The Young grandfather: a new approach to the reliability of expert evidence in Scottish criminal cases or business as usual?

Donald Nicolson* Rhonda Wheate**

Journal Article - Submitted to the Juridical Review

Subject

Criminal evidence

Other related subjects

Science

Keywords

Admissibility; Comparative law; Expert evidence; Forensic evidence; Reliability; Scientific methodology; Scotland; United States

Cases cited

Young (Thomas Ross) v HM Advocate [2013] HCJAC 145; 2014 S.L.T. 21; [2013] 11 WLUK 444 (HCJ)

Daubert v Merrell Dow Pharmaceuticals Inc 509 U.S. 579 (1993) (Sup Ct (US)) Frye v United States 297 F. 1013 (1923) (DC Cir (US))

Kennedy v Cordia (Services) LLP [2016] UKSC 6; [2016] 1 W.L.R. 597; [2016] 2 WLUK 287 (SC (SC))

R. v Bonython (1984) 38 S.A.S.R. 45

Scientific1 and other experts have played a useful role in legal proceedings since at least as far back as Archimedes' "Eureka" moment,2 but increasingly so with the Enlightenment's scientific discoveries, the explosive development of new technology during the Industrial Revolution and the scientific development of medicine and later psychiatry.3 Indeed, certain disciplines were developed specifically—and in some cases exclusively—to serve legal needs. But it was from the nineteenth century that there increasingly developed "armies of experts in medicine, dentistry, pathology, psychiatry, psychology, fingerprinting, toxicology, biological sciences (blood, hair, and bodily fluids analysis) genetics, ballistics, narcotics, trace mark examination (paint, glass, fibre, toolmarks, and footprint inspection), and document and handwriting analysis, to mention only some of the most common expertises".4

For the period 2018–19, the Scottish Legal Aid Board reports that over £4.2 million was spent on outlays including expert evidence and court reports in criminal proceedings,5 and while there are no figures for the incidence of expert testimony in Scotland, an English and Welsh study in the early 1990s found that around one-third of all contested trials on indictment involved scientific evidence.6

These figures in fact underplay the actual use of scientific expertise in criminal proceedings, given that "most experts make their most regular and significant contributions prior to the trial".7 This is most marked in relation to the constantly improving and ever-expanding range of forensic disciplines which have become a central part of investigating and proving crimes and other legally relevant acts.8

This "creeping scientisation" 9 of legal fact-finding is perhaps not surprising given that scientific and other forms of specialist knowledge and technology are now so embedded in modern life that legal actors often have little option but to rely on experts to make sense of our brave new world. The law has, however, also long utilised a variety of non-scientific experts like art historians who rely on non-scientific knowledge, engineers who rely on technical know-how, health and safety officers who received specialised training and police officers who gain relevant knowledge though experience. Moreover, the law is also gradually opening its door to the relatively new disciplines which focus on understanding human behaviour—often pejoratively, and unfairly10 called the "soft" sciences. But here the law has been far more cautious, not least because it has always maintained—albeit not consistently11 —that, in the words of an English judge, fact- finders do not need "psychiatrists to tell them how ordinary folk who are not suffering from any mental illness are likely to react to the stresses and strains of life".12

Yet, as this article will show, there are also grounds for being wary about the admissibility of expertise from the natural or physical ("hard") sciences. This is especially so because the "mystic Infallibility”13 of science may lead to fact-finders being so dazzled by scientific evidence that they automatically defer (or at least give undue prominence) to expert testimony, especially if it is bathed in the "aura of precision"14 associated with statistical information and because fact-finders are conditioned to expect forensic evidence in criminal cases by the so-called CSI effect. 15 Thus, while scientific evidence is undoubtedly an incredibly useful means of ascertaining and proving facts, and uncovering miscarriages of justice, erroneous or misleading scientific evidence is also a significant cause of wrongful convictions16 —indeed, second only to mistaken identification, according to US studies.17 Whether the problem is as acute in Scotland is impossible to gauge given the lack of relevant empirical research, but the recent case of Shirley McKie,18 a police officer whose perjury conviction was based on mistaken fingerprint evidence, shows that such miscarriages do occur in Scotland. This case led to the Fingerprint Inquiry,19 which made a number of important recommendations on the use of fingerprint evidence, but without raising its potential inadmissibility due to unreliability.

This non-interventionist approach to admissibility was consistent with then prevailing Scots law,20 which, like other common law jurisdictions, provides experts (traditionally called "skilled witnesses") with special legal status. Most notably,21 experts are exempt from the normal prohibitions on witnesses expressing opinions on (rather than just reporting their observation of) facts, and relying on hearsay evidence (here, the reports of others encountered in studying and practising their discipline). Despite these privileges and the likely influential impact of experts, Scots law traditionally made no attempt to ensure that expert evidence was sufficiently reliable to justify its special status. Instead, it required only that it be relevant, not usurp the fact-finders’ authority to decide the "ultimate issue", necessary for a decision to be made, and that it came from an impartial witness with the relevant qualifications or experience to assist the court. The courts were not totally oblivious to the importance of the reliability of expert evidence,22 but, until recently, there was no indication that reliability was a matter of admissibility rather than merely affecting the weight of expert evidence. Instead, the approach was to admit the evidence subject to testing through cross-examination and, if necessary, warnings in jury directions, with the backstop of possible correction on appeal.

Then, in 2013, in Young v HM Advocate,23 the High Court of Justiciary (HCJ) was faced with the relatively recent and still controversial forensic psychology technique of case linkage analysis (CLA), which purports to show that different crimes were likely to have been committed by the same person because of the unique characteristics of the crimes and perpetrator (often called "offender profiling").24 The defence argued that the appellant could not have committed a particular murder because CLA had identified it as one of a number of crimes committed by someone else. The HCJ, however, accepted the Crown's submission that CLA was inadmissible as not being "sufficiently reliable", and went on to set out a five-pronged test of "sufficient reliability", which requires that expert evidence"[1] must be based on a recognised and developed academic discipline. [2] It must proceed on theories which have been tested (both by academic review and in practice) and [3] found to have a practical and measurable consequence in real life. [4] It must follow a developed methodology which is explicable and open to possible challenge; and [5] it must produce a result capable of being assessed and given more or less weight in light of all the evidence before the finder of fact".25 However, in addition to introducing this test, Lord Menzies also declared—obiter and without supporting argument—that "scientific evidence about DNA comparisons, fingerprint evidence, evidence of medical practitioners or pathologists is evidence based on a sufficiently clear and reliable basis that it may assist the finder of fact, and will be admitted", thus appearing to "grandfather" these types of expert evidence.26 Technically grandfathering occurs when existing practices are exempted (by "grandfather clauses") from being invalidated under new legislative schemes. Cole, however, has described the US courts as engaging in de facto grandfathering by upholding the admissibility of myriad types of expertise of dubious validity that appears to conflict with modern legal tests of expert reliability.27

The Young decision was delivered without any reference to these tests or the vast related literature in other jurisdictions. Indeed, only one case was cited by counsel: the influential and much discussed US precedent of Daubert v Merrell Dow Pharmaceuticals Inc.28 Here, the Supreme Court replaced the 70 year old precedent of Frye v United States 29 which required only that any principle relied on by the expert "be sufficiently established to have gained general acceptance in the particular field in which it belongs". This "general acceptance" test has been called an external or extrinsic,30 or deference,31 approach in delegating responsibility for evaluating reliability to legal outsiders. As such, it has been criticised for allowing the courts to avoid their responsibility to ensure that only sufficiently reliable expert evidence is admitted and, as we shall see,32 for being poorly drafted. Although purporting to be more liberal, the Daubert test is potentially also more rigorous.33 Thus, while retaining general acceptance as "having some bearing on the inquiry", the Supreme Court in Daubert also stated that courts should inquire as to whether "the theory or technique can be (and has been) tested" and has "been subjected to peer review and publication", the technique's "known or potential rate of error", and the "existence and maintenance of standards controlling its operation".34

Whereas the Young decision resembles the internal gatekeeping approach of Daubert far more than the deference approach of Frye, the HCJ's failure to discuss any other decisions renders it impossible to know whether or not the similarities and differences with Daubert were deliberate. Nor have subsequent cases brought clarity. The only Scottish criminal case to cite Young is Procurator Fiscal, Alloa v C,35 where Sheriff McGowan relied on the first requirement in Young to exclude evidence from a police officer about the likely use of drugs found in the accused's possession. However, he also found helpful the guidance on the admissibility test produced by the Law Commission of England and Wales in its Report on Expert Evidence in Criminal Proceedings in England and Wales,36 which itself was heavily influenced by Daubert.37

The only other reference to Young has been the Supreme Court civil law case of Kennedy v Cordia

(Services) LLP,38 which has subsequently been cited in criminal cases as the authority on the conditions for admissibility of expert evidence in Scotland, albeit not in cases where reliability has been at stake.39 This is curious, given that, as the Supreme Court itself acknowledged, it had no jurisdiction to rule on criminal law and only cited criminal cases as "context for our consideration of the law of evidence in civil cases".40 In fact, it placed more reliance on English cases, which in turn also "somewhat curiously"41 seem very taken with the South Australian case of R. v Bonython,42 which specifies that to be admissible the subject matter of expert opinions must form "part of a body of knowledge or experience which is sufficiently organised or recognised to be accepted as a reliable body of knowledge or experience".

Because the Bonython test was adopted in Kennedy 43 and it correlates only with the first element of the Young test, there now seem to be different Scottish rules on the reliability of expert evidence for criminal and civil cases. Given, however, that the Kennedy dicta on reliability were obiter (the issue being whether the expert evidence assisted the court rather than whether it was reliable), the HCJ in Young commented on a wider range of types of expert evidence, adopted a more uniquely Scottish approach, and the consequences of relying on unreliable expert evidence will usually be more serious in criminal cases, this article will focus primarily on Young and criminal cases, though using Kennedy where relevant as a comparative foil, along with the approach adopted in other Anglo-American jurisdictions.

More specifically, we will evaluate whether the Young test provides a suitably rigorous means of ensuring the reliability of expert evidence. This will be done by analysing its ambit, application and the meaning of each of its five elements. We will then turn to interrogate the implicit assumption that having a rigorous reliability test is likely to improve the quality of expert evidence offered in court. We will argue that there are at least two reasons for pessimism in this regard. The first is that the High Court of Justiciary (HCJ) itself was quick to "grandfather" DNA comparisons, fingerprint evidence, evidence of medical practitioners and that of pathologists. Yet, as we shall show, their reliability is hardly beyond doubt. It is, of course, possible that the courts will in fact assess these and other forms of scientific evidence far more rigorously than suggested by the HCJ's obiter comment. However, as a second reason for not expecting Young to make much difference to the admission of invalid and unreliable expert evidence, we will show that the record in other Anglo-American jurisdictions suggests that admissibility practices do not differ much according to the rigour or indeed existence of an admissibility test scrutinising the validity and reliability of the scientific evidence.

By exploring the reasons for this judicial complacency, we will argue that the Young test may not dramatically improve the validity and reliability of scientific evidence in Scottish criminal cases. Consequently, we will explore various alternative means of ensuring that they are better evaluated. Nevertheless, we will also argue that it remains important to ensure the most effective admissibility test possible, not least in order to provide legal counsel with the tools to challenge the various problems with expert evidence, and hopefully thereby force problematic disciplines to put their house in order. Consequently, we begin the article with a detailed and comparative evaluation of the Young test, starting with the general questions of what it covers and how it applies.

II. The ambit and application of the Young test

As regards the Young test's ambit, given that it was developed in the context of forensic psychology, it clearly regulates social science disciplines. And, notwithstanding the free passes given to the existing physical (albeit applied) sciences of DNA profiling and fingerprinting, and the "squishy"44 sciences of medicine and pathology, one can assume that the HCJ intended it to apply to physical and medical sciences, not least because its first element refers to "recognised and developed academic discipline[s]".

What is less clear is whether Young also applies to what is often called "experience-based" expertise.45 Thus, the Scottish courts have long allowed testimony which is based, not on formal education or training, but on having sufficient experience with the relevant phenomena.46 Thus, police officers have testified as to typical drug consumption by drug users,47 and even more dubiously, whether pornographic photographs involved girls under the age of 16.48 Moreover, as we shall see in section V below, many methods adopted by forensic scientists are largely or even purely based on little more than their experience. Even the testimony of medical practitioners and pathologists may rely as much on their experience as their formal training.

Because Young did not involve or mention experience-based expertise, it is unclear whether such expertise is completely exempt from its test, or conversely, whether it applies in its full rigour, or in some modified form. Prima facie, the rigorous option seems to have been taken in Procurator Fiscal, Alloa, where the police officer's evidence (based on his research into the practice of drug dealers) was excluded as not involving "a recognised and developed academic discipline", not gathered in "anything approaching a scientific way" or being "of sufficient quantity to provide a statistically valid opinion".49 In fact, however, it was unnecessary to rely on Young (or scientific unreliability more generally) because arguably the police officer lacked sufficient experience to count as an expert.50 More generally, if applied literally, the Young test seems to exclude all forms of knowledge which are unlikely ever to develop into an academic discipline, nor be subjected to its testing requirements, irrespective of how reliable they may be. This raises the issue of how the Young test is meant to be applied. Its wording suggests that each and every requirement must be satisfied, whether or not this is possible or necessary.

By contrast, the US Supreme Court in Kumho Tire Co v Carmichael,51 confirmed that the four elements of the Daubert test are not fixed or determinate52 and should be applied flexibly as only indications —rather than mandatory requirements—of reliability. Moreover, they apply, where appropriate, to "technical" and "other specialised knowledge",53 which includes experience- based knowledge. Similarly, in Kennedy, the UK Supreme Court stated that "[w]hat amounts to a reliable body of knowledge or experience depends on the subject matter of the proposed skilled evidence".54 It went on to state that where expert evidence does not derive from a recognised scientific discipline, the party adducing it "would need to set up by investigation and evidence … the methodology and validity of that field of knowledge or science".55 This flexible approach was implicitly followed in Jones v HM Advocate 56 where the HCJ referred to Kennedy and allowed opinion evidence by a police officer in relation to drugs offences after investigating the extent of his experience. Young, however, was not cited. Nor was the reliability of the evidence raised, still less the issue of scientific versus experience-based expert evidence.

Even more relaxed was the Law Commission of England and Wales, which accepted that for experts like accountants, lip readers, and health and safety and police officers, relevant qualifications, experience, skill and/or reliance on well-accepted practices would obviate the need for reliability tests.57 Indeed, it expressly reversed its initial suggestion of different tests for scientific and experience-based evidence on the grounds that there is a continuum, rather than a clear dividing line, between the two, and that many experts rely on a mixture of both.58 Equally, it made clear that in applying its proposed test of reliability courts should look only at those factors relevant to the particular type of evidence in question.59

A flexible application of reliability "indicia"60 to all forms of purported expertise seems far preferable to Young 's apparent approach of requiring fulfilment of each element, which would exclude many areas of sufficiently reliable expertise that may never become academic disciplines or ever be capable of fulfilling the other elements of Young, including—as we shall see in section V

below—much expert evidence traditionally admitted. Conversely, exempting experience-based expertise from the Young admissibility test altogether would undermine the goal of enhancing reliability and may encourage experts to attempt to claim that their expertise is based on experience rather than scientific research, in order to avoid having to establish its reliability.

