capable of making the mark (the RMP). Questions to be considered would include:

- How likely would it be to observe a mark of this nature in the soil if it had been made by this shoe at the time of the offence?
- How likely would it be to observe a mark of this nature in the soil if it had been made by this shoe four weeks earlier?

A scientist following CAI protocols might, or might not, have the necessary knowledge and expertise to answer these questions. Expertise in the changing appearance over time of footwear marks made in soil would be required.

4.57 A forensic scientist with the requisite expertise, and with access to adequate records of the crime scene mark suitable for detailed examination, might be able to express the opinion that the appearance of the mark is far more probable if it had been made at the relevant time (during the assault) rather than four weeks earlier; i.e. the probability of E assuming H_P , would be considerably greater than the probability of E assuming H_D .

The magnitude of this ratio might be difficult to quantify with precision. Nonetheless, if the broad ratio of the probabilities is scientifically justifiable, the LR would be greater than 1 and the expert could testify in good faith that the evidence supported the prosecution's, rather than the defence's, proposition – albeit to an unspecified, because unspecifiable, degree.

Conversely, in other cases with different facts – perhaps involving other soil types or adverse environmental conditions – a forensic scientist with expertise in soil analysis might conclude that the probabilities of the evidence conditioned on either of the two propositions in the proposition pair were quite similar, within the tolerance of whatever precision this type of forensic examination is capable of attaining. In that case, the LR would be approximately 1,

and the scientist would be able to report that the evidence was incapable of discriminating between the truth of the prosecution's and defence's rival contentions. Confined to this *precisely specified* issue (and corresponding proposition pair), the evidence would be irrelevant.

4.58 (ii) Semen swabs

Turning now to the semen evidence recovered from V's post-mortem vaginal swabs, CAI analysis begins, as always, with a pair of appropriately specified, mutually exclusive (but not necessarily exhaustive) propositions. In the first part of her analysis, the forensic scientist judged that somebody had vaginally penetrated V with ejaculation, and the only question was as to identity: who did it? A pair of propositions reflecting this initial judgement might be:

H_P – S penetrated V's vagina with his penis at the material time

H_D - Some other man penetrated V's vagina with his penis at the material time

The 'material time' is a deliberately vague conditioning presupposition of this proposition pair. Given the stated facts of *Intruder Murder*, we know that 'the material time' is likely to correspond to 'during the burglary' – this is the pertinent time interval. However, it is *not* logically necessary for a forensic scientist to condition DNA profiling on the assumption that a burglary had taken place; nor would it be appropriate to do so, strictly speaking. 'Burglary', after all, is a *legal* rather than a scientific category.⁶¹

4.59 The scientific examination for this part of the analysis would involve comparison of S's DNA reference sample with the results of DNA profiling from the semen recovered from V's post-mortem vaginal swabs.

Table 4.6 summarises the forensic scientist's assessment of potential outcomes for this examination. The figures quoted are illustrative, assuming full SGM Plus™ profiles were subsequently obtained. If only partially matching profiles were obtained, the probability of a match, conditioned on the defence proposition, would be higher; and, the LR would

⁶¹ Notably, the legal definition stipulated by s.9 of the Theft Act 1968 (England and Wales) is much broader than conventional linguistic usage.

consequently be smaller than LR_s for full profiles. The value of the LR, representing the probative value of the evidence of ‘matching’ profiles, depends on the degree of completeness of the crime scene profile in question: incomplete profiles are less probative than full or fuller matching profiles. (*Practitioner Guide No 2* gave extended consideration to the calculation and interpretation of likelihood ratios for full, partial and mixed DNA profiles.)

Table 4.6 Assessment of DNA profiling on semen swabs (assuming full SGM Plus™ profile with RMP of the order 1 in 1 billion)

Outcome of DNA profiling (E)	Probability of outcome, assuming truth of the prosecution’s proposition $Pr[E H_p, I]$	Probability of outcome, assuming truth of alternative proposition $Pr[E H_d, I]$	Likelihood ratio, $Pr[E H_p, I]/Pr[E H_d, I]$, for fully matching SGM Plus™ profile
1) No profile obtained	0.01	0.01	1
2) Single profile matching S	0.97	1 in a billion	→ 1 billion
3) Single profile not matching S	0.01	0.98	0.01 (1/98)
4) Mixed profile containing components matching S	0.01	0.01 x (1/ 1 billion) = 1 in 100 billion	1 billion
5) Mixed profile containing components not matching S	Very small (0.0001)	0.01	0.01 (1/100)
Total	1	1	

4.60 Recall that this table of probabilities for the evidence, E, and associated likelihood ratios, is conditioned on a pair of *activity* level propositions, the activity being sexual penetration of V. Some of the outcomes reproduce the very large LR_s typical for *sub-source* evaluations. However, others provide LR_s smaller than 1, which support the alternative proposition,

Inspecting the values summarised in Table 4.6 more closely supports several interesting logical inferences.