Instead, we think a reliability test should exclude all forms of expertise—scientific as well as non-scientific, based on formal study or on experience—if the theories and methods underlying that expertise have not been substantiated in the most rigorous way possible.61 What this means in practice will be explored when we turn to the relevant elements of Young. For now, however, it can be noted that this flexible "best evidential support" does not entail that alternative indicia of reliability can be satisfied by the mere ipse dixit of purported experts62 (even if it reflects a consensus within the relevant field), not least because such opinions are difficult to challenge.63 Relevant to such a flexible approach would be whether the expert testimony involves an opinion (such as forensic scientists stating that in their view trace evidence found at a crime scene matched samples from the accused) or merely conveys facts (such as engineers explaining how a piece of machinery operates). This distinction was not raised in either Young or Procurator Fiscal, Alloa. However, in Kennedy the Supreme Court stated that the "special rules that govern the admissibly of expert opinion evidence also cover such evidence of fact",64 which seems to suggest that it must also be reliable. In contrast, the Law Commission opined that expert evidence of fact is unlikely to raise issues of reliability and would make a reliability test more difficult to formulate.65
A better reason for distinguishing testimony of opinion and fact is that, while as noted in the "Introduction" above, expert evidence, especially of a scientific nature, can be held in unjustifiably high esteem by fact-finders, opinion evidence impinges far more on their jurisdiction. In the case of evidence of fact, fact-finders still need to draw their own inference from facts testified to by the expert and hence it is only the hearsay rule, and not also the opinion evidence rule, which is breached—and then only if the expert relies on information provided by others (such as through formal training or in published studies). Moreover, experience-based testimony of fact never involves hearsay as it relies on the expert's own observations, and hence can be interrogated for reliability just like any other non-expert testimony. Indeed, it is arguable that such testimony is not in fact expert evidence and hence that a reliability test is unnecessary.

If this approach is followed, however, and courts evaluate only the reliability of the science behind experts' opinions and reported statements of fact, they should also follow the English and Welsh Law Commission suggestion that expert evidence should be treated as involving an opinion whenever it is unclear that testimony is not confined to statements of fact.66 In fact, this is likely to occur frequently because there is no clear distinction between fact and opinion.67 This is not merely to make the epistemological point that all statements of fact are merely someone's opinion about what they think the facts are,68 but also to recognise that it is often difficult in practice to decide when someone is simply stating what they know, or actually drawing inferences from such knowledge. Moreover, often the distinction between fact and opinion will be effectively meaningless. For instance, if police officers state that particular gang members use particular hand signals to denote their gang membership and the court is then shown a photograph of someone making those signals,69 it makes little difference whether matters are left there or the officers go on to state their opinion that the signals in the photograph denote membership of that gang. The effect on the fact-finder is likely to be the same.

III. Desperately seeking sense: analysing the Young elements

1. Young and the language of science

Having explored the general questions of how far the Young test extends and how it might be applied, we turn to analyse the details of its requirements for expert evidence to qualify as "sufficiently reliable". For scientists, this involves a number of different inquiries.70 The first is

whether the theory, principle or other knowledge claim (henceforth, theory for short) accurately reflects the phenomenon it describes: what is called "validity". Often, however, and especially as regards forensic evidence, courts are not interested in theories of how the world works, but in some method, technique or procedure (henceforth "method") which applies the theory to relevant phenomena to produce an opinion or information relevant to the facts in dispute. Such method must also be valid in terms of being capable of accurately doing what it is claimed to do. By contrast, in scientific terms "reliability" usually71 refers not to theories but to methods, and involves the question of whether a method consistently produces accurate results irrespective of who is using it. Both validity and reliability in this sense thus involve the question of whether, in general, the theory and method are sufficiently supported to warrant being relied upon. And it is important that both aspects of what (following Daubert),72 we can call "evidentiary reliability" are established. A method is useless if it consistently delivers results, but does so on the basis of an invalid theory; equally a theory might be valid, but it would be dangerous to rely upon it if the method used to apply it does not reliably deliver accurate results.

Problems may nevertheless still arise, however, because valid and reliable theories and methods are not appropriately applied in the specific case at hand. Thus, the chosen theory and method must also be appropriate for the specific issues at stake; what in Daubert was called the question of "fit",73 but which was not raised by the facts or mentioned in Young. Given that any fact or opinion delivered by way of an inappropriate theory or method is irrelevant, and of no help to the fact-finder, the question of fit, like that of reliability, is a question of admissibility.74

Finally, problems of proficiency arise when a valid, reliable and appropriate method is applied incorrectly to produce wrong results. To mention just the more obvious causes, humans can make isolated errors, or be unqualified or insufficiently skilled in conducting techniques, equipment might be faulty, and procedures designed to ensure accuracy might be flawed or ignored. In relating to the results and not the validity, reliability or fit of the methods, however, proficiency problems are usually regarded as affecting the weight and not the admissibility of expert evidence75 and therefore, as matters for the fact-finder, not a judicial gatekeeper, are rightly excluded from the Young test.

On other hand, as we shall see, the HCJ failed to distinguish validity from reliability strictu sensu. It also confusingly distinguished between scientific "theory" and "methodology", notwithstanding that in scientific terms "methodology" usually refers to a combination of a particular theory and methods derived from it, rather than to the methods alone.76 Nor did the court make clear how its conclusions about CLA were related to the various elements of the new test. Consequently, its treatment of the various aspects of the validity and reliability of expert evidence was patchy, confusing and at times incomprehensible, as we shall now see.

2. Young and maturity

In fact, the first element of the Young test refers to neither the validity and reliability of theories or methods, nor even to methodologies, but requires expert evidence to be based on "recognised and developed" academic disciplines more generally. Although the HCJ may have seen this as an obviously plausible means to exclude novel, undeveloped and controversial methodologies like CLA, it does not in fact do so, since CLA is a methodology which falls within the recognised and developed academic discipline of forensic psychology. The same problem besets Bonython ’s "sufficiently organised body of knowledge or experience" test77 (though not Frye ’s78 similarly motivated general acceptance test which applies to the relevant scientific principle and not its originating discipline as a whole).

Moreover, all three tests are plagued by definitional uncertainties.79 Focusing on Young, one can ask what is meant by a "recognised" "discipline". How is recognition evidenced: in academic journals or university courses? It cannot be by the ipse dixit of the very expert whose reliability is in question, since this would allow the claims of a few partisan supporters working in narrow fields to be self-validating, especially if they can organise themselves professionally to exclude dissenting voices.80 But even if recognition must be more widespread, how widespread? Take the example of CLA: does one look to forensic psychology, psychology as a whole, or even social science in general? And what about methodologies which are built on more than one discipline, such as DNA profiling, which involves both biology and statistical mathematics. Is recognition required in one, some, or all underlying disciplines? Certainly, the ongoing debate in DNA profiling between biologists and statisticians over how a "match" should be reported for legal purposes has been fierce.81 Further questions in relation to what counts as recognition flow from the fact that in large disciplines it would be impractical to require complete consensus—there will always be mavericks. Indeed, as Galileo famously (and literally) illustrated, sometimes today's heresy becomes tomorrow's orthodoxy (and vice versa).82 But what level of dissent is compatible with "recognition"? Depending on its meaning, the requirement of being "developed" seems more useful than a recognition test which might just degenerate into simply "counting noses".83 But does it simply require a sufficiently lengthy existence; in which case, CLA's 30 year pedigree was longer than that of DNA when it was first accepted (prematurely as it turned out).84 Or, more substantively, must the theories and methods have been tested and re-tested, in which case, the first element of Young does not add much to a properly formulated requirement of testing.

In any event, irrespective of its exact meaning, this element is simultaneously too demanding and too lenient. Thus a mandatory requirement that expertise is part of a recognised and developed discipline will exclude evidence based on technical training and experience even though they might be more reliable than those taught at university,85 but simultaneously validate disciplines formally taught for commercial reasons or even in order to subject them to critical scrutiny. Equally, the requirement of maturity might exclude useful novel discoveries too new or radical to garner recognition, or, conversely, cocoon long-standing methodologies from challenges posed by new discoveries. Indeed, recognition (or general acceptance, etc.) can be gained and maintained by processes having little to do with inherent reliability.86 In reality, there is simply insufficient time, resources or prestige to be gained from attempting to test all scientific findings. At some stage (sometimes, almost immediately), the scientific community will accept that there has been sufficiently adequate testing to treat theories as established (at least provisionally, until new evidence emerges). Deciding when this point has been reached involves tacit understanding and negotiation within the relevant scientific community as to what constitutes effective testing methods and criteria, and whether tests confirm or falsify relevant hypotheses. The decision may also be influenced by intrinsic factors like the theory's explanatory power, simplicity or elegance, and extrinsic factors like the reputation of the theories' proponents, funders' interests and the theories' political or moral acceptability.

This is not to assert that recognition, general acceptance and the maturity of claims to knowledge and effective methods have no value as indicators of reliability. Courts should indeed be wary of theories and methods which have not stood the test of time, particularly if controversial in their field.87 Nevertheless, it is generally recognised,88 not only that theories or methods initially rejected by mainstream science can ultimately become received knowledge,89 but also more generally that recognition, development, organisation and general acceptance are only proxies for actual scientific validity and reliability, and that the most important question is whether such evidentiary reliability has been established through relevant forms of testing. We therefore turn to how the Young test dealt with this inquiry.

3. Testing, testing, testing

Of the Young elements, the second and third refer only to the theory on which expert evidence proceeds and the fourth and fifth only to the methodology used to produce the evidence offered to court. Because of this distinction, and because only the second and third elements deal with the "key question"90 of testing, we will discuss them together, before analysing the final two. Cumulatively, the testing elements require that "expert evidence must proceed on theories which have been tested (both by academic review and in practice) and found to have a practical and measurable consequence in real life".91

More specifically, the first of these elements requires two types of testing. One is "academic review" which can precede a funding application or publication in a peer-reviewed journal. But like general acceptance, academic review involves a considerable degree of deference to discipline insiders, and only indirectly evaluates the validity and reliability of expert evidence. Moreover, the value of "peer review" is limited, not least because it can be affected by informal factors like institutional affiliation and established reputations.92 Indeed, as the Daubert court recognised,93 failing this element should not necessarily preclude admissibility, since reliable theories might not have yet made it through the sometimes lengthy peer review processes, their value might not yet be appreciated, or their subject matter, while valuable, might not be of sufficient interest to a wide enough readership to warrant publication in peer-reviewed journals.94

More importantly, for practical reasons, peer review does not mean that the journal, reviewers, or readers have replicated the published experiment or conducted other forms of testing, and hence publication in a peer-reviewed journal by no means guarantees the reliability of research findings. What peer review does is create a public record of research and give others in the field the opportunity to challenge findings either by replicating the research or by scrutinising it from the perspective of their own research and knowledge.95 In this way, despite not providing a guarantee of reliability, peer review nevertheless contributes to the important scientific criterion of falsifiability, discussed further below. Moreover, rejection by reviewers or, worse still, the failure to even submit research to a peer-reviewed publication, provides useful clues as to its possible lack of evidentiary reliability.96 And here the most glaring problem with Young is that its direct testing requirements ("tested … in practice") appear to relate only to the validity of theories on which the evidence proceeds and not to the validity or reliability of the methods by which it is obtained. Further, it is unclear as to whether external validity, internal validity, or both, are required in relation to theories, or methods, or both.

Thus, being very charitable to the HCJ, perhaps the requirement that a theory has "a practical and measurable consequence in real life" was intended to refer to valid methods emanating from the theory. But even so—like the reference to testing "in practice" and the court's criticism of CLA research as being "concerned only with closed, or solved, crime" and not safely applicable to "predictions, or unsolved crimes"—this requirement only relates to what is called "ecological validity", which is itself only one aspect of "external validity". The latter requires that research findings are not confined to the specific phenomena tested (the actual fingerprints used, actual crimes committed, etc.), but are generalisable to all relevant phenomena (all fingerprints, all committed crimes). More specifically, ecological validity requires that results are not confined to controlled conditions, such as in laboratories, but also apply in real life conditions (such as crime scenes).

However, ecological validity is often bought at the expense of what is called "internal validity", which involves controlling study variables in order to ensure a causal connection between the factor studied and the results obtained, such as by excluding rival hypotheses that provide plausible alternative explanations for the results.97 In addition, particularly in tests regarding human subjects,

steps need to be taken to recognise and minimise the behaviour of experimenters and subjects which affects the outcome of tests. Most notably, random selection of test subjects ensures that results are not biased by the experimenter's choice of subjects and "double masking" prevents both experimenter and subject subconsciously or even consciously altering their behaviour to ensure the expected results and to avoid the placebo effect, in terms of which subjects may subconsciously respond to some ineffective treatment deliberately administered in order to compare with the treatment being tested.98

Although a concern with internal validity was not explicitly raised by the HCJ, it was implied by its criticisms that most studies of CLA do not take account of "the effect of a victim's reaction and behaviour on the behaviour of the offender" and that "[i]nter-rater reliability remains a real issue".99 In fact, this latter comment relates to methods and not underlying theory. But if the HCJ intended to require testing of both forms of validity and reliability, it did not make this clear, notwithstanding being exposed to at least the concept of "ecological validity".100 It also referenced the distinction between foundational theories and methods, in noting the Crown experts' criticisms of the lack of "empirical scrutiny" for the "two theories underpinning the practice of case linkage".101 But, given their limited scientific training and time available for the judges as well as counsel to grasp unfamiliar background concepts, it is not surprising that they displayed little understanding of the important distinctions between theory and methods, reliability and validity, and internal and external validity.

The third requirement of the Young test, that theories have been "found to have a practical and measurable consequence in real life", is similarly problematic. It might refer to a need for ecological validity. Alternatively, it might have been intended to emphasise that courts are interested only in theories that have real-life consequences, even though theories usually explain, rather than have consequences for, real-life phenomena. It is the results of applying the methods to which theories give rise, however, which are most likely to feature in court and only rarely the theories themselves, and it is thus methods which require ecological validity.

By contrast, the HCJ's justifiably102 most pressing103 criticism of CLA's "known high error rate" does implicitly make clear that it is not just testing per se which is important, but testing which actually actively examines the validity and reliability of theories and methods. Error rates quantify the frequency with which errors occur in using a particular method.104 This is calculated by comparing the number of times it produces wrong results with the number of times it is right. Scientists do not expect a theory, method, its application, the equipment used or the skills of relevant personnel all to be perfect. In reality, the process of arriving at results may well be affected by errors of one type or another. These may involve random errors (for example, a stopwatch is started too early or too late); known and hence correctable systematic errors (for instance, a thermometer is systematically reading at 2°C too high); unknown and hence irremediable systematic errors; or even "blunders" (for example, writing down the wrong number). Although repeating a method many times may help to identify systematic errors and blunders, and help random errors to cancel one another out, the results may nevertheless be clearly wrong sometimes, or may not be exactly the same every single time, without necessarily invalidating the method.

This, however, still leaves the question as to what level of error rate is compatible with a valid and reliable method. It was easy to describe CLA's error rate of 25 per cent as being unacceptably high, but there is no standard accepted error rate in science generally. Much depends on the context and the type of error. Thus, the consequence of false positives (commonly called type I errors) may be very different to false negatives (type II errors). For instance, a higher error rate for false positives is far more acceptable when investigating a case because these errors can be corrected by further

investigation, whereas type I errors are unacceptable at trial because of the risk of wrongful conviction. On the other hand, false negatives are problematic pre-trial because they might close down appropriate lines of inquiry, whereas at trial they may lead to mistaken acquittals which are generally regarded as less egregious than wrongful convictions. This still does not help, however, in identifying a rule specifying an acceptable error rate. Here, perhaps all the court can do is draw on, and subject to critical examination, views from the relevant scientific community as to what error rates are acceptable in their domain.

Another issue is whether "measurable consequences" in the Young test require that the results of testing the theory are measurable in numbers (as suggested by the emphasis on the CLA's error rate). On the other hand, no other jurisdiction has such a requirement. Even if "measurable" simply means that tests can be judged to be successful or not, there needs to be a benchmark against which to measure such testing. For reliability, one simply needs sufficient repeated applications of the method in question in order to confirm its consistency. For validity, however, the benchmark is whether the theory does in fact reflect how things actually are; in other words, true facts. But, even if accessing "true" facts was epistemologically possible,105 it is not always practically possible outside of controlled experiments where such "ground truth" is known.106

For instance, to assess the capacity to spot dishonest witnesses, one can instruct some mock witnesses to lie and others to be honest, and then compare the evaluations of the former with the latter. In many disciplines relevant to the law, however, this is impossible. One cannot create an experimental crime scene identical to the real crime scene, to validate an expert's theories about blood spatter patterns or the degree of violent force used on a victim, or subject people to trauma to test the accuracy of alleged symptoms of post-traumatic stress disorder. As we shall see in section V below, even the background theories underlying expert evidence such as DNA profiling and fingerprinting can never be established definitively, but only in terms of probabilities. Once again, this suggests the need for a "best testing" approach which would recognise that not all testing can be benchmarked directly. Again, this suggests that there needs to be a "best testing" requirement related here, not to the methods of testing, but to the benchmark for results.

4. A falsifiability test and a false reliability test

The final two elements of the Young test are no less opaque than the first three, not least because it is unclear whether their reference to "methodology" was used loosely as a synonym for methods or, less plausibly, whether the HCJ understood that technically it also encompasses theory. Aside from this, there are numerous problems with the fourth requirement that expert evidence "follow a developed methodology which is explicable and open to possible challenge".107

If "developed" denotes positive testing, there might be an overlap with the previous two elements (assuming they apply to methods as well as theory). If, however, it means that the method (or methodology) has to be accepted and applied for some (albeit unspecified) time, then it provides for a maturity requirement which is appropriately focused on the means by which opinions are produced rather than disciplines as a whole. It is also possible that "developed" was prompted by the HCJ's criticism of CLA as lacking an "agreed or uniform procedure (either within the UK or worldwide) to check and certify [inter-rater reliability]" and, more generally, "agreed international or national standards".108 Admittedly, many forensic and other sciences have developed forms of accreditation, protocols to guide forensic procedures and other means of monitoring quality control, such as requiring documentation of all steps taken in the collection, analysis and reporting of forensic evidence,109 but these relate to proficiency and not validity or reliability.110

The requirement that the methodology is "explicable" is also ambiguous. If it means simply that

experts must be able to explain how it works,111 then again this is not about evidentiary reliability, but about ensuring that fact-finders do not have their jurisdiction usurped by being unable to independently assess the value of expert evidence.112 Such a motivation far more clearly applies to the final element of the Young test which requires that the expert's methodology "must produce a result capable of being assessed and given more or less weight in light of all the evidence before the finder of fact",113 and therefore does not seem to be about evidentiary reliability at all.
Less plausibly, "explicability" could be interpreted, not as an independent requirement, but conjunctively with the requirement that the methodology must be "open to possible challenge".114 If so, it would echo Daubert 115 in referencing the important scientific concept of falsifiability and have some value in ensuring evidentiary reliability (though falsifiability usually relates more to theories than methods).116 Thus, given Hume's argument that hypotheses about the world based on observation can be invalidated by just one disconfirming instance,117 Popper argued that scientists should seek not repeated confirmations of hypotheses, but repeated attempts at falsification.118 Consequently, only knowledge claims that can be proven wrong deserve the status of science. This excludes theories which cannot be stated precisely enough to be tested, such as Freud's theory that childhood trauma may result in sublimation, projection or repression.119 Because this theory does not specify what conditions would lead to one rather than another response, virtually any symptom can be interpreted as confirming the theory, and hence one cannot test the claimed causal connection to see if particular responses do or do not confirm the hypothesis.

While it seems sensible to exclude such vague theories, a falsifiability requirement may not in fact hit CLA, which suffers more from high error rates than being unfalsifiable. In any event, there are numerous problems with falsifiability as a mandatory requirement.120 For one thing, logically, no theory can ever be falsified since empirical observations which allegedly falsify some theory are themselves subject to falsification by another empirical observation, which is subject to falsification, and so on. Secondly, given that the different experiments for the same phenomena rarely reproduce each other exactly, every theory is prone to falsification by minor differences between experiments.121 More importantly, as we have already seen,122 many accepted theories and methods are incapable of being tested either at all or in any reasonable sense, or there may be no independent criteria by which to measure their falsification or confirmation. Yet this may not stop theories or methods becoming received wisdom, such as in the case of Darwin's "survival of the fittest" theory. Consequently, researchers in many scientific fields pursue alternative means of confirming theories that cannot be directly tested.123 In fact, science rarely displays the culture of "organised scepticism"124 required to subject every new scientific claim to falsification attempts.125 Whether or not claims are subject to attempts at falsification and what is required for them to become confirmed can be an arbitrary and patchy process, not least due to the vagaries of the peer-reviewed publication process.

This suggests that ideally experiments in controlled environments should be conducted where possible, but if this is impossible, the next best method for confirming and attempting to falsify the claim should be required and, again where possible, as many relevant means of confirmation could be sought, as a way of achieving what scientists call consilience or convergence.126 In fact, courts evaluating evidentiary reliability could also look to other factors scientists regard as hallmarks of good science.127 One is the number of times that a test is positively repeated by the original experimenter or more usefully by others attempting to falsify it (what are sometimes called "repeatability" and "reproducibility" respectively). Courts could also look for comprehensiveness in the sense that a theory should be able to coherently explain all relevant data and, finally, for consistency in the sense that it does not conflict with established theories in its own or neighboring fields.

IV. Reforming Young

We can see that the Young test of evidentiary reliability is highly problematic. It is riven with terminological vagueness, conceptual confusion and uncertainty as regards its ambit and application. Its fifth and possibly also fourth elements conflate the issue of reliability with the need to ensure that fact-finders retain their jurisdiction to evaluate the evidence, and its first element is arguably otiose, while simultaneously overly rigorous and unduly lax. Moreover, unless we give those elements which do provide relevant tests for actual validity and reliability (the second, third and fourth) a very charitable interpretation, they seem to apply in a haphazard and partial way. Thus, empirical testing and certainly peer review seem required only in relation to the theory underlying the expert's method and not the method itself—hence only partially ensuring validity but not reliability—whereas if falsifiability is required, it seems to apply only to experts' methods and not their underlying theory.

These inadequacies raise the question of why the HCJ appears not to have utilised the vast literature, case law and experience in other relevant jurisdictions. Admittedly, we have seen that the Frye and Bonython tests are plagued by uncertainty and are very blunt instruments in being both over- and under-inclusive.128 By contrast, while Daubert 's internal gatekeeping approach is far from perfect and has been criticised for requiring far too much of judges lacking expertise in scientific method, not understanding how science works, increasing the complexity, time and cost of trials, and for being too vague to offer sufficient guidance,129 it is far clearer than the Young test and does not noticeably require more of judicial gatekeepers.

On the other hand, it was fortunate that Young appears not to have been influenced by the English and Welsh Law Commission.130 In its proposed Bill dealing with expert evidence, it defined sufficiently reliable expert evidence as (a) being "soundly based" and (b) where "the strength of the opinion is warranted having regard to the grounds on which it is based".131 It then specified five ("higher-order") examples of reasons why evidence might be insufficiently reliable, namely where the opinion is based (a) "on a hypothesis which has not been subjected to sufficient scrutiny (including, where appropriate, experimental or other testing), or which has failed to stand up to scrutiny", (b) "on an unjustifiable assumption", or (c) on "flawed data", or (d) "where the opinion relies on examination, technique, method or process which was not properly carried out or applied, or was not appropriate for use in the particular case", or (e) "on an inference or conclusion which has not been properly reached".132 These examples were augmented by the eight "lower-order" factors133 which the Procurator Fiscal, Alloa court134 found useful and which were stated to apply "where relevant", along with any other relevant factors, including possible factors specific to particular fields which might be (but never have been) developed.135 The Bill was not made into law, however, apparently because of concerns about the cost of training judges to gain sufficient understanding of scientific methods to apply such "a detailed and complex" test, lengthy trials and "floods of appeals".136 Instead, its recommendations were incorporated into Criminal Practice Directions which encourage, but do not require, juridical screening of expert evidence reliability and without the Commission's definition of sufficient reliability.137

Despite deliberating for months, consulting widely and being more scientifically literate than the HCJ, the Law Commission was, however, unable to avoid the sort of uncertainty, patchy coverage of validity and reliability, and encroachment on matters of weight traditionally reserved for fact-finders138 which plague the Young test.139 Nevertheless, details aside, there is much to be said for having general "higher order" evidentiary reliability requirements fleshed out by more specific "lower order" indicia which help judges apply the higher order requirements.140

Building on this, we would suggest that a reformed test commences by stating that in order to be admissible, the validity of the scientific or other theory or form of knowledge on which expert

testimony (whether based on knowledge gained from formal training or experience141 or a mixture of both) is based and the methods or techniques which flow from the valid theory must prima facie be sufficiently established by the most appropriate forms of testing. Where, however, testing is impossible or inappropriate, the proponent of the theory or method must provide alternative means of justifying validity or reliability, respectively. Then, in order to provide guidance on these requirements and to help courts apply the deliberately chosen "weasel words" like "sufficient", "appropriate" and "relevant", the test should specify lower order indicia of evidentiary reliability while making clear that they are also not always relevant. At a minimum such indicia should include those specified by Daubert, namely falsifiability, general acceptance, peer review, and error rates, as well as other hallmarks of good science like repetition and reproduction of confirmatory tests, and coherence and consistency, and the various means to ensure that human subjects and those doing the testing do not personally influence the results.

Admittedly, such a test would impose a heavy burden on judges to understand these factors, when they apply and how to apply them. But this seems inevitable since no clear prior indication can be given of what is sufficient, appropriate and relevant for every form of expert testimony because of the myriad types of scientific evidence which can be used in court and the even more protean forms of experience-based testimony.

Nevertheless, whatever the merits or problems with our suggested approach, the complex debate over the various judicial tests devised for the evidentiary reliability of expert evidence—not to mention limitations to the scientific literacy of lawyers and judges—suggests that reforming the test is best left to a body like the Scottish Law Commission,142 which can draw upon the extensive relevant academic literature in the UK and beyond (particularly the US) which has explored the issue in even greater depth and with greater profit than the English and Welsh Law Commission.

This, however, assumes that the exact wording of an evidentiary reliability test matters, which in turn assumes that the Young test—even if appropriately reformed—will be applied in a way which will improve the validity and reliability of expert evidence relied upon in Scottish criminal cases. We now turn to exploring how likely this is to occur. In doing so, we will confine our discussion to evidence which can be loosely defined143 as (or is claimed by its proponents to be)144 scientific, since such evidence was the clear target of the Young test and is, in any event, far more common than reliance on "experience-based" expertise.

V. Should DNA comparisons, fingerprint evidence, medicine and pathology have been grandfathered?

One reason for being pessimistic about the prospects for the rigorous judicial scrutiny of scientific evidence, at least those associated with the "hard" sciences145 is the apparent willingness of the HCJ in Young to grandfather "scientific evidence about DNA comparisons, fingerprint evidence, [and] evidence of medical practitioners or pathologists [as] evidence based on a sufficiently clear and reliable basis that it may assist the finder of fact".146

However, when we look at the validity and reliability of the four disciplines grandfathered by the HCJ, only DNA profiling ought to have been given a clean bill of health, and even then with some caveats.147 After being exposed to successful challenge in the courts (mainly in the US— the so-called DNA wars),148 researchers went to considerable lengths to establish the validity and reliability of DNA profiling by compiling impressive statistical databases allowing analysts to provide a relatively accurate, but not entirely discretion-free evaluation of the statistical chance of two DNA profiles coming from the same source and to represent this evaluation by a numerical statement.149 As a result, DNA profiling has now replaced fingerprinting as the "gold standard" amongst the common forensic tasks of what is variously called source attribution comparison, feature comparison methods or identification evidence. Here, forensic scientists seek to identify people or

objects suspected of involvement in criminal offences through investigating whether some trace (for example, a finger, palm, ear or shoe print; a bite, tool or tire mark; or a bloodstain, hair or fibre), left by some relevant legal actor matches the "reference" or "exemplar" sample derived from the suspect, shoe, tool, etc.

In the case of DNA evidence, identification profiling is premised on the universally accepted scientific theory that, while most DNA is shared by all humans (and indeed other biological entities), there are regions unique to each individual (even identical twins).150 Within these regions, using clear and well-specified methods, analysists compare "short tandem repeats" (STRs) at a specific number of sites (loci) along the DNA from the crime scene and reference samples.151 Based on the prior examination of millions of samples,152 scientists have estimated that there is less than a one in a billion chance of a random match between unrelated individuals at all examined loci. In ideal conditions, profiling can be done relatively automatically by computer and can be said to be highly reliable. Such conditions do not, however, always prevail. Samples may be degraded, may not produce results at all of the loci examined, or may derive from more than one source (such as when more than one person uses the same cup, or samples become contaminated). Here uncertainty increases as the size of the sample decreases and its sources increase, with a consequent increasingly greater reliance on subjective interpretation, with all the associated dangers detailed below. There is also far less reliability as regards what is called low copy number (or low template) techniques which allow analysis of samples that would previously have not been expected to yield results, because of factors like their age, size, or biological degradation.153 Nevertheless, the HCJ's general opinion of the reliability of DNA profiling finds support in a major investigation by the US National Research Council of the National Academies which concluded that nuclear DNA analysis154 has been "built on solid bases of theory and research".155
In stark contrast, the status of fingerprint analysis is based almost entirely on confident assertion by its practitioners and blind judicial faith.156 As such it is more of a "national treasure"157 than even a usurped previous holder of the gold standard accolade. Unlike with DNA, no testable theory underlies the claim of fingerprinting analysts to be able to reliably determine whether or not the mark left by the ridges on fingers (as well as palms and soles) matches that of the reference sample. While it has long been assumed that no two people have the same ridge patterns, such uniqueness has never been established. Nor could it. Even if one could collect all existing prints, there may be past or future examples of identical prints. The best that can be done is to develop databases of as many samples as possible from which to develop probabilistic claims about the possibility of coincidental matches. But the attempt to do so has only recently begun.

A similar dearth of published and peer-reviewed research undermines the idea that fingerprint analysis is sufficiently reliable to be legally admissible—not to mention the assertions by many forensic identification examiners (frequently accepted by the UK courts,158 but now discouraged by the 2011 Fingerprint Inquiry)159 that they can infallibly match suspects with finger marks, notwithstanding that in many cases the same source can produce non-identical marks. For instance, differences in the skin's elasticity and the pressure imparted when leaving the mark mean that no two marks are ever identical, even if derived from the same finger. In addition, marks may be smudged or emanate from only part of the finger.