Firstly, the outcome judged overwhelmingly most likely (97%), if S were the man who had sexually penetrated V, is a single matching profile (row 2). Assuming a full SGM Plus™ profile with an RMP of 1 in a billion, this would give an LR of the order 1 billion; clearly

very powerful evidence of the identity of the person who penetrated V (to repeat, the *activity* on which the relevant proposition-pair was conditioned).

Secondly, three other potential outcomes are judged equally (un)likely conditioned on H_p , i.e. assuming that S performed the act of penetration. Each of these outcomes, however, is thought to have only a 1% probability of occurrence on the assumed facts. One of these outcomes, namely a mixed profile containing DNA matching S's profile, could potentially also support an LR of (the order of) 1 billion, supposing that the DNA mixture could be fully resolved into a full major and a separate minor profile (as discussed in the second *Guide* in this series). On the face of it, this is still impressively powerful evidence in favour of S's being the man who penetrated V. However, the presence of a second donor clearly implies the possibility of a second hypothesis, namely that the donor of the minor profile was actually the one who penetrated V. Or perhaps they both did? The police would very likely make further enquiries into the circumstances of the crime. This should spark the scientist, in conjunction with the police, to move into *investigative* mode and seek further background information that might help to explain the presence of multiple donors in the semen sample.

4.61 Thirdly, by far the most likely outcome (approaching a probability of 1, i.e. practical certainty) if, in fact, some other man had sexually penetrated V, is a single non-matching profile (third column, row 3 in Table 4.6). This gives an LR value of approximately 100 in favour of the alternative (defence) proposition (third column, row 4) – recalling that any LR value less than 1 favours the alternative proposition, with a magnitude inversely proportional to LRs greater than 1 supporting the prosecution proposition. (So here, an LR of 0.01 or 1/98 favours the defence proposition to the same extent, in terms of weight of evidence, as an LR of 100 would favour the prosecution's proposition.) The same LR value is given by the fifth potential outcome (third column, row 5), involving a mixed profile in which none of the components matches S's DNA. Neither of these outcomes conclusively eliminates S from suspicion (an LR of 100 is significant, but far from compelling), but they would both provide material evidence to support S's denials. And in either of these scenarios, the police would obviously want to pursue further lines of enquiry, presumably starting with a search of the National DNA Database to see if it contains any profiles matching the unknown donor in the instant case.

Finally, the outcome ‘no profile obtained’ (row 1), unsurprisingly, would be neutral evidence favouring neither the prosecution proposition nor its alternative (LR = 1). Notably, however, the forensic scientist judges that the probability of obtaining this outcome on either conditioning assumption is small (only 1%).

4.62 The probabilities assigned by the scientist and presented in Table 4.6 would reflect expert knowledge and experience of DNA profiling techniques and their evidential value (utilising relevant statistics), particularly with regard to physical properties of transfer, persistence and detection of semen and DNA (as *Practitioner Guide No 2* explained). Different forensic scientists might well assign somewhat different probabilities, based on their own personal experiences and knowledge and professional judgement, including judgements about the best way to characterise and quantify uncertainties. There are no ‘right’ answers or ‘correct’ values for assigning these probabilities.

It follows that such opinions, though genuinely expert, are inherently challengeable. Individual forensic scientists must explain and be prepared to justify their assessments if called upon to do so, by a client, an adversarial opponent or the court. There are well-defined data on frequencies of occurrence of genetic profiles in particular (ethnically stratified) populations and widely accepted forensic protocols for assigning match probabilities. Even so, there is an irremediably subjective element in deciding how to adjust raw frequency data to meet the demands of the instant case. In the final analysis, expert opinion evidence derives its authority and credibility from the personal standing and professional judgements of particular experts, as the criminal courts have recognised.⁶²

4.63 The final three steps of the CAI protocol involve (5) Carrying out the examination; (6) Interpreting the findings; and (7) Communicating the findings and their interpretation to clients and/or courts in expert reports. We need say no more about step 5, and steps 6 and 7 can now be addressed fairly summarily, building on previous examples.

⁶² Recently, see e.g. *R v Nicholson* [2012] 2 Cr App R 31, [2012] EWCA Crim 1568, [45], where the Court of Appeal endorsed the admissibility of expert evidence where ‘the nature of the evidence was such that no statistical evaluation could realistically be attempted.... The experts were, in the main, reporting the clinical experience of themselves and their colleagues and comparing the available research with the present cases’.

Let us say that the forensic scientist in *Intruder Murder* performs DNA profiling on genetic material recovered from V's post-mortem vaginal swabs, and obtains a full single profile matching S. As previously explained, the two conventional formats for presenting this result in expert reports or testimony are:

The findings are approximately a billion times more likely to be obtained if it were S, rather than some other unidentified man [genetically] unrelated to S, who penetrated V's vagina with his penis at the material time.