More fundamentally, it can be argued that identification based purely on observation can never be sufficiently reliable, given that the method amounts to little more than a highly subjective process of noticing similarities, and that questions of similarity and difference are matters of social construction rather than essence. Analysts in various jurisdictions developed the practice of requiring a specified number of similarities in a sample before declaring a match (16 points in Scotland). But the relatively recent move to a non-numeric approach was justifiably endorsed by

the Fingerprint Inquiry160 given that the number of points is entirely arbitrary, there is no evidence as to what threshold delivers infallible matches, and even high thresholds have been shown to allow misidentifications.161 As Koehler notes: "Where a method depends as heavily on subjective human judgment as does fingerprint examination, the method literally is the people who employ it."162 Admittedly, examiners receive extensive training and will improve with experience, but experience is only relevant if results can be compared with the "ground truth" about the actual origin of traces, and rarely does the criminal process have conclusive evidence of who left the fingerprints.

To establish the reliability of fingerprint examination, there needs to be "black-box" studies involving numerous analysts exposed to the same samples, where the origin of the samples is known only to the experimenters and, importantly, where the analysts do not know whether the test samples come from the source in question. Few such studies have been conducted, however, and then only very recently.163 Whilst these established that fingerprint experts are considerably more accurate than novices, the studies report substantial error rates (possibly as much as 1 in 18 cases of false positives) that are likely to be much higher than most fact-finders will assume.164 Whether they are too high to justify admission as reliable evidence is a difficult question, but even then it is arguable that there has been insufficient testing to justify acceptance of any error rate.

Moreover, such validation tests are usually conducted in ideal laboratory situations. In actual forensic casework, analysts are exposed to varying degrees of information about their cases. Such information may cause "cognitive contamination (where interpretations and judgments are swayed, often without awareness or conscious control, by contextual cues, irrelevant details of the case, prior experiences, expectations and institutional pressures)".165
This in turn leads to various types of biases which cause errors in analysis.166 "Hot biases" may flow from forensic investigators learning of the heinousness of the crime being investigated or other factors likely to arouse their emotions. More common are "cold" biases which stem from examiners learning, for instance, that the alleged sources of forensic material have a criminal record, have confessed or been positively identified by other examiners. These may lead them, for instance, to search for information that confirms prior beliefs (the confirmation bias), interpret ambiguous information as supporting what one expects (the expectation effect), or being reluctant to depart from the first item of information encountered and sufficiently adjust first impressions when new information emerges (the anchoring and adjustment bias). Cumulatively, such biases cause people to attend to, exaggerate and emphasise what is expected and ignore or downplay that which is not. This is confirmed by numerous studies in which contextual information has caused forensic examiners to change their analysis of previously examined samples. While cognitive biases may also affect DNA analysts,167 they are far more likely to affect fingerprint examination which is based purely on subjective judgment rather than a measurable basis for analysis and a statistical basis for making comparisons.

In this light, it is hardly surprising that the US National Research Council included fingerprinting along with all other forms of source attribution as "lacking the capacity to consistently and with a high degree of certainty, demonstrate a connection between an evidentiary sample and a specific individual or source".168 Notwithstanding these problems, fingerprint evidence is in fact considerably more trustworthy than a lot of other highly speculative source attribution techniques (such as those involving firearms, toolmarks, blood splatter, and hair, fibre, shoe and teeth marks) especially where the source of marks are manufactured items rather than biological organisms.169

While our argument here is based on Young ’s treatment of DNA and fingerprint evidence, it is worth noting in passing that all forensic sciences share common features which, with the exception of DNA, should prevent their satisfaction of the Young test for the foreseeable future.

One such feature is the fact that, unlike medical practitioners, pathologists and those who developed and refined DNA profiling, most forensic practitioners lack professional training in scientific knowledge and methodology, and an inculcation into scientific norms which are thought to ensure adherence to accepted scientific methods.170 Indeed, unlike the pure or research sciences, forensic science does not involve a disinterested search for knowledge about the world. Instead, especially in the case of private (as opposed to state) providers, techniques are developed as "products" which can be "sold" to meet the instrumental needs of the "legal masters" to which the providers are "inextricably tethered" if not mere "handmaidens".171 Thus, reward structures in forensic science undermine any motivation to gain independent testing of the validity and reliability of their techniques or to be honest about problems which emerge. Where forensic scientists do engage in research it is usually to find new ways to serve their customers. Testing the validity and reliability of these new methods, if it occurs at all, usually follows rather than precedes its use by legal actors, as we saw in the case of DNA profiling.172 As a result, most forensic "science" techniques still lack "an underlying scientific theory, experiments designed to test the uncertainties and reliability of the method, or sufficient data that are collected and analysed scientifically".173 In fact, it has been argued to be "a mistake … to view forensic science as a science".174

By comparison, both medicine in general and the sub-specialism pathology clearly fit within the scientific domains (notwithstanding interminable debates about whether medicine is in fact more of an art),175 at least in the sense that its practitioners are trained in various scientific disciplines and in scientific methodology, and profess scientific norms. Determining the validity and reliability of the theories and techniques they apply is, however, an impossible task given the variety of the goals pursued (diagnosis, prognosis, treatment, prevention, etc.), types of doctors (surgeons and physicians, general practitioners, hospital doctors, etc.), specialisms (obstetrics, pediatrics, oncology, etc.) and the sheer scale of the different areas of knowledge involved (anatomy, biochemistry, immunology, epidemiology, etc.). This complexity does not, however, justify the HCJ's blanket grandfathering of both medicine and pathology given that it should have been aware of at least occasional problems with cutting edge theories such as the highly controversial "shaken baby syndrome".176

Yet problems are by no means confined to idiosyncratic theories or techniques. Despite medicine's ancient history, only recently have there been concerted efforts to establish the validity and reliability of medical knowledge and techniques through rigorous and repeated, double-masked testing involving large samples published in peer-reviewed journals.177 Thus in the 1990s it was realised that many traditional theories and techniques had never been scientifically established,178 many treatments were either useless or positively dangerous, and many diagnoses were based either on partial and subjective "clinical experience", or at best on "mechanistic reasoning" involving drawing logical but often misleading inferences from knowledge about underlying pathological or physiological causes of disease.179 Since then, medicine has been put on a more scientific basis, and medical research has been heavily regulated and is unlikely to be authorised unless based on sound research methodology. Accusations remain, however, of researchers succumbing to career pressures to carry out research for which they are ill equipped and consequently to "use the wrong techniques (either wilfully or in ignorance), use the right techniques wrongly, misinterpret their results, report their results selectively, cite the literature selectively, and draw unjustified conclusions".180 We thus see that the scientific foundations of medical knowledge are patchy.

Admittedly, problems with the validity and reliability of the theories and techniques relied on by medical experts—as opposed to the much more controversial theories and interpretations of psychiatry and psychology which the HCJ presumably did not intend to "grandfather"181 —are more likely to arise in civil cases where diagnoses and prognoses are more relevant than criminal cases. In the latter, medical practitioners and pathologists largely draw on biological, chemical and various

other forms of medical knowledge about how the body works (both in life and after death) and how it interacts with various weapons and other means of causing death and injury, and with various environmental clues which can reveal important information about death and injury.

An immediate problem for forensic pathologists (who are more narrowly concerned with causes of death)182 is that for some issues—most notably the identification of victims—they may delegate questions of source attribution to forensic scientists who rely on unsubstantiated theories and techniques. More than that, though, medical and pathological knowledge relevant to criminal cases is not based on the sort of double-blind testing of the reliability of interpretations about issues where the ground truth is known; nor could it be. It is unthinkable to stage death or injury in order to test relevant theories in conditions where its causes are known, but even where auxillary theories which inform predictions about death or injury are capable of being tested, most relevant medical and pathogical evidence is based on mechanistic reasoning and clinical experience.

Mechanistic reasoning can undoubtedly be based on highly reliable information about human bodies and other relevant biological organisms (such as the rate of development of maggots in corpses), and physical materials (like gunshot residue, etc.). Nevertheless, medical history is replete with numerous examples of apparently watertight logical inferences from solid knowledge being confounded by rigorous testing. One reason for this is that mechanistic reasoning can cope only with relatively simple biological or other physical or chemical mechanisms and not situations involving multiple mechanisms or even just a few variables which affect how the mechanism works (temperature, age of the body, underlying physical conditions, etc.). Another is that mechanistic reasoning is stochastic in not applying to all people in exactly the same way and hence its conclusions can be delivered only in, usually unnumerated, probabalistic terms.

As regards clinical experience, there are circumstances in which it can deliver relatively reliable theories or opinions, such as through repeated observation of something like the patterns of knife or gunshot wounds, but only if pathologists are certain of relevant variables such as the type of knife (serrated or smooth edge, etc.) or the distance from which the gun was discharged. These problems of being able to link observation to ground truth are particularly acute in the case of live human subjects who might trigger the expectation effect183 by subconsciously or even consciously altering their behaviour in response to prompts and clues conveyed by the medical examiner, Moreover, practitioners will be unlikely to observe modest but nevertheless important effects of phenomena and, as we have seen,184 are cognitively primed to ignore evidence which confounds their theories. But, again, even when observations are based on many similar cases, human biology and the natural environment can always throw up unique cases. On many issues relating to the cause, manner and especially time of death and injuries, medical practitioners and pathologists may work with their own theories or presumptions formed in less than perfect conditions where they are susceptible to contextual biases. Indeed many such theories are used without any (or with merely just one) test.185 Even more problematic are theories such as shaken baby syndrome which are based on a wider range of symptoms which may have many alternative causes.186

This is not to say that mechanistic reasoning and medical experience are not useful bases for expert evidence on death and injury. Certainly, research187 suggests that forensic pathologists assess the cause and manner of death more accurately than other doctors; though only suggests as ground truth about the cases compared is almost never known. Moreover, it is obvious that courts are far better relying on the opinion of all forms of medical practitioners on medical matters than drawing their own inferences from the facts presented to them. It can be argued, however, that at the very least much more needs to be done to base medical theories and techniques on the best possible testing and that the courts should be prepared to examine the evidentiary reliability of particular areas of medical and pathological knowledge and opinion instead of blithely assuming its existence.

Whether they will do so is another matter. The fact that the HCJ provided an undifferentiated clean bill of health to medicine and pathology, and seemed completely unaware of the problems with fingerprint evidence and (more understandably) marginal probems with DNA, can perhaps be explained by the fact its comments were made obiter rather than fully considered. However, as we shall now show, the record in other jurisdictions does not provide greater hope for rigorous scrutiny if the evidentiary reliabilty of scientific evidence is centre stage.

VI. Gatekeeper or grandfather: How likely are the Scottish courts to rigorously assess the reliability of scientific expertise?

1. The judicial record in other jurisdictions

Based on their extensive survey of cases in the UK, US, Canada and Australia, Edmond et al concluded that "the introduction of new admissibility standards in the U.S. and Canada has not radically disrupted historical settlements around admissibility practices".188 In other words, it seems to make little difference whether a jurisdiction adopts the relatively rigorous judicial gatekeeping test laid down in Daubert (or its Canadian equivalent),189 or those like Frye or Bonython 190 which delegate questions of evidentiary reliability to the discipline in question, or eschew any evidentiary reliability. Indeed, admissibility standards have not contributed to the exclusion (or informed systematic evaluation) of unreliable expert evidence, or had much discernible impact on the quality of forensic and medical evidence.

Thus, instead of rigorously examining the validity and reliability of forensic techniques by, for instance, calling for evidence of validation studies and error rates, and taking seriously various reports condemning a wide range of forensic science,191 judges have replaced scientific indicia of evidentiary reliability with what Edmond and his co-authors have called tests of legal or forensic reliability.192 These look to whether the technique has been judicially admitted, resulted in convictions or upheld on appeal, thus scaffolding ongoing admission on cases where there was no independent verification of the relevant scientific evidence. Similarly, the courts apply what they variously call "heuristics" or "legal proxies" involving "epiphenomenal", "secondary" or "indirect" reliability criteria.

These include:

How long the technique has been used and by how many other users;

Whether the expert or their laboratory has been certified or recognised by relevant external organisations;

The experts' credentials such as formal qualifications, training, and years of practice, experience in applying the relevant technique,193 independence, impartiality, response to cross-examination and even demeanour.194

Such factors are either irrelevant (the second and third) or at best only indirect indicators of the validity and reliability of the techniques used (the first and third). Indeed, they are largely about expert proficiency and, in some cases, proficiency as a witness rather than as an expert. Moreover, being highly subjective, most factors allow courts great leeway to expand (or contract) the net of admissible evidence according to factors unrelated to evidentiary reliability.

Similarly problematic195 are a number196 of other judicial practices adopted either individually or in combination. Thus, studies report a greater judicial willingness to admit expert evidence if it does not deal with an ultimate issue, but some other "intermediary" issue, like the origins rather than interpretation of forensic evidence, or where it is corroborated by non-expert evidence, so that the overall strength of the evidence is high even though that of the expert might be less reliable than might otherwise be required.197

These practices might have some justification if applied to defence-led expert evidence given that they would accord with the idea that criminal suspects and accused need to be "overprotected"198 by mechanisms such as the higher standard of criminal proof, the privilege against self-incrimination and the corroboration requirement, which have been developed in recognition of the greater power of the state over the individual and the much harsher consequences of wrongful convictions as opposed to incorrect acquittals. However, easing the path of the prosecution's expert evidence simply adds to the advantages which created the need for this "principled asymmetry"199 in the first place. Moreover, once admitted in circumstances thought to justify reduced scrutiny, admissibility decisions may be binding in cases absent such circumstances. More specifically, as regards the practice of allowing in weak expert evidence if corroborated by other evidence, the latter is not just irrelevant as regards evidentiary reliability,200 but may interact with it in problematic ways such as where accused confess when confronted with apparently insurmountable, but in fact vulnerable, scientific evidence.201

There is still less justification for two final ways in which courts sidestep scrutiny of dubious expert evidence. One is the "Solomnic compromise"202 of neither excluding such evidence nor giving it a clean bill of health, but restricting the strength of experts' opinion about the evidence by, for instance, precluding them from declaring a definite as opposed to possible match between samples, even though research suggests that watering down the strength of an expert's opinion does not result in it being given less weight when combined with other incriminating evidence.203 Indeed, paradoxically, this "half-way house" might make untested theories and techniques seem more justified than they are, or at least discourage fact-finders from reducing the weight attached tothem.204 The final problematic judicial practice is the ubiquitous one of sidestepping problems with the evidentiary reliability of experts' evidence by concentrating on the wrongdoing of individual experts or the institutions that employ them.205

2. Understanding the judicial record

One reason why judges fail to scrutinise expert evidence for evidentiary reliability is the failure by lawyers to challenge it.206 Instead, lawyers tend to focus on problems with individual witnesses (qualifications, partiality, etc.), the procedures they use (such as the integrity of the chain of custody), or their equipment and laboratories.207 No doubt, this may owe much to a lack of knowledge about what makes good science208 or, as we shall see in the case of defence counsel,209 the time and resources necessary to mount effective challenges and instruct their own experts to challenge prosecution experts. But even if lawyers were more active in challenging scientific evidence it is not clear that the courts would respond accordingly. For one thing, studies of US judges (who are given similar levels of relevant background knowledge to those in Scotland)210 reveal a poor understanding of the sort of basic scientific criteria relied on in Daubert.211

Other factors which might play a role are respect for precedent or simple inertia.212 More fundamentally, however, it is argued that judges suffer from a fundamental "lack of will" when asked to rule scientific evidence inadmissible.213 This may owe something to traditional judicial deference towards knowledge from the gentlemanly pursuits of "natural philosophy, medicine and earlier manifestations of engineering", which was initially taken on trust as coming from fellow elites,214 and then assumed to be equally infallible when extended to the applied medical and physical sciences which followed.