The findings provide extremely strong support for the proposition that it was S, rather than some other unidentified man [genetically] unrelated to S, who penetrated V's vagina with his penis at the material time.

Expert witnesses testifying to the results of DNA profiling in court, as well as lawyers adducing written DNA reports as expert evidence, must take care to ensure that the format chosen to express profiling results complies with formal legal requirements. The applicable law may vary from place to place. For example, some legal jurisdictions may demand that only a match probability or 'random occurrence ratio' (as the courts sometimes describe it) should be quoted to the jury, and might forbid the citation of likelihood ratios. These are (local) rules of law, rather than (universal) rules of logic. *Practitioner Guide No 2* reviewed the current situation in England and Wales more comprehensively.

4.64 *An instructive (and realistic) factual variation*

As for Example #2, it is instructive to conclude this extended illustration of the CAI approach in practical case work by briefly exploring an – entirely realistic – variation on the facts.

Imagine that, by the time *Intruder Murder* progresses to trial, the defendant is admitting the burglary, but denying sexual penetration or murder. His explanation for the presence of his semen on the victim is that, post-mortem, he removed V's underclothing, took out his penis and masturbated over her body, ejaculating over V's pubic area.

This material change in the relevant case circumstances, requiring major revisions to the conditioning information ('I') on which the scientist's original assessment and interpretation

were predicated, would require the scientist to re-evaluate her findings. There is no longer any valid objection to conditioning the analysis on a burglary's having taken place, since the defendant has admitted it.

4.65 A new proposition pair must be formulated, to reflect the defence's revised contentions, thus:

H_P – S penetrated V's vagina with his penis during the burglary

H_D – S ejaculated over V's pubic area, but did not penetrate her vagina with his penis

The scientist must now ask herself: how likely is it that I would have obtained my findings if S had penetrated V? And how likely is that I would have obtained my findings if he had ejaculated over her pubic area without penile penetration?

If swabs had been taken from exposed areas of V's body, they could be examined for the presence of semen. The forensic scientist could also review the distribution of semen-staining on the woman's clothing, and compare the relative quantities of semen that were found on the external and internal vaginal swabs. Taking all these observations into account, and drawing on her relevant forensic knowledge and experience, the scientist might well conclude that the evidential observations and findings (E) are more probable if the prosecution proposition (*H_P*) were true rather than if the defence proposition (*H_D*) were true. On that basis, the semen evidence would favour the prosecution's allegation (penetration) over the defence contention (masturbation). The likelihood ratio in that case would be greater than 1. How much greater? The scientist might judge it legitimate to try to quantify E's degree of support for *H_P* by assigning values to the probabilities of obtaining the observations. This would enable a numerical LR to be calculated. Alternatively, the scientist might form the opinion that there are insufficient relevant data to assign sufficiently precise numerical probabilities for her observations in this case. In that event, her opinion would have to be formulated in looser, qualitative terms, utilising an appropriate scale of verbal proxies for LR values (e.g. 'this evidence provides limited/moderate/strong, etc support for the prosecution's allegation that S sexually penetrated V'). Yet this must still be an intellectually disciplined reasoning process. The scientist would need to guard against the ineradicably human temptation towards intuitively latching onto a convenient verbal descriptor of evidential strength and then simply concocting sympathetic LR values by way of *post-hoc* self-justification.

Note that the random match probability for a matching DNA profile entirely drops out of this re-assessment of the semen swab evidence. We are no longer concerned with *identifying* the donor of the semen, but rather with the question of what a known donor might have done at the material time. The semen's relevance becomes a function of its physical transfer, persistence and detection under different conditions, rather than a marker of its biological origin. This assessment can be carried out coherently only at *activity* level, addressing the question: what did S do? S's conduct, as opposed to the offender's identity, cannot logically be addressed as a *source* or *sub-source* level issue.

5. How to Read Forensic Reports and Interpret Expert Opinions

- 5.1 This concluding Part draws together and concisely summarises material presented earlier in this and the previous three *Practitioner Guides*, in order to offer practical guidance on how to read forensic reports intelligently and interpret expert opinions without tears. Our suggested approach is logical, systematic and flexible, reflecting the robust methodological underpinnings of CAI. It can be used not only for initial appraisal of forensic science statements and reports, but also as a basis for cross-examining expert witnesses in court.

Our guidance divides into the following four stages:

Stage 1: Classify the type of expert opinion being given

Stage 2: Identify the level in the hierarchy of issues that the opinion addresses.

Stage 3: Test the class and level of that particular type of opinion.

Stage 4: Assess whether the expert's examination strategy was appropriate to address (particularised) contested facts in issue.