Perhaps even more fundamentally judges are likely to be discouraged from raising doubts about long-accepted types of expertise because of the institutional and political embarrassment of admitting that countless convictions based on untested forms of expertise might be suspect.215

More urgently, severe disruption to the criminal justice system would ensue if ongoing cases were abandoned, recent convictions challenged and those minded to plead guilty in the face of apparently sound forensic evidence decide to proceed to trial.216 No doubt judges oriented towards a "due process" model of the criminal justice system,217 which celebrates the above noted218 "principled asymmetry" and "overprotection" of criminal suspects and accused as a way of placing hurdles in the way of unjust convictions, would not baulk at prioritising the integrity of decision-making and civil liberties over expediency. However, the record of courts in other jurisdictions regarding expert evidence,219 and in Scotland regarding criminal evidence more generally,220 suggests that judges may well adopt a more "crime-control" orientation, which prioritises the suppression of crime through high rates of detection and conviction, achieved with maximum efficiency and speed, and at minimal cost. Indeed, courts elsewhere have been far more rigorous in screening for reliability of expert evidence led by the defence as opposed to the state.221
A final reason why judges do not rigorously screen the evidentiary reliability of scientific evidence is the assumption that the adversarial process will ensure that convictions are based only on reliable evidence. Judges seems to believe that problems with reliability will be raised by defence lawyers in cross-examination or by leading counter-witnesses, conveyed in judicial directions to the jury, understood and appropriately acted upon by jurors or, if all else fails, remedied by appeal courts. Studies, however, have repeatedly shown that none of these assumptions hold water.222
First, even if the defence focus on questions of evidentiary reliability rather than factors like the expert's partiality and qualifications, and the security of the custody chain of forensic evidence, cross-examination has been shown to be ineffective at exposing problems with the evidentiary reliability of expert evidence.223 Secondly, the option of defence counsel leading experts in rebuttal is undermined by the fact that, compared to the prosecution, which has easy access to its own experts, private experts are thin on the ground (especially outside Scottish cities) and increasingly likely to be beyond the means of many accused as legal aid shrinks. But, even if the defence can access an expert, they nevertheless face numerous disadvantages.224

For example, forensic materials may be destroyed during their examination by the prosecution or degrade with time. Equally, the party holding the forensic materials (including private laboratories utilised by the prosecution) may prevent access to them or control tests conducted by opposition scientists on their turf.225 Then, when it comes to preparing reports and presenting evidence, state experts are very likely to offer evidence subconsciously or even consciously disposed towards the prosecution. By contrast, being paid and often coming from the academic rather than the forensic world, defence experts can more easily be portrayed as lacking relevant "real world" experience.226
In addition, defence experts' testimony may not be buttressed by other non-expert evidence and thus made more persuasive by being part of an overall story. Not only does research show that holistic stories are more persuasive than atomistic challenges to parts of opposition stories,227 but they may also encourage fact-finders to use their non-scientific knowledge to fill in missing gaps in the overall story.228 Finally, the mere leading of scientific evidence by an apparently disinterested prosecution, seeking to bring criminals to justice, may suggest evidentiary reliability, which is then reinforced by its endorsement in examination-in-chief and directions to the jury.229

As regards such directions themselves, there is little evidence they have a positive impact in guarding against unreliable expert evidence.230 Limited scientific literacy may prevent judges from effectively highlighting problems of reliability and validity,231 whereas jury directions often come too late to counter already formed conclusions or help jurors struggling to cope with information overload. And whereas all judicial directions are prone to comprehensibility problems,232 jurors are unlikely to find them useful or pay much heed to them when (as is usually the case)233 they involve only abstract warnings about unreliability problems in general, rather than guidance on case-specific problems with scientific evidence. Moreover, any message about possible evidentiary unreliability might be countered by the fact that the expertise in question has been regularly admitted in court.

In addition to the limitations of the various supposed safeguards against reliance on unreliable evidence, the adversarial trial (and still less pre-trial proceedings) is not in general well-suited to exposing the flaws with expert evidence, given the incentives for the parties to present evidence in a one-sided fashion and the tendency of battles of experts to confuse jurors.234 And this is before one takes into account the fact that the fact-finder may lack the ability and not just the information necessary to make sense of challenges to evidentiary reliability.235

Thus, while research shows that fact-finders are not, as feared by some,236 routinely so dazzled by scientific evidence that they automatically defer to it or at least give it undue prominence, this does not necessarily mean that they will understand and pay due regard to criticisms of its evidentiary reliability, still less notice problems for themselves.237 Instead, when the evidence is complex or contradictory, they tend to focus on peripheral issues like expert demeanour, and when the evidence is of a statistical nature, on various misleading heuristics and cognitive biases. Indeed, when it comes to scientific evidence presented in statistical form, as is—or ought to be—the case with all identification and much other forensic evidence, the performance of research subjects is particularly poor,238 and that of lawyers and judges hardly better, even if exposed to post-school mathematical training.239

The final formal safeguard against fact-finders being misled by unreliable expert evidence is the appeal process. However, the chances of successful appeals are not only rather slim in general,240 but also likely to be affected by the very same causes of the relaxed judicial approach to admissibility noted above. Indeed, if any of these formal safeguards were as effective as assumed, we would be unlikely to see so many miscarriages of justice caused by unreliable scientific evidence.241 The same logic undermines faith in more informal safeguards represented by the duty of prosecutors to act as officers of the court rather than zealous pursuers of convictions, and various ethical codes and procedural obligations on experts.242

VII. Reforming the assessment of evidentiary reliability

The above problems undermining the accurate assessment of the validity and reliability of scientific evidence raise the question of how matters can be improved.243 Here we can distinguish between reforms to general court procedures and those which focus on the question of admissibility itself.

1. Going beyond admissibility

An obvious starting point for reforms to general court procedures would be to attempt to ensure that judicial assumptions about the effectiveness of the adversarial system are not simply an article of blind "judicial faith".244 The courts and professional regulators could in theory take seriously the duty of prosecutors not to submit unreliable evidence, though in practice this would radically curtail their ability to bring prosecutions if interpreted to require validation studies of expert theories and techniques.245 More realistically, lawyers could do more to ensure scientific evidence is more effectively evaluated in court, such as by encouraging experts to instruct adjudicators on how to approach statistical and other technical evidence, and to present evidence in a comprehensible manner, using multi-media aids.246 The defence should also be guaranteed better access to state expert reports, disclosure of which should be required to include negative test findings and known error rates. Potentially wider in its impact, but less realistic, is an increase to legal aid provision to enable the defence equal access to experts and more time to ensure zealous representation.

There is little that can be done through legal requirements to ensure defence lawyers and courts take evidentiary reliability more seriously. Less formally, problems with scientific and other expert evidence could be made more central to both legal and judicial training. Judicial directions to the

jury might then become more effective—especially if combined with even brief247 training on how to evaluate scientific evidence. Nor is it only lawyers who would benefit from training. Experts themselves arguably could be trained in court procedures before appearing as witnesses.248

Other procedural reforms are more controversial in altering the adversarial nature of criminal trials. Least problematic would be extending and strengthening the mechanisms249 requiring experts to meet so as to agree as far as possible evidence before trial.250 Rather more controversial would be to group expert testimony around issues, rather than according to who calls the witness,251 and exempting experts from the constraints of the fragmented style of testimony,252 such as by using concurrent evidence ("hot tub") sessions involving all experts from similar or closely related fields initially testifying together without lawyers and cross-examination.253

By comparison, while not affecting the adversarial nature of trials, their characteristic continuous nature would be disrupted if one was to follow the initial, but subsequently dropped, recommendations of the English and Welsh Law Commission to allow interlocutory appeals on questions of reliability.254 This might avoid lengthy trials and wrongful convictions based on expert evidence which is ruled inadmissible later on appeal. Appeal courts might also be less likely than trial courts to allow problems with reliability to be compensated by other corroborating evidence.255 Also likely to reduce miscarriages of justice would be to follow another of the Commission's proposals and allow appeal courts to treat decisions on evidentiary reliability as rulings of law, thus allowing de novo reconsiderations rather than simply review for error.256 While this would modify the traditional approach to finality in Scottish trials, other suggestions257 are likely to be more controversial in challenging more revered procedural values.

One is the idea of replacing jurors with judges when cases involve complex science issues, or even establishing "science courts" staffed by scientifically-trained judges. These could occur only if requested by the parties, or alternatively, be mandatory in all cases where scientific evidence is challenged, or at least258 where novel scientific evidence is proffered. Such courts could decide the entire case, only issues involving scientific expertise, or simply rule on the admissibility of allegedly invalid or unreliable expert evidence. In support, it can be noted that it is easier to educate judges than jurors, and that judges may acquire relevant knowledge over time, use research assistants or even educate themselves.259 On the other hand, studies reveal that knowledgeable jurors may tutor their fellow jurors and group deliberation may improve the quality of jurors' evaluation of scientific evidence.260 Also problematic is the dilution of the principle of being judged by community representatives who can "import a social sense of justice".261 Although this problem applies less to "blue ribbon" juries comprised exclusively of scientifically-qualified jurors,262 the latter's relative scarcity in a small jurisdiction like Scotland makes such an option unfeasible. Instead of using scientists as adjudicators, they could act as "assessors"263 sitting alongside judges, so that the latter could be educated as needs arise, rather than having to undergo more wide-ranging and hence either time-consuming or unduly-generalised training on all potential forms of expertise. More controversially, neutral experts could testify in court and even, as on

the Continent, investigate cases from the outset.264 Having neutral investigators would alleviate problems regarding inequalities between parties in investigating and preparing cases, while all neutral experts would eradicate problems associated with the adversarial selection, preparation and questioning of experts, thus helping to level the uneven playing field between state and defence.265 However, given that they might still have partisan views on scientific controversies, neutral experts should still be subject to adversarial testing. Indeed, many commentators accept that parties should retain the right to lead their own experts, which would then further increase the time and cost of legal proceedings and the confusion caused by multiple experts. A court-appointed expert might also

exacerbate problems of fact-finder deference to expertise, in that jurors might be influenced by the expert's official status.266 There is also the thorny issue of expert selection. Judges are not themselves qualified to choose the most suitable experts, but delegation of this task to scientists may cause personal and professional rivalries to distort selection and existing elites to block those in the vanguard of new developments.267 In addition, there may be insufficient experts in some disciplines to both accredit and be accredited.

2. Going back to admissibility

Given these problems, and doubts about whether busy268 judges can ever be sufficiently educated about scientific methods and standards to make sensible assessments of evidential reliability,269 and the court system sufficiently reformed to facilitate this role, some recommend that decisions about the reliability of scientific disciplines or their techniques are referred to an independent standing body before or even during trial, preferably sitting without representatives from the discipline in question.270 Depending on how appropriate it is thought to allow the courts to defer to external experts and how hopeful one is that judges can be assisted to make informed evaluations, conclusions can be either binding or merely advisory and, given the traditional division of labour between judge and jury, should be confined to the issue of the admissibility rather than weight of the evidence. Moreover, such a body could make rulings, not merely when issues arise or even more narrowly on novel forms of expertise, but also act proactively to review all those disciplines which have been used for years without appropriate challenge.

Indeed, it can be argued that expert evidence should only be admissible if it is based on knowledge established, where relevant,271 by rigorous and peer-reveiwed validation studies, based on knowledge about the ground truth of the phenemonon being tested and which produce error rates.272 Given the judicial record elsewhere, the apparent judicial unwillingness to scrutinise scientific expertise and disrupt existing state and prosecution practices, and the HCJ's apparent grandfathering of some of the most prominent, but not necessarily scientifically secure, disciplines, such a rule offers a potentially more effective means of ensuring the evidentiary reliability of scientific evidence than relying on the Young test, especially if combined with an independent body to decide whether scientific evidence meets validity and reliability standards.

However, it seems unlikely that the Scottish Government will be willing to prioritise this issue over more pressing concerns, undermine many ongoing prosecutions, raise doubts about past convictions, and devote scarce resources to fund a panel of experts. Consequently, questions of admissibility seem certain to remain in judicial hands for the foreseeable future. Accordingly, it remains important to ensure the best possible formulation of an evidentiary reliability test. Here it is worth noting that whilst the Daubert and other admissibilty tests which require judges to independently evaluate the validity and reliability of scientific theories and techniques have made only modest differences to admissibility patterns, they have not lacked any impact at all.273

In any event, preventing even a few unjustified convictions is arguably worth the effort of seeking to strengthen the Young test. Morever, excluding scientific evidence is likely to have a ripple effect in warning those disciplines who have hitherto blithely peddled their expert wares to put their scientific house in order through research and refinements to their theories and techniques.274 Equally, a few victories might encourage defence lawyers to use Young and to use it effectively by becoming better acquainted with what constitutes good science, and apprising judges of the importance of evidential reliability275 and about other matters like experts over-claiming the certainty of their opinions,276 while also encouraging prosecutors to seek more reliable evidence. Making issues of validity and reliability relevant to admissibility may also ease the job of fact-finders in allowing them to concentrate on their main job of assessing the accuracy and weight of scientific evidence.277 Finally, the possibility of decisions on admissibilty creating precedents will not only ensure greater procedural justice through certainty and consistency of approach, but also

be efficient in reducing the number of judges who will have to get up to speed on complicated issues of evidentiary reliability. Hopefully, this will also lead to better decisions if judges and lawyers invest more time in familiarising themselves with the relevant scientific knowledge when the outcome will set a precedent.

In addition to strengthening the Young test along the lines suggested in section IV of this article, there are procedural reforms which could enhance the protection of criminal accused and reduce the chances of invalid and unreliable scientific evidence causing miscarriages of justice, as well as futile prosecutions. One is to enhance existing requirements for disclosure and agreeing evidence pre-trial by specifially requiring those leading expert evidence278 to disclose any known "limitiations and uncertainties, and the error rates associated with the technique" or other forms of expert knowledge.279 Evidence from other jurisdictions suggests, however, that experts are either unaware of or resistant to obligations to report evidence undermining their conclusions and lawyers have done little to ensure that they do so.280
A potentially more effective protection for criminal accused would be to impose a legal burden of proving the validity and reliability of scientific evidence on its proponent and to extend the idea of "principled asymmetry" to the issues of evidentiary reliability of scientific evidence by requiring that the prosecution establish the validity of novel or even all scientific evidence beyond a reasonable doubt, while subjecting the defence to the lower civil standard281 or even exempt it completely from having to establish a lack of validity or reliability of expert evidence.282 More informally, courts could be encouraged to reverse the approach seen in other jurisdictions where prosecution expertise is treated more leniently than that of the defence and allow the latter to adduce evidence even if it does not strictly satisfy all aspects of the original or a revised Young test.283 And lest this is regarded as unfair on the prosecution, Edmond suggests a "mirror image" provision284 extending the same relaxed standard for any scientific evidence used to rebut that of the defence.