Expert testimony and reports often cover multiple issues, and these may be of quite different types, ranging from hard scientific fact to speculative opinion. It is recommended that the structured approach to interpreting expert evidence set out in this Part should be applied, sequentially, to each discrete conclusion or opinion expressed by the expert.

5.2 *Stage 1: Classify the Type of Expert Opinion*

Part 2 of this *Guide* explained how expert opinions could be classified according to the format and phraseology in which they are expressed. In particular, we distinguished (a) explanations, (b) posterior probabilities (for explanations), (c) categorical opinions, and (d) probabilities for observations and likelihood ratios. Concrete examples were provided in *Practitioner Guide No 1* and earlier in this *Guide*. Further illustrations of each category are summarised below.

There is also a group of expressions which – for want of a better term – we classify as (e) indeterminate opinions. These expressions come with major health warnings, and are probably best avoided altogether.

Admittedly, precise classification of an opinion is not always a straightforward or uncontentious matter. Some words or phrases do not fit snugly into a single category. There may be borderline cases, or opinions straddling more than one category, or genuine instances of ambiguity. It might be necessary, when reading an expert report, to make a critical judgment regarding what the expert intended to convey, and classify the opinion accordingly. But it is always worthwhile, when approaching the task of making sense of expert evidence, to begin by seeking to clarify the nature of the opinion the expert is giving, and the following checklists of words and phrases are tried-and-tested starting points.

5.3 (a) *Explanations*

Typical phraseology includes:

'... could have come from ...'

'... consistent with ...'

'... entirely consistent with ...'

'... possible sources include ...'

'... I cannot exclude the possibility that ...'

'... I cannot completely exclude the possibility that ...'

'... [some specified pertinent event] may have happened ...'

'...[some specified pertinent event] is a possibility ...'

'... the findings could be explained by ...'

'... possible explanations include ...'

A characteristic feature of explanations is that they are stated as bald propositions, with no attempt to indicate (much less quantify) how likely, or unlikely, they might be on the evidence.

5.4 (b) *Posterior probabilities (for explanations)*

Typical phraseology includes:

'... likely to have come from ...'

'... [some specified pertinent event] probably occurred ...'

'... the most probable cause is ...'

'... the observations are suggestive of ...'

'... the findings point towards ...'

'... the findings indicate ...'

Posterior probabilities come in different forms. Some formulations (e.g. those employing the language 'likely'/'likelihood' or 'probable'/'probability') are more obviously statements of posterior probability than other, more subtle variants. Irrespective of the precise language used to express them, posterior probabilities are inferential conclusions about contested facts in issue, formulated after the expert has taken account of pertinent scientific findings.

5.5 (c) *Categorical opinions*

Typical phraseology includes:

'... has come from ...'

'... [some specified pertinent event] has occurred ...'

'... the cause is ...'

'... the only explanation is ...'

'... the source is ...'

'...[some specified cause] is excluded as a source of the trace.'

'...[some specified pertinent event] cannot have happened.'

Categorical opinions (purport to) allow no room for residual uncertainty. They are bald, conclusory statements of fact. When stated in the negative (like the last two examples) they are also known as ***exclusions***.

5.6 (d) *Probabilities for the observations and likelihood ratios*

Typical phraseology includes:

- For *probabilities of observations* (where *x* and *y* refer to events, states of affairs, propositions, hypotheses, etc, according to context):

'... the results are likely to have been obtained if x had occurred ...'

'... the findings are very probable if y is true ...'

- For *likelihood ratios* (where *x* and *y* refer to events, states of affairs, propositions, hypotheses, etc, according to context):

'... the results are more likely to have been obtained if x rather than y were true ...'

'...the findings are approximately ten times more probable if x rather than y had occurred...'

'... the findings provide strong support for the view that x rather than y is true...'

The last example is an expression of the magnitude of the likelihood ratio, translated into a qualitative or 'fuzzy' verbal proxy ('strong support') taken from a scale of verbal equivalents for quantified LRs.

5.7 (e) *Indeterminate opinions*

There is a residual collection of phrases, sometimes encountered in forensic practice, that are difficult to understand and therefore hard to classify. Typical examples include:

'... there is an association between ...'

'... provides strong evidence of a link between ...'

'... no support for ...'

Such expressions are prone to mislead, and possibly always dubious. Whenever an expert uses any of these, or similar, phrases, it would be prudent to seek clarification of meaning. For example, what kind of 'link' or 'association' is intended? (Everything can be linked to *anything* under some imaginable description, as illustrated by the parlour game Six Degrees of Kevin Bacon!)

To say that evidence provides 'no support for' a proposition is, at best, very incomplete. Is this meant to be a posterior probability? Or does it imply a likelihood ratio smaller than one, in which case the evidence does support the contrary proposition – and the expert is only telling us half the story. Alternatively, and especially if encountered when the expert is operating in *investigative* mode, the opinion could be simply an expression of a single *probability for the observations*, i.e. the findings are very unlikely if the proposition were true. If so, this is how the opinion should, more accurately and transparently, be expressed – emphasising the assumed status of the conditioning assumptions.

5.8 ***Stage 2: Identify the Level in the Hierarchy of Issues***

The characterisation of the issue addressed, or purportedly addressed, by expert evidence is often the best indication of its place in the hierarchy of issues. This can sometimes be more or less ‘read off’ from the expert’s specification of the relevant proposition pair. On other occasions, it must be reconstructed by inference from the expert’s own looser formulations.

Briefly recapitulating on the four levels in the hierarchy:

- (a) Offence level: If expert evidence addresses the offence, then that self-evidently would be an *offence* level opinion (e.g. D raped V; D burgled the house, etc).
- (b) Activity level: If the expert evidence is concerned with particular conduct (whether or not on its own amounting to a criminal offence) by named person(s), this would be an *activity* level opinion (e.g. D penetrated V sexually; D smashed the window, etc).
- (c) Source level: If expert evidence is concerned with the physical origin or provenance of recovered evidential material or samples, or with establishing the identity of a particular person of interest (e.g. an offender or victim), this would be a *source* level opinion (e.g. Who was the donor of the semen recovered from V? What is the source of the fibres in combings from D’s head hair?).
- (d) Sub-source level: If expert evidence reports analytical results which cannot be directly associated with a particular physical medium, this would be a *sub-source* opinion. One example would be a DNA profile obtained by taking tapings from a garment, where the bodily fluid facilitating the transfer of genetic material to the garment cannot be ascertained.

5.9 ***Stage 3: Testing the Class and the Level of an Expert Opinion***

Having classified the particular kind of opinion being offered (step 1), and ascertained the level in the hierarchy of issues to which it is addressed (step 2), the third step in a rigorously logical approach to reading and fully comprehending expert evidence involves critical interrogation of these classifications.

More specifically, critical questions can be addressed to each dimension of the opinion. Different questions should be posed depending on whether the opinion, properly interpreted and classified, is (a) an explanation, (b) a posterior probability, (c) a categorical opinion, or (d) a probability for the observations, employing likelihood ratios. Likewise, tailored questions should be posed to explore the relevant level(s) in the hierarchy of issues to which the opinion is supposedly addressed.

5.10 Key questions for each of the four principal types of expert opinion would prominently include the following:

(a) *Explanations*

- Are there any other possible explanations for the findings (consistent with known background case information)?
- If so, have these been mentioned by the expert? If not, why not?
- Do the explanations offered betray any – conscious or unconscious - bias?

(b) *Posterior probabilities*

- Has the expert considered a full range of realistic hypotheses or propositions before assigning prior probabilities?
- Has the expert fully appreciated the impact of background information on the formulation of their opinion? Has the expert consciously guarded against the risks of (unconscious) bias in their choice of hypotheses or propositions to consider and in assigning prior probabilities?
- Has the expert expressly considered and assigned probabilities for the scientific findings or observations, taking account of those matters listed in (d), below? (Such assignments are in fact a logical precondition of generating posterior probabilities. The only question, in reality, is whether the expert did this knowingly and explicitly, or implicitly and without due consideration.)
- Has the expert clearly explained how a posterior probability for the explanation(s) was derived?
- Where a conditional has been transposed, was this deliberate and a legitimate application of Bayes' theorem? Or has the expert perpetrated a logical reasoning

fallacy (such as ‘the prosecutor’s fallacy’, described in *Practitioner Guides* Nos 1 and 2)?

(c) *Categorical opinions*

- Is the opinion a genuine logical deduction, a personal belief, or a manifestation of a zero probability for the findings assuming the truth of a proposition?
- If the opinion is deductive in form, are its premises robust and empirically justified?
- If the opinion is a belief, how does the expert justify holding and expressing it? What is the epistemic *warrant* for the belief?
- If the opinion is based on a zero probability for the occurrence of the observations, how does the expert justify that probability assignment? What is the epistemic basis for assigning that outcome a probability of zero (i.e. categorical impossibility)?

(d) *Probabilities for the observations and statements employing likelihood ratios*

- What is the statistical, empirical or logical basis of the expert’s assignments of probabilities?
- Are these probabilities based solely on the expert’s personal knowledge and/or experience? And/or on databases? And/or on further experimentation?
- If personal knowledge and experience were relied on, what is the nature, depth and breadth of that knowledge and experience? How reliable is the expert’s memory of their own professional experience? Is there documentary corroboration of personal recollections? (E.g. if an expert testifies to having made a particular observation or finding ‘only once or twice in thirty years of my clinical experience’, is it safe to take that statement literally and at face value?)
- If any databases were consulted, do they appropriately model a population relevant to the issue in question? Is the sample size sufficient to support a reliable estimate of relative frequencies?
- If further experimentation was undertaken, how many trials were performed? How closely did the experimental conditions replicate actual case facts and circumstances? Are any identifiable discrepancies between case-facts and experimental conditions significant?
- If multiple techniques were employed, has the expert taken into account any conditional dependency between the results produced by different techniques? Has the

logical independence of findings been robustly verified before employing the product rule to combine two or more conditional probabilities forming likelihood ratios for different pieces of evidence or analytical results?