One final admissibility-related reform would be to encourage judges to raise concerns about evidentiary reliability ex proprio motu and resolve them pre-trial if noticed in time.285 While this conflicts with their traditionally passive role it would go some way towards remedying the apparent286 current lack of initiative by defence lawyers to regularly challenge the validity and reliability of scientific evidence, while saving the time, expense and stress of trials where this leads to prosecutions being discontinued.

VIII. Concluding thoughts

Especially if supported by the procedural reform canvassed above, it would thus seem that it is worth reforming the Young test of admissibility both as a direct means of ensuring that only valid and reliable scientific evidence is led in court and more indirectly in order to encourage scientific experts themselves to enhance the evidentiary reliability of the evidence used in the criminal justice system. At the same time, however, two caveats can be noted.

First, it is not suggested that these reforms would solve all problems with scientific expertise or end resultant miscarriages of justice, even if lawyers and judges were to understand fully and take seriously problems with evidentiary reliability. On the contrary, it must be recognised that these problems may in fact be less serious than those that arise in the processes of collecting, analysing and reporting scientific evidence.287 Consequently, it is arguable that we need to pay as much, if not more, attention to "front-end" reforms288 to the process of dealing with expert evidence "upstream"289 of the courts compared to the issues of the admissibility and evaluation of expert evidence in court. This is because the benefits of such reforms will potentially affect many more suspects affected by scientific evidence than just those capable of challenging admissibility. Furthermore, the problems with the handling of scientific evidence apply to all forms of evidence, including those that are based on valid and reliable scientific theories and methods, and hence the

dangers they pose are far more widespread than those caused by invalid and unreliable scientific evidence. It is no doubt little comfort to those convicted with contaminated, misinterpreted or misreported scientific evidence that it is based on valid and reliable theories and methods.
An exploration of these problems and possible solutions would, however, involve an article in its own right. In any event, just as it is no comfort to those convicted on improperly applied methods to know that they are valid and reliable, and based on valid theories, so is it no comfort to learn that properly applied methods are unlikely to result in accurate results because of problems of evidential unreliablity. The Young decision creates the potential for remedying the latter problem, and it is thus worth seeking to ensure its best possible formulation and most effective implementation.

The second caveat to note is that, contary to the possible impression conveyed by this article, it it not denied that the sciences can claim to "constitute the richest and most extensive body of human knowledge"290 and that, even with all its problems, scientific evidence is "the most reliable of the organised knowledges at our disposal".291 It is certainly more reliable than other common forms of evidence in criminal cases such as witness observation and confessions.292 At the same time, however, this suggests that scientific evidence deserves "respect rather than deference"293 and that such respect would be even more deserved if judicial scrutiny of evidentiary reliability encouraged the forensic sciences, medicine and pathology to strengthen the scientific foundations of their disciplines. Hopefully this article will help in this aim.

Donald Nicolson
Rhonda Wheate

Footnotes

1 By this we mean those areas of knowledge which test ideas through investigating evidence from the natural and social world (cf. the definition adopted by the Science Council: https://sciencecouncil.org/about-science/our-definition-of-science/ [Accessed 9 November 2020], though, etymologically and historically, "science" can be interpreted more broadly as any organised critical examination of knowledge: D. Nicolson, Evidence and Proof in Scotland: Critique and Context (Edinburgh: Edinburgh University Press, 2019), p.180).
2 Prompted by discovering how to prove a coin was not made of gold as claimed in a legal dispute.
3 See, e.g. C.M. Bowers, "The History of Experts in English Common Law, with Practice Advice for Beginning Experts" in C.M. Bowers (ed.), Forensic Testimony: Science, Law and Expert Evidence (Oxford: Elsevier, 2014); C. Jones, Expert Witnesses: Science, Medicine and the Practice of Law (Oxford: Clarendon, 1994), Chs 2–6 passim.
4 P. Roberts and A. Zuckerman, Criminal Evidence, 2nd edn (Oxford: Oxford University Press, 2010), p.470.
5 Scottish Legal Aid Board, Annual Report 2018–2019 (SLAB, 2019), Appendix 1, "Key Statistics", p.8, https://www.slab.org.uk/app/ uploads/2019/09/SLAB-Annual-Report-2018-19-Appendix-1-Key-Statistics.pdf [Accessed 21 November 2020].
6 M. Zander and P. Henderson, Crown Court Study (HMSO, 1993), Royal Commission on Criminal Justice Research Study No.19, pp.84– 85.
7 P. Roberts, "Science, Experts, and Criminal Justice" in M. McConville and G. Wilson (eds), The Handbook of the Criminal Justice Process (Oxford: Oxford University Press, 2002), p.259.
8 See, e.g. J. Fraser and R. Williams (eds), Handbook of Forensic Science (Cullompton: Willan, 2009).
9 M.R. Damaška, Evidence Law Adrift (New Haven, Connecticut: Yale University Press, 1997), pp.143–144.

10 See Nicolson, Evidence and Proof in Scotland (2019), pp.198–204 passim.
11 See the more recent cases discussed by D. Nicolson and D. Auchie, "Assessing Witness Credibility and Reliability: Engaging Experts and Disengaging Gage?" in P. Duff and P. Ferguson (eds), Current Developments in Scottish Criminal Evidence Law (Edinburgh: Edinburgh University Press, 2018), pp.170–172.
12 R. v Turner [1975] Q.B. 834 at 841, quoted with approval in Gage v HM Advocate [2011] HCJAC 40; 2011 S.C.L. 645 at [25].
13 J. Schklar and S.S. Diamond, "Juror Reactions to DNA Evidence: Errors and Expectancies" (1999) 23 Law and Human Behavior 159, 159.
14 Schklar and Diamond, "Juror Reactions to DNA Evidence: Errors and Expectancies" (1999) 23 Law and Human Behavior 159, 160.
15 See, e.g. R. Wheate, "The Importance of DNA Evidence to Juries in Criminal Trials" (2010) 14 International Journal of Evidence & Proof 129; R. Baskin and I.B. Sommer, "Crime-Show-Viewing Habits and Public Attitudes Toward Forensic Evidence: The 'CSI Effect Revisited'" (2010) 31 Justice System Journal 97.
16 For UK examples, see M. Redmayne, "Expert Evidence and Disagreement" (1996) 30 UC Davis Law Review 1027, 1039–1046; C. Walker and R. Stockdale, "Forensic Evidence" in C. Walker and K. Starmer (eds), Miscarriages of Justice: A Review of Justice in Error

(Oxford: Blackstone, 1999); L. Hackman, "Miscarriages of Justice and the Role of the Expert" in L. Hackman, F. Raitt and S. Black (eds), The Expert Witness, Forensic Science and the Criminal Justice Systems of the UK (London: CRC Press, 2019).
17 M.J. Saks and J.J. Koehler, "The Coming Paradigm Shift in Forensic Identification Science" (2005) 309 Science 892, 893; B.L. Garret and P. Neufeld, "Invalid Forensic Science Testimony and Wrongful Convictions" (2009) 95 Va. L. Rev. 1.
18 See Raitt on Evidence: Principles, Policy and Practice, edited by E. Keane and F. Davidson, 3rd edn (Edinburgh: W. Green, 2018), p.71
and the more detailed discussion in Sir Anthony Campbell, The Fingerprint Inquiry Report (Edinburgh: APS Group Scotland, 2011).
19 Campbell, The Fingerprint Inquiry Report (2011).
20 See, e.g. Raitt on Evidence: Principles, Policy and Practice (2018), Ch.4; F.P. Davidson, Evidence (Edinburgh: W. Green, 2007), Ch.11;
M. Ross and J. Chalmers, Walker and Walker: The Law of Evidence, 4th edn (Haywards Heath: Bloomsbury Professional, 2015), Ch.16.
21 Unlike other witnesses, experts are also permitted to charge for their testimony and to sit in court before testifying: Ross and Chalmers, Walker and Walker: The Law of Evidence (2015), Ch.16.
22 e.g. in HM Advocate v McGinty, 1984 S.C.C.R. 176 the HCJ relied on alleged flaws in the methods used by forensic scientists to test banknotes for traces of cannabis in holding that the accused's conviction for being concerned in the supply of cannabis resin was unsafe.
23 Young v HM Advocate [2013] HCJAC 145; 2014 S.L.T. 21. See brief discussions by G. Maher, Guarding the Gate: Some Problems in Expert Evidence in Scots Law (Edinburgh: Edinburgh University Press, 2015), Edinburgh Law School Working Papers No.2015/07;
T. Welsh, "The Impact of Case Law and Inquiries on the Role of the Expert Witness" in The Expert Witness, Forensic Science and the Criminal Justice Systems of the UK (2019), pp.21–23.
24 See, e.g. D. Howitt, Introduction to Forensic and Criminal Psychology, 6th edn (Harlow: Pearson, 2018), Chs 14–15.
25 Young, 2014 S.L.T. 21 at [54] (numbering added).
26 Young, 2014 S.L.T. 21 at [55].
27 S.A. Cole, "Grandfathering Evidence: Fingerprint Admissibility Rulings from Jennings to Llera Plaza and Back Again" (2004) 41 American Criminal Law Review 1189.
28 Daubert v Merrell Dow Pharmaceuticals Inc 509 U.S. 579 (1993) (Sup. Ct).

29 Frye v United States 297 F. 1013 (D.C. Cir. 1923).
30 G. Edmond and D. Mercer, "Keeping 'Junk' History, Philosophy and Sociology of Science Out of the Courtroom: Problems with the Reception of Daubert v Merrell Dow Pharmaceuticals Inc" (1997) 20 University of New South Wales Law Journal 48, 52.
31 G. Edmond et al, "Admissibility Compared: The Reception of Incriminating Expert Evidence (i.e., Forensic Science) in Four Adversarial Jurisdictions" (2014) 3 University of Denver Criminal Law Review 31, 39–40.
32 See text at fn.79 below. On the other hand, in not requiring courts to be experts in the very area they need expertise, the test can be said to involve a realistic epistemology of admissibility: B. Leiter, "The Epistemology of Admissibility: Why Even Good Philosophy of Science Would Not Make for Good Philosophy of Evidence" (1997) Brigham Young University Law Review 803, 819.
33 See, e.g. M.I.H. Graham, "The Daubert Dilemma: At Last a Viable Solution" (1998) 2 International Journal of Evidence & Proof 211, 215.
34 Daubert 509 U.S. 579 (1993) at 593–594.
35 Procurator Fiscal, Alloa v C, unreported, 1 December 2014, Alloa Sheriff Court.
36 Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales (TSO, 2011), Law Com. No.325.
37 See also A. Roberts, "Rejecting General Acceptance, Confounding the Gatekeeper: The Law Commission on Expert Evidence" [2009] Crim. L.R. 551, 558; G. Edmond and K. Roach, "A Contextual Approach to the Admissibility of The State's Forensic Science and Medical Evidence" (2011) 61 University of Toronto Law Journal 343, 375.
38 Kennedy v Cordia (Services) LLP [2016] UKSC 6; 2016 S.C. (U.K.S.C.) 59.
39 Jones v HM Advocate [2016] HCJAC 65; 2016 S.C.L. 743; Graham v HM Advocate [2018] HCJAC 57; 2018 S.C.C.R. 347 at 115, 123 and 124.
40 Kennedy, 2016 S.C. (U.K.S.C.) 59 at [37].
41 Roberts and Zuckerman, Criminal Evidence (2010), p.496.
42 R. v Bonython (1984) 38 SASR 45 at 46.
43 Kennedy, 2016 S.C. (U.K.S.C.) 59 at [43].
44 A.C. Love, "History, Scientific Methodology, and the 'Squishy' Sciences" (2006) 49 Perspectives in Biology and Medicine 452.
45 See, e.g. Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, para.133ff.
46 Sometimes called ad hoc experts in other jurisdictions: Edmond et al, "Admissibility Compared: The Reception of Incriminating Expert Evidence (i.e., Forensic Science) in Four Adversarial Jurisdictions" (2014) 3 University of Denver Criminal Law Review 3, especially 80; Roberts and Zuckerman, Criminal Evidence (2010), p.477.
47 e.g. White v HM Advocate, 1986 S.C.C.R. 224.
48 Griffiths v Hart [2005] HCJAC 51; 2005 1 J.C. 313. Even more dubiously, in Hopes and Lavery v HM Advocate, 1960 J.C. 104, a court stenographer was held to be an expert in interpreting an audio tape recording of the accused and victim despite never having interpreted tapes before.
49 Procurator Fiscal, Alloa v C, unreported, 1 December 2014, Alloa Sheriff Court at [41] and [42].
50 See Davidson, Evidence (2007), para.11.12; Ross and Chalmers, Walker and Walker: The Law of Evidence (2015), para.16.3.5.
51 Kumho Tire Co v Carmichael 526 U.S. 137 (1999).
52 But evidently many courts tend to apply each in turn: J.A. Moreno, "Eyes Wide Shut: Hidden Problems and Future Consequences of the Fact-Based Validity Standard" (2003) 34 Seton Hall Law Review 89, 92.
53 Expressions taken from r.702 of the Federal Rules of Evidence, which formed the basis of the Daubert decision.
54 Kennedy, 2016 S.C. (U.K.S.C.) 59 at [54] (emphasis added).

55 Kennedy, 2016 S.C. (U.K.S.C.) 59 at [55].
56 Jones, 2016 S.C.L. 743.
57 Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, paras 5.71–15.82, see also para.3.41ff
58 Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, paras 3.40–3.48, 5.71–5.82. See also Roberts, "Rejecting General Acceptance, Confounding the Gatekeeper: The Law Commission on Expert Evidence" [2009] Crim.
L.R. 551; D.W. Shuman, and B.D. Sales, "The Admissibility of Expert Testimony Based Upon Clinical Judgment and Scientific Research" (1998) 4 Psychology, Public Policy, and Law 1226.
59 Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, especially paras 3.49–3.52.
60 Term taken from Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, para.3.41.

61 M. Redmayne, Expert Evidence and Criminal Justice (Oxford: Oxford University Press, 2001), p.133. See also Edmond and Roach, "A Contextual Approach to the Admissibility of The State's Forensic Science and Medical Evidence" (2011) 61 University of Toronto Law Journal 343, 401; Shuman, and Sales, "The Admissibility of Expert Testimony Based Upon Clinical Judgment and Scientific Research" (1998) 4 Psychology, Public Policy, and Law 1226, 1248–1250, passim.
62 As apparently occurs in the US: Edmond et al, "Admissibility Compared: The Reception of Incriminating Expert Evidence (i.e., Forensic Science) in Four Adversarial Jurisdictions" (2014) 3 University of Denver Criminal Law Review 31, 94.
63 Edmond et al, "Admissibility Compared: The Reception of Incriminating Expert Evidence (i.e., Forensic Science) in Four Adversarial Jurisdictions" (2014) 3 University of Denver Criminal Law Review 31, 94.
64 Kennedy, 2016 S.C. (U.K.S.C.) 59 at [41].
65 Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, para.3.37.
66 Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, para 3.37–3.39, 5.4; see also paras 2.19– 2.23, 5.81.
67 See, e.g. Raitt on Evidence: Principles, Policy and Practice (2018), para.4.02, Davidson, Evidence (2007), para.11.01.
68 See Nicolson, Evidence and Proof in Scotland (2019), pp.55–63 passim, but especially, pp.56–57.
69 As in Myers v The Queen [2015] UKPC 40; [2016] A.C. 314.
70 See, e.g. K.R. Kreiling, "Scientific Evidence: Toward Providing the Lay Trier with the Comprehensible And Reliable Evidence Necessary to Meet the Goals of the Rules of Evidence" (1990) 32 Arizona Law Review 915, 968ff.
71 Though sometimes "foundational reliability" is used to describe the validity of methods: see US President's Council of Advisors on Science and Technology, Report to the President Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature- Comparison Methods (PCAST 2016), pp.42–44.
72 Daubert 509 U.S. 579 (1993) at 590–591, fn.9.
73 Daubert 509 U.S. 579 (1993) at 591.
74 See Developments in the Law, "Confronting the New Challenges of Scientific Evidence" (1994–1995) 108 Harv. L. Rev. 1481, 1548, though noting that the distinction between relevance and reliability will not necessarily be clear.
75 See D.L. Faigman, "Expert Evidence: The Rules and the Rationality the Law Applies (or Should Apply) to Psychological Expertise" in

D. Carson and R. Bull (eds), Handbook of Psychology in Legal Contexts, 2nd edn (Chichester: John Wiley and Sons, 2003), pp.376– 378, though noting that this distinction is not always easy to make in practice.