5.11 Critical questions to consider in relation to each level in the hierarchy of issues would prominently include the following:

(a) *Offence level*

- Is it logically competent and legally appropriate for an expert witness to address this level of issue? Or are offence level issues strictly reserved to the fact-finder in the applicable legal jurisdiction? (Alternatively, may experts comment, provided that the fact-finder always has the last word?)
- Has the expert strayed into commenting on issues (e.g. normative questions of moral or legal culpability) beyond the scope of an expert witness's genuine expertise, knowledge or professional competence?

(b) *Activity level*

- Are the propositions formulated by the expert appropriate to the conduct (acts or omissions) constituting the offence?
- Have the expert's propositions been formulated in a way that fairly reflects the prosecution's allegations, and whatever might be known at the material time about the defence's contentions?
- Does the expert have adequate and sufficiently robust knowledge about transfer, persistence and detection of particular materials or substances to be able to make confident assignments of probabilities for the findings, allowing for relevant contextual features and circumstances of the instant case?

(c) *Source level*

- Why has the expert chosen to express an opinion (only) at source level? Why has scientific analysis and interpretation of results not advanced to the more legally useful *activity* level?
- Is this progression being blocked by inadequate data and understanding about transfer, persistence and detection of the transfer material?

- How, in the absence of expert interpretation of results at activity level, should source level findings be presented to clients or fact-finders? Is there serious risk that clients or juries will over-estimate (or, conversely, underestimate) the probative value of the expert's opinion? Can such risks be ameliorated by choice of presentation formats?
- Where the absence of 'matching' material precludes any source level evaluation, might an *activity* level evaluation provide further information useful to the client or fact-finder?
- Where an issue is properly addressed at *source* level, do the chosen proposition pairs fairly reflect the prosecution's allegations and whatever might then be known about the defence's contentions?
- Does the expert have sufficient, reliable data and knowledge both as to (i) 'within-sample' variability (typically influencing the numerator in a likelihood ratio); and (ii) 'between-sample' variability (typically influencing an LR's denominator)?

(d) *Sub-source level*

- Why has the expert chosen to express an opinion (only) at sub-source level? Why has scientific analysis and interpretation of results not advanced to the more legally useful *activity* level?
- Is there a problem attributing the analytical results to specific material (e.g. a particular bodily fluid)?
- How, in the absence of expert interpretation of results at activity level, should sub-source level findings be presented to clients or fact-finders? Is there serious risk that clients or juries will over-estimate (or, conversely, underestimate) the probative value of the expert's opinion? Can such risks be ameliorated by choice of presentation formats?
- On further questioning, is the expert able and willing to express a view on the value of his observations or scientific findings at *activity* level?
- Where an issue is properly addressed at *sub-source* level, do the chosen proposition pairs fairly reflect the prosecution's allegations and whatever might then be known about the defence's contentions?
- Does the expert have sufficient, reliable data and knowledge both as to (i) 'within-sample' variability (typically influencing the numerator in a likelihood ratio); and (ii) 'between-sample' variability (typically influencing an LR's denominator)?

5.12 ***Stage 4: Assessing Whether the Examination Strategy was Appropriate to Address Facts in Issue before a Court***

These questions relate specifically to contested trials at which an expert has expressed an opinion in the form of a likelihood ratio, either in a written report or in oral testimony.

- Would alternative forensic strategies or techniques employed on the same evidential items have produced (even) more informative likelihood ratios (irrespective of whether those LRs would have favoured the prosecution or defence)?
- Might forensic examination of additional, or different, evidential items have provided more informative likelihood ratios?
- In either event, why were these alternative tests or examinations not conducted?

Appendix A: CAI Aide-memoire

This aide-memoire, formulated as a series of questions designed to assist forensic practitioners to employ the Case Assessment and Interpretation (CAI) model in their own case work, is structured around the seven stages of the CAI model set out in Part 3 of this *Guide*; and draws on Jackson and Jones (2009).

Stage 1. Acquire information regarding all *relevant* circumstances of the case

- a. What are the relevant case circumstances, e.g. relevant timings, aspects of witness statements, allegations, medical and other specialist evidence?
- b. What are the issues – with what aspects of the case does the client need help?
- c. What is the potential charge?
- d. Is there a suspect? Has he/she been charged?
- e. What is the suspect's story, if any?
- f. What are the client's deadlines – statutory requirements; other needs?
- g. If the client is willing to share the information, and requires help in ensuring cost-effective strategies, what is the budget for the case?
- h. Has sufficient information to help identify the key issues in the case, and to inform probabilities of findings, been submitted? If not, make efforts to obtain it.

Stage 2. Clarify and define the client's requirements

Based on your understanding of the case circumstances:

- a. Are the issues investigative and/or evaluative?
- b. What level in the hierarchy of issues (sub-source/source/activity/offence) will you be addressing?
- c. What type of opinion (explanations, posterior probabilities, probabilities for the observations and likelihood ratios) will be justified?

Stage 3. Develop a proposal for a cost-effective examination strategy

- a. If the issue is investigative, develop list of possible explanations with, if possible, realistic prior probabilities. Considering the techniques available and the probabilities of the different outcomes of the techniques, decide which examinations on which items have the best chance of providing the most effective information to help direct the investigation.
- b. If the issue is evaluative, specify a relevant pair of propositions based on prosecution and defence positions. Develop likelihood ratio tables, or other depictions of probability, for the different items and techniques, previewing items if necessary.
- c. Identify the relative cost-effectiveness of the different examinations.

Stage 4. Consult client, explain potential outcomes, costs and timescales, and agree an examination strategy

- a. Contact client, explain potential outcomes and their respective probabilities, costs and timescales
- b. Address clients' expectations.
- c. Discuss options.
- d. Refine purpose of examination and agree the examination strategy.

Stage 5. Conduct the examination(s)

- a. Perform the examinations; make observations; obtain results of analytical techniques.
- b. Review the observations and/or the results of any analyses.

Stage 6. Interpret the scientific findings

- a. If operating in investigative mode, revisit original list of explanations.
- b. If providing posterior probabilities, incorporate the probabilities of the observations with the prior probabilities, and revise list in order of posterior probability.
- c. If not offering posterior probabilities, then simply revise list and ensure all possible, feasible explanations are given.

- d. If operating in evaluative mode, compare results with previous assessment and refine the likelihood ratio.

Stage 7. Communicate findings and explain their interpretation to the client

- a. Check your 'Conclusion' aligns with your 'Purpose' as agreed in the case strategy.
- b. Convey to the client your findings and conclusion in the most appropriate way.
- c. Ensure the basis and the strengths and limitations of your opinion are clearly set out and understood by the client.

Appendix B - Bibliography

ACPO and CPS (2013) *National Streamlined Forensic Reporting Guidance*, Version 2.0.

www.cps.gov.uk/legal/s_to_u/scientific_evidence/ [accessed 8 December 2014]

Association of Forensic Science Providers (2009) 'Standards for the Formulation of Evaluative Forensic Science Expert Opinion' 49 *Science and Justice* 161.

Budowle, B, Bottrell, M C, Bunch, S G, Fram, R, Harrison, D, Meagher, S, Oien, C T, Peterson, P E, Seiger, D P, Smith, M B, Smrz, M, Soltis G L and Stacey, R B (2009) 'A Perspective on Errors, Bias, and Interpretation in the Forensic Sciences and Direction for Continuing Advancement' 54 *Journal of Forensic Sciences* 798.

Champod, C, Evett, I W and Jackson, G (2004) 'Establishing the Most Appropriate Databases for Addressing Source Level Propositions' 44 *Science & Justice* 153.

Cook, R, Evett, I W, Jackson, G, Jones, P J and Lambert, J A (1998) 'A Model for Case Assessment and Interpretation' 38 *Science & Justice* 151.

Evett, I W (1983) 'What is the Probability that this Blood Came from that Person? A Meaningful Question?' 23 *Journal of the Forensic Science Society* 35.

Evett, I W (1991) 'Interpretation: A personal Odyssey' in C G G Aitken and D A Stoney (eds) *The Use of Statistics in Forensic Science*. Chichester: Ellis Horwood.

Evett, I W, Gill, P D, Jackson, G, Whitaker, J and Champod, C (2002) 'Interpreting Small Quantities of DNA: The Hierarchy of Propositions and the Use of Bayesian Networks' 47 *Journal of Forensic Sciences* 520.

Evett, I W, Jackson, G, Lambert, J A and McCrossan, S (2000) 'The Impact of the Principles of Evidence Interpretation on the Structure and Content of Statements' 40 *Science & Justice* 233.

Finkelstein, M O and Fairley, W B (1970) 'A Bayesian Approach to Identification Evidence' 83 *Harvard Law Review* 489.

Forensic Regulator (2014) *Draft Guidance: Cognitive Bias Effects Relevant to Forensic Science Examinations*. www.gov.uk/government/consultations/cognitive-bias-effects-relevant-to-forensic-science-examinations-draft-guidance [accessed 8 December 2014].