76 K.E. Howell, Introduction to the Philosophy of Methodology (London: Sage Publications, 2013) and cf. Campbell v HM Advocate, 2004
S.L.T. 397 at [83], where theory and methods were correctly distinguished.

77 Bonython (1984) 38 SASR 45.

78 Frye 297 F. 1013 (D.C. Cir. 1923).

79 For similar criticism of Frye, see, e.g. Faigman, "Expert Evidence: The Rules and the Rationality the Law Applies (or Should Apply) to Psychological Expertise" in Handbook of Psychology in Legal Contexts (2003), pp.383–384; Redmayne, Expert Evidence and
Criminal Justice (2001), pp.25, 101–107, 115–116; A. Roberts, "Rejecting General Acceptance, Confounding the Gatekeeper: The Law Commission on Expert Evidence" [2009] Crim. L.R. 551 and "Drawing on Expertise: Legal Decision-making and the Reception of Expert Evidence" [2008] Crim. L.R. 443, 455–457. Slightly less problematic is the Bonython requirement of "sufficiently organized body of knowledge or experience", but then questions arise as to what constitutes sufficient organisation and more importantly how organisation translates into validity and reliability: cf. Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, para.5.105, noting that many experts, such as amateur lip readers, do not benefit from organisational structures.

80 E. Beecher-Monas, Evaluating Scientific Evidence: An Interdisciplinary Framework for Intellectual Due Process (Cambridge: Cambridge University Press, 2007), p.8.

81 e.g. D. Taylor et al, "An illustration of the Effect of Various Sources of Uncertainty on DNA Likelihood Ratio Calculations" (2014) 11 Forensic Science International: Genetics 56; P. Gill et al, "Interpretation of Complex DNA Profiles Using Empirical Models and a Method To Measure Their Robustness" (2008) 2 Forensic Science International: Genetics 91.

82 As famously described by Thomas Kuhn in The Structure of Scientific Revolutions 4th edn (Chicago: Chicago University Press, 2012). See also, e.g. Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, para.3.104.

83 Beecher-Monas, Evaluating Scientific Evidence (2007), pp.3, 8–9.

84 See fn.148 below.

85 e.g. many forensic sciences studied at university are—as we shall see—less reliable than, for instance, cooking or hairdressing techniques.

86 See, e.g. Nicolson, Evidence and Proof in Scotland (2019), pp.191–198 passim; Beecher-Monas, Evaluating Scientific Evidence (2007), p.37ff.

87 See Roberts, "Science, Experts, and Criminal Justice" in The Handbook of the Criminal Justice Process (2002), p.263: "As a rough guide, any scientific knowledge or theory that is likely to require material revision before a defendant has served his gaol sentence is unlikely to be considered an acceptable basis for criminal conviction and punishment."

88 Including by Daubert itself: Daubert 509 U.S. 579 (1993) at 593. See also, e.g. Edmond and Roach, "A Contextual Approach to the Admissibility of The State's Forensic Science and Medical Evidence" (2011) 61 University of Toronto Law Journal 343, 400; Faigman, "Expert Evidence: The Rules and the Rationality the Law Applies (or Should Apply) to Psychological Expertise" in Handbook of Psychology in Legal Contexts (2003), p.374; Leiter, "The Epistemology of Admissibility: Why Even Good Philosophy of Science Would Not Make for Good Philosophy of Evidence" (1997) Brigham Young University Law Review 803, 804, 818–819; B. Black, F.J. Ayala, and C. Saffran-Brinks, "Science and the Law in the Wake of Daubert: A New Search for Scientific Knowledge" (1994) 72 Texas L. Rev. 715, 718–719.

89 See text at fn.82, above.

90 Daubert 509 U.S. 579 (1993) at 593–594.

91 Young, 2014 S.L.T. 21 at [54].

92 See, e.g. S.A. Cole, "Forensic Culture as Epistemic Culture: The Sociology of Forensic Science" (2013) 44 Studies in History and Philosophy of Biology & Biomedical Science 36, 40; D. Crane, "The Gatekeepers of Science: Some Factors Affecting the Selection of Articles for Scientific Journals" (1967) 2 The American Sociologist 195.
93 Daubert 509 U.S. 579 (1993) at 593.
94 See Black, Ayala, and Saffran-Brinks, "Science and the Law in the Wake of Daubert: A New Search for Scientific Knowledge" (1994) 72 Texas L. Rev. 715, 731 noting that a hydrogeology study of a specific aquifer might be very significant in a particular case but not be

publishable as advancing scientific knowledge. See also Moreno, "Eyes Wide Shut: Hidden Problems and Future Consequences of the Fact-Based Validity Standard" (2003) 34 Seton Hall Law Review 89, 99–100.
95 R.M. Wheate and A. Jamieson, "A Tale of Two Approaches—The NAS Report and the Law Commission Consultation Paper on Forensic Science" (2009) 7 International Commentary on Evidence 1, 5.
96 Faigman, "Expert Evidence: The Rules and the Rationality the Law Applies (or Should Apply) to Psychological Expertise" in Handbook of Psychology in Legal Contexts (2003), p.382; Black, Ayala, and Saffran-Brinks, "Science and the Law in the Wake of Daubert: A New Search for Scientific Knowledge" (1994) 72 Texas L. Rev. 715, 778.
97 The classic discussion is that of D.T Campbell and J.C. Stanley, Experimental and Quasi-Experimental Designs for Research (Boston: Houghton Mifflin Co, 1966).
98 See, e.g. J. Howick, The Philosophy of Evidence-Based Medicine (Chichester: Wiley-Blackwell, 2011), Chs 4–8. 99 Young, 2014 S.L.T. 21 at [57].
100 Young, 2014 S.L.T. 21 at [28].
101 Young, 2014 S.L.T. 21 at [25] (emphasis added).
102 cf. for example, G. Edmond, "Re-assessing Reliability" in P. Roberts and M. Stockdale, Forensic Science and Expert Testimony (Cheltenham: Edgar Elgar Publishing, 2018), pp.101–103.
103 See Young, 2014 S.L.T. 21 at [58]: "Perhaps the aspect of CLS which is of most concern …".
104 See, e.g. Faigman, "Expert Evidence: The Rules and the Rationality the Law Applies (or Should Apply) to Psychological Expertise" in Handbook of Psychology in Legal Contexts (2003), pp.380–381; Developments in the Law, "Confronting the New Challenges of Scientific Evidence" (1994–1995) 108 Harv. L. Rev. 1481, 1540ff; M. Grabe, Measurement Uncertainties in Science and Technology (Berlin: Springer, 2005), Ch.1.
105 See, e.g. Nicolson, Evidence and Proof in Scotland (2019), pp.55–63.
106 K. Martire and G. Edmond, "Rethinking Expert Opinion Evidence" (2017) 40 Melbourne University Law Review 967, 988. 107 Young, 2014 S.L.T. 21 at [54].
108 Young, 2014 S.L.T. 21 at [57]
109 A. Kershaw, "Professional Standards, Public Protection and the Administration of Justice" in Handbook of Forensic Science (2009).
110 See Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, para.5.27.
111 cf. Young, 2014 S.L.T. 21 at [57] where the HCJ refers to "the lack of any means whereby a fact finder may weigh and assess the evidence in a particular case".
112 cf. R.J. Allen, "Expertise and the Daubert Decision" (1993) 84 Journal of Criminal Law and Criminology 1157, especially 1174–1175, who criticises Daubert for failing to require expertise which can be understood by fact-finders who must therefore decide to either totally defer to, or reject, the expert testimony and can do so only on irrationality grounds.
113 See also the positive reference to evidence like DNA, etc., which "may be weighed and assessed by a finder of fact": Young, 2014 S.L.T. 21 at [55].
114 Unless, by "challenge" the HCJ was thinking of that mounted in court, in which case it was again seeking to protect fact finders' authority.

115 Daubert 509 U.S. 579 (1993) at 593–594, where the court refer to principles which "can be … tested".
116 The following discussion of falsifiability draws on Allen, "Expertise and the Daubert Decision" (1993) 84 Journal of Criminal Law and Criminology 1157, 1169–1171; Edmond and Mercer, "Keeping 'Junk' History, Philosophy and Sociology of Science Out of the Courtroom: Problems with the Reception of Daubert v Merrell Dow Pharmaceuticals Inc" (1997) 20 University of New South Wales Law Journal 48, 82–97; A. Schwartz, "A 'Dogma of Empiricism' Revisited: Daubert v. Merrell Dow Pharmaceuticals, Inc. and the Need to Resurrect the Philosophical Insight of Frye v. United States" (1997) 10 Harv. J. L. & Tech. 149, 165–192.
117 Hume's Treatise of Human Nature, edited by L.A. Selby Bigge (Oxford: Clarendon Press, 1978), Book I, Pt III, section VI.
118 e.g. K. Popper, The Logic of Scientific Discovery (Abingdon: Routledge, 1992).
119 See J.T. Richardson et al, "The Problems of Applying Daubert to Psychological Syndrome Evidence" (1995) 79 Judicature 10, 13.
120 See the references in fn.116, above.
121 J. Ziman, Reliable Knowledge: An Exploration of the Grounds for Belief in Science (Cambridge: Cambridge University Press, 1978), pp.34–35, and more generally at pp.56–70, on the importance of scientific consensus to establish experimental verification.
122 See "Testing, testing, testing", above; see also section V, below.
123 e.g. climate scientists and epidemiologists substitute the paradigmatic scientific method of experimentation with "statistical analyses of populations, reviews of long-term trend data, clinical studies of illness in individuals, observations of organizational behaviour, computer simulations, and even historical, literary, or cultural records": S. Jasanoff, "Law's Knowledge: Science for Justice in Legal Settings" (2005) 95 American Journal of Public Health, S49, S54. See also Redmayne, Expert Evidence and Criminal Justice (2001), pp.114ff; Faigman, "Expert Evidence: The Rules and the Rationality the Law Applies (or Should Apply) to Psychological Expertise" in Handbook of Psychology in Legal Contexts (2003), p.387.
124 See R.K. Merton, The Sociology of Science: Theoretical and Empirical Investigations, edited by N.W. Storer (Chicago: University of Chicago Press, 1973), Ch.13.
125 In addition to references in fn.116, above, see, e.g. Nicolson, Evidence and Proof in Scotland (2019), pp.191–198 passim, especially p.195; Beecher-Monas, Evaluating Scientific Evidence (2007), especially Ch.3; M. Mulkay, Science and the Sociology of Knowledge (Abingdon: Routledge, 1985), pp.50–59, 76ff.
126 W.S. Bainbridge and M.C. Roco (eds), Handbook of Science and Technology Convergence (Switzerland: Springer, 2016).
127 See, e.g. Mulkay, Science and the Sociology of Knowledge (1985), especially Ch.2; D. Crump, "The Trouble with Daubert-Kumho: Reconsidering the Supreme Court's Philosophy of Science" (2003) 68 Missouri Law Review 1, 32ff.
128 See "Young and maturity" above; but cf. Schwartz, "A 'Dogma of Empiricism' Revisited: Daubert v. Merrell Dow Pharmaceuticals, Inc. and the Need to Resurrect the Philosophical Insight of Frye v. United States" (1997) 10 Harv. J. L. & Tech. 149, 206ff who provides a much more elaborate scheme.
129 See, e.g. Jasanoff, "Law's Knowledge: Science for Justice in Legal Settings" (2005) 95 American Journal of Public Health, S49, S49–S53 and "What Judges Should Know about the Sociology of Science" (1993) 77 Judicature 77; Edmond and Mercer, "Keeping 'Junk' History, Philosophy and Sociology of Science Out of the Courtroom: Problems with the Reception of Daubert v Merrell Dow Pharmaceuticals Inc" (1997) 20 University of New South Wales Law Journal 48, 49, 94ff; Leiter, "The Epistemology of Admissibility: Why Even Good Philosophy of Science Would Not Make for Good Philosophy of Evidence" (1997) Brigham Young University Law Review 803, 815–817; Crump, "The Trouble with Daubert-Kumho: Reconsidering the Supreme Court's Philosophy of Science" (2003)

68 Missouri Law Review 1. See also Allen, "Expertise and the Daubert Decision" (1993) 84 Journal of Criminal Law and Criminology 1157; the articles in (1994) 15 Cardozo Law Review 1745.
130 Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales.
131 Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, Appendix A, cl.4.
132 Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, Appendix A, cl.4.
133 Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, Appendix A, Schedule para.1.
134 Procurator Fiscal, Alloa v C, unreported, 1 December 2014, Alloa Sheriff Court.
135 cf. Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, paras 3.48–3.59, relegating this to a long-term goal.
136 I. Dennis, Law of Evidence, 6th edn (London: Sweet and Maxwell, 2017), p.897.
137 See Criminal Practice Directions 2015 at "V Evidence Part 19A Expert Evidence", available at https://www.judiciary.uk/wp-content/ uploads/2015/09/crim-pd-2015.pdf [Accessed 9 November 2020].
138 Consciously so: see Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, para.3.30.
139 See, e.g. O. Scallavaci, The Impact of Scientific Evidence on the Criminal Trial: The Case of DNA Evidence (Abingdon: Routledge, 2014), p.84ff; G. Edmond, "Is Reliability Sufficient—The Law Commission and Expert Evidence in International and Interdisciplinary Perspective (Part 1)" (2012) 16 International Journal of Evidence & Proof 30; but cf. the more positive reception by T. Ward, "'A New and More Rigorous Approach' to Expert Evidence in England and Wales" (2015) 19 International Journal of Evidence and Proof 228.
140 Though it is unclear why it regarded certain factors as "higher" rather than "lower order", or vice versa, and in fact the Practice Directions reverse the order of each level.
141 Except in cases where "experts" provide testimony of fact rather than opinion which is based on actual observations of relevant phenomena gained through actual experience rather than study or training.
142 See Lady Scott, "Submissions for the Ninth Programme of Law Reform" (16 August 2014), https://www.scotlawcom.gov.uk/ files/7114/3161/1713/Lady_Scott.pdf [Accessed 9 November 2020].
143 See the definition in fn.1, above.
144 Thus, as argued in section V below, many forensic "sciences" scarcely deserve that honorific: see especially at fn.174 below.
145 As noted at fnn.11–12 above, the courts have always been far more reluctant to admit the evidence of social scientists. 146 Young, 2014 S.L.T. 21 at [55].
147 The following discussion of DNA profiling draws primarily on S. Carr et al, "Clarifying the 'Reliability' Continuum and Testing its Limits: Biometetric (Fingerprint and DNA) Expert Evidence" in Forensic Science and Expert Testimony (2018), pp.173–181 and US President's Council of Advisors on Science and Technology, Report to the President Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods, pp.69–82. For more detail, see, e.g. P. Gill and T. Clayton, "The Current Status of DNA Profiling in the UK" in Handbook of Forensic Science (2009); J.M. Taupin, Introduction to Forensic DNA Evidence for Criminal Justice Professionals (London: CRC Press, 2013).
148 See, e.g. Edmond et al, "Admissibility Compared: The Reception of Incriminating Expert Evidence (i.e., Forensic Science) in Four Adversarial Jurisdictions (2014) 3 University of Denver Criminal Law Review 31, 45ff.
149 But less so in relation to population sub-groups: Edmond et al, "Admissibility Compared: The Reception of Incriminating Expert Evidence (i.e., Forensic Science) in Four Adversarial Jurisdictions (2014) 3 University of Denver Criminal Law Review 31, 35.