Haack, S (2009) 'Irreconcilable Differences? The Troubled Marriage of Science and Law' 72 *Law and Contemporary Problems* 1.

Ho, H L (2000) 'The Judicial Duty to Give Reasons' 20 *Legal Studies* 42.

House of Lords Select Committee on Science and Technology (1993) *Forensic Science* 5th Report of Session 1992-93, HL Paper 24 (Chairman: Lord Dainton). London: HMSO.

Jackson, G (2009) 'Understanding Forensic Science Opinions' in J Fraser and R Williams (eds), *Handbook of Forensic Science*. Cullompton, Devon: Willan Publishing.

Jackson, G (2013) 'The Impact of Commercialization on the Evaluation of DNA Evidence' 4 *Frontiers in Genetics* Article 227.

Jackson, G. and Jones, P J (2009) 'Case Assessment and Interpretation' in A Jamieson and A Moenssens (eds), *Wiley Encyclopedia of Forensic Science*. Chichester, UK: Wiley.

Jackson, G, Jones, S, Booth, G, Champod, C and Evett, I W (2006) 'The Nature of Forensic Science Opinion - A Possible Framework to Guide Thinking and Practice in Investigations and in Court Proceedings' 46 *Science and Justice* 33.

Jones, C G (1986) 'Men of Science v. Men of Law: Some Comments on Recent Cases' 26 *Medicine, Science and the Law* 13.

JUSTICE (1991) (Chair: HHJ Christopher Oddie), *Science and the Administration of Justice*. London: JUSTICE.

The Hon Fred Kaufman, CM, QC (1998) *Report of the Kaufman Commission on Proceedings Involving Guy Paul Morin*. Ontario: Ministry of the Attorney General.

www.attorneygeneral.jus.gov.on.ca/english/about/pubs/morin/ [Accessed 8 December 2014].

Lawless, C J (2011) 'Policing Markets: The Contested Shaping of Neo-Liberal Forensic Science' 51 *British Journal of Criminology* 671.

Lempert, R O (1977) 'Analyzing Relevance' 75 *Michigan Law Review* 1021.

Lindley, D V (1977) 'A Problem in Forensic Science' 64 *Biometrika* 207.

National Research Council (2009) *Strengthening Forensic Science in the United States: A Path Forward*. Washington, DC: National Academies Press.

Oesterhelweg, L, Kröber, S, Rottmann, K, Willhöft, J, Braun, C, Thies, N, Püschel, K, Silkenath, J and Gehl, A (2008) 'Cadaver Dogs—A Study on Detection of Contaminated Carpet Squares' 174 *Forensic Science International* 35.

Ramsay, R (1987) *The Effectiveness of the Forensic Science Service*, Home Office Research Study No.92. London: HMSO.

Redmayne, M (1998) 'Bayesianism and Proof' in M Freeman and H Reece (eds), *Science in Court*. Aldershot: Ashgate.

Risinger, D M, Saks, M J, Thompson W C and Rosenthal, R (2002) 'The *Daubert/Kumho* Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion' 90 *California Law Review* 1.

Roberts, P (1996) 'What Price a Free Market in Forensic Science Services? The Organization and Regulation of Science in the Criminal Process' 36 *British Journal of Criminology* 37.

Roberts, P (2013) 'Renegotiating Forensic Cultures: Between Law, Science and Criminal Justice' 44 *Studies in the History and Philosophy of Biological and Biomedical Sciences* 47.

Roberts, P (2014) (ed), *Expert Evidence and Scientific Proof in Criminal Trials* (Ashgate, 2014).

Roberts, P, Aitken, C, Jackson, J and Redmayne, M (2011) 'When the Numbers Count' 175 *Criminal Law & Justice Weekly* 304 (21 May).

Roberts, P and Zuckerman, A (2010) *Criminal Evidence*, 2nd edn. Oxford: OUP.

Royal Commission on Criminal Justice (1993) *Report* Cm 2263 (Chair: Viscount Runciman). London: HMSO.

Saks, M J and Koehler, J J (2008) 'The Individualization Fallacy in Forensic Science Evidence' 61 *Vanderbilt Law Review* 199.

Schum, D A (2001) 'Species of Abductive Reasoning in Fact Investigation in Law' 22 *Cardozo Law Review* 1645.

Tillers, P and Schum, D (1991) 'A Theory of Preliminary Fact Investigation' 24 *UC Davis Law Review* 931.

Touche Ross (1987) *Review of Scientific Support for the Police*. London: Home Office.

Walton, D (1992) 'Nonfallacious Arguments From Ignorance' 29 *American Philosophical Quarterly* 381.

Wonder, A K Y (1989) 'Science and Law, A Marriage of Opposites' 29 *Journal of the Forensic Science Society* 75.