150 J. Weber-Lehmann et al, "Finding the Needle in the Haystack: Differentiating 'Identical' Twins in Paternity Testing and Forensics by Ultra-Deep Next Generation Sequencing" (2013) 9 Forensic Science International: Genetics 42.
151 In 2015, Scotland began testing 21 STR loci and three sex markers in response to recommendations by the European Network of Forensic Science Institutes and the European DNA Profiling Group.
152 The UKDNA database contains approximately 6 million profiles as at 2020: Home Office, "National DNA Database statistics", https:// www.gov.uk/government/statistics/national-dna-database-statistics [Accessed 9 November 2020].
153 C. Lawless, "The Low Template DNA Profiling Controversy: Biolegality and Boundary Work Among Forensic Scientists" (2013) 43 Social Studies of Science 191.
154 As opposed to emerging techniques such as mitochondrial DNA analysis, which have not reached the same levels of ubiquity, testing or acceptance.
155 Committee on Identifying the Needs of the Forensic Sciences Community, National Research Council, Strengthening Forensic Science in the United States: A Path Forward (National Academies Press, 2009), p.128.
156 The following draws on the concise analyses in Carr et al, "Clarifying the 'Reliability' Continuum and Testing its Limits: Biometetric (Fingerprint and DNA) Expert Evidence" in Forensic Science and Expert Testimony (2018), pp.165–173, US President's Council of Advisors on Science and Technology, Report to the President Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods, pp.165–167. For more detail, see e.g. Campbell, The Fingerprint Inquiry Report (2011).
157 Carr et al, "Clarifying the 'Reliability' Continuum and Testing its Limits: Biometetric (Fingerprint and DNA) Expert Evidence" in Forensic Science and Expert Testimony (2018), p.155.
158 See, e.g. C. Champod and P. Chamberlain, "Fingerprints" in Handbook of Forensic Science (2009), pp.78–79.
159 Campbell, The Fingerprint Inquiry Report (2011), Ch.38.
160 Campbell, The Fingerprint Inquiry Report (2011), Chs 33 and 35.
161 See, e.g. the notorious wrongful conviction of Shirley McKie investigated by Campbell, The Fingerprint Inquiry Report (2011), and Cole, "Grandfathering Evidence: Fingerprint Admissibility Rulings from Jennings to Llera Plaza and Back Again" (2004) 41 American Criminal Law Review 1189, 1206, regarding three other UK miscarriages of justice involving this standard.
162 J.J. Koehler, "Fingerprint Error Rates and Proficiency Tests: What They Are and Why They Matter" (2008) 59 Hastings L.J. 1077, 1090. See S.A. Cole "More than Zero: Accounting for Error in Latent Fingerprint Identification" (2004) 95 Journal of Criminal Law and Criminology 985, 1039.
163 Discussed by US President's Council of Advisors on Science and Technology, Report to the President Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods, pp.94–95.
164 US President's Council of Advisors on Science and Technology, Report to the President Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods, p.101.
165 G. Edmond et al, "Contextual Bias and Cross-Contamination in the Forensic Sciences: The Corrosive Implications for Investigations, Plea Bargains, Trials and Appeals" (2015) 14 Law, Probability and Risk 1, 2.

166 Edmond et al, "Contextual Bias and Cross-Contamination in the Forensic Sciences: The Corrosive Implications for Investigations, Plea Bargains, Trials and Appeals" (2015) 14 Law, Probability and Risk 1. See also Redmayne, Expert Evidence and Criminal Justice (2001), pp.13–16; I.E. Dror and S.A. Cole, "The Vision in 'Blind' Justice: Expert Perception, Judgment and Visual Cognition in Forensic Pattern Recognition"(2010) 17 Psychonomic Bulletin & Review 161; I.E. Dror, D.

Charlton and A.E. Peron, "Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications" (2006) 156 Forensic Science International 74.

167 See Nicolson, Evidence and Proof in Scotland (2019), pp.208–209.

168 Committee on Identifying the Needs of the Forensic Sciences Community, National Research Council, Strengthening Forensic Science in the United States: A Path Forward, p.100.

169 Thus, US President's Council of Advisors on Science and Technology, Report to the President Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods, p.101, rather cautiously exempted fingerprinting analysis from the above conclusion by the National Research Council.

170 See, e.g. Nicolson, Evidence and Proof in Scotland (2019), pp.194–197.

171 Committee on Identifying the Needs of the Forensic Sciences Community, National Research Council, Strengthening Forensic Science in the United States: A Path Forward, p.52.

172 See text at fn.148, above.

173 Committee on Identifying the Needs of the Forensic Sciences Community, National Research Council, Strengthening Forensic Science in the United States: A Path Forward, p.128.

174 W.C. Thompson, "A Sociological Perspective on the Science of Forensic Testing" (1997) 30 University of California Davis Law Review
1113, 1131.

175 cf. M. Solomon, Making Medical Knowledge (Oxford: Oxford University Press, 2015), "Introduction".

176 E. Imwinkelreid, "Shaken Baby Syndrome: A Genuine Battle of the Scientific (and Non-Scientific) Experts" (2010) 46 Criminal Law Bulletin 156.

177 See, e.g. Howick, The Philosophy of Evidence-Based Medicine (2011), especially Ch.2; J. Ridderikhoff, Methods in Medicine (Dordrecht, The Netherlands: Kluwer Academic Publishers, 1989), especially Ch.1.

178 e.g. R.J. Smith, "Where is the Wisdom? The Poverty of Medical Evidence" (1991) 303 B.M.J. 798 estimated that only 15% of medical interventions were based on solid scientific evidence.

179 See, e.g. Howick, The Philosophy of Evidence-Based Medicine (2011), especially Chs 10 and 11.

180 D.G. Altman, "The Scandal of Poor Medical Research" (1994) 308 B.M.J. 283. See also, e.g. Howick, The Philosophy of Evidence- Based Medicine (2011), p.53.

181 This seems to be implicit in the use of the term "medical" and the fact that the courts are well aware of the often controversial nature of assessments of an accused and victim's states of mind: see F.E. Raitt and M.S. Zeedyk, The Implicit Relation of Psychology and Law: Women and Syndrome Evidence (Abingdon: Rouledge, 2000).

182 See P.C. Giannelli et al, Scientific Evidence, 5th edn (London: LexisNexis, 2012), Vol.2, especially para.19.04.

183 See text at fn.166, above.

184 See text at fn.166, above.

185 Giannelli et al, Scientific Evidence (2012), Vol.2, para.19.10(b).

186 See Imwinkelreid, "Shaken Baby Syndrome: A Genuine Battle of the Scientific (and Non-Scientific) Experts" (2010) 46 Criminal Law Bulletin 156.

187 Giannelli et al, Scientific Evidence (2012), Vol.2, para.19.01.

188 Edmonds et al, "Admissibility Compared: The Reception of Incriminating Expert Evidence (i.e., Forensic Science) in Four Adversarial Jurisdictions" (2014) 3 University of Denver Criminal Law Review 31, 33. For similar studies see, e.g. Developments in the Law, "Confronting the New Challenges of Scientific Evidence" (1994–95) 108 Harv. L. Rev. 1481, 1492ff; J. Groscup et al, "The Effects of Daubert on the Admissibility of Expert Testimony in State and Federal Criminal Cases" (2002) 8 Psychology, Public Policy, and Law 339.

189 See E. Cunliffe, "A New Canadian Paradigm? Judicial Gatekeeping and the Reliability of Expert Evidence" in Forensic Science and Expert Testimony (2018).

190 Frye 297 F. 1013 (D.C. Cir 1923); Bonython (1984) 38 SASR 45, discussed at fnn.29, 42 and 77–79 above.

191 e.g. Campbell, The Fingerprint Inquiry Report (2011), has yet to be cited in a Scottish court, though it has been cited in the English case of R. v Smith [2011] EWCA Crim 1296; [2011] 2 Cr. App. R. 16. For the judicial response in other jurisdictions to reports like those cited at fnn.147 and 155 above, see Edmond, "Re-assessing Reliability" in Forensic Science and Expert Testimony (2018), pp.89–96;
S.A. Cole and G. Edmond, "Science without Precedent: The Impact of the National Research Council Report on the Admissibility and Use of Forensic Science Evidence" (2015) 4 British Journal of American Legal Studies 585.

192 Edmond et al, "Admissibility Compared: The Reception of Incriminating Expert Evidence (i.e., Forensic Science) in Four Adversarial Jurisdictions" (2014) 3 University of Denver Criminal Law Review 31, especially at 91–96; Edmond, "Re-assessing Reliability"
in Forensic Science and Expert Testimony (2018), pp.73–82, 102, G. Edmond, "Forensic Science Evidence and the Conditions for Rational (Jury) Evaluation" (2015) 39 Melbourne University Law Review 77, 95–99; Martire and Edmond, "Rethinking Expert Opinion Evidence" (2017) 40 Melbourne University Law Review 967, 984–987.

193 A factor observed by one Scottish judge (Lady Scott, "Submissions for the Ninth Programme of Law Reform" (16 August 2014), https://
www.scotlawcom.gov.uk/files/7114/3161/1713/Lady_Scott.pdf [Accessed 9 November 2020]), and one which advantages the Scottish prosecution as the state employs the most experienced practitioners.

194 A practice confirmed in Scotland by Lady Scott, "Submissions for the Ninth Programme of Law Reform" (16 August 2014), p.6, https://
www.scotlawcom.gov.uk/files/7114/3161/1713/Lady_Scott.pdf [Accessed 9 November 2020].

195 More justifiable is the greater receptivity to expert evidence which allows fact-finders to examine for themselves the facts upon which it is based.

196 An additional one—being more amenable to expert evidence where it is the sole evidence relied upon—should be less problematic in Scotland given the corroboration requirement.

197 In addition to the references at fn.192 above, see Developments in the Law, "Confronting the New Challenges of Scientific Evidence" (1994–95) 108 Harv. L. Rev. 1481, 1500–1507.

198 D. Luban, Lawyers and Justice: An Ethical Study (Princeton, New Jersey: Princeton University Press, 1988), pp.60–63.

199 Roberts and Zuckerman, Criminal Evidence (2010), p.19.

200 See Lady Scott, "Submissions for the Ninth Programme of Law Reform" (16 August 2014), p.5, https://www.scotlawcom.gov.uk/ files/7114/3161/1713/Lady_Scott.pdf [Accessed 9 November 2020]; Edmond, "Is Reliability Sufficient—The Law Commission and Expert Evidence in International and Interdisciplinary Perspective (Part 1)" (2012) 16 International Journal of Evidence and Proof 30, 52, 56.

201 Edmond, "Forensic Science Evidence and the Conditions for Rational (Jury) Evaluation" (2015) 39 Melbourne University Law Review
77, 108.

202 D.L. Faigman, "Anecdotal Forensics, Phrenology, and Other Abject Lessons From the History of Science" (2008) 59 Hastings
L.J. 979, 992, 997. See also, e.g. Edmond et al, "Admissibility Compared: The Reception of Incriminating Expert Evidence (i.e., Forensic Science) in Four Adversarial Jurisdictions" (2014) 3 University of Denver Criminal Law Review 31, 33; Edmond, "Forensic Science Evidence and the Conditions for Rational (Jury) Evaluation" (2015) 39 Melbourne University Law Review 77, 100 and "Is Reliability Sufficient—The Law Commission and Expert Evidence in International and Interdisciplinary Perspective (Part 1)" (2012) 16 International Journal of Evidence and Proof 30, 55.

203 D. McQuiston-Surrett and M. Saks, "Communicating Opinion Evidence in the Forensic Identification Sciences: Accuracy and Impact" (2008) 59 Hastings L.J. 1159.
204 Edmond, "Forensic Science Evidence and the Conditions for Rational (Jury) Evaluation" (2015) 39 Melbourne University Law Review
77, 100.
205 Edmond et al, "Admissibility Compared: The Reception of Incriminating Expert Evidence (i.e., Forensic Science) in Four Adversarial Jurisdictions" (2014) 3 University of Denver Criminal Law Review 31, 94 and 98; Lady Scott, "Submissions for the Ninth Programme of Law Reform" (16 August 2014), pp.6–7, https://www.scotlawcom.gov.uk/files/7114/3161/1713/Lady_Scott.pdf [Accessed 9 November 2020]. See, in relation to science more generally, e.g. Cole "More than Zero: Accounting for Error in Latent Fingerprint Identification" (2004) 95 Journal of Criminal Law and Criminology 985, 1034ff; Jones, Expert Witnesses: Science, Medicine and the Practice of Law (1994), especially pp.13–14, 97–101; W.C. Thompson, "Beyond Bad Apples: Analyzing the Role of Forensic Science in Wrongful Convictions" (2008) 37 Southwestern University Law Review 1027.
206 See Edmond et al, "Admissibility Compared: The Reception of Incriminating Expert Evidence (i.e., Forensic Science) in Four Adversarial Jurisdictions" (2014) 3 University of Denver Criminal Law Review 31, 96; Edmond, "Re-assessing Reliability" in Forensic Science and Expert Testimony (2018), p.98.
207 See, e.g. J.S. Oteri, M.G. Weinberg and M.S. Pinales, "Cross-Examination of Chemists in Narcotic and Marijuana Cases" (1973) 2 Contemporary Drug Problems 225; J.M. Shellow, "The Limits of Cross-Examination" (2003) 34 Seton Hall L. Rev. 317.
208 For anecdotal support, see Lady Scott, "Submissions for the Ninth Programme of Law Reform" (16 August 2014), pp.6 and 11, https://
www.scotlawcom.gov.uk/files/7114/3161/1713/Lady_Scott.pdf [Accessed 9 November 2020].
209 See text at fnn.222–227 below, passim.
210 While there is no compulsory judicial training in the UK, there are primers on various aspects of scientific evidence, "Courtroom science primers launched today" (The Royal Society, 22 November 2017), https://royalsociety.org/news/2017/11/royal-society-launches- courtroom-science-primers/ [Accessed 9 November 2020], as in other Anglo-American jurisdictions, some of which also provide training: see Hon. T.A. Cromwell, "The Challenges of Scientific Evidence" (Scottish Council of Law Reporting, 2 March 2011), https:// scottishlawreports.org.uk/publications/macfadyen-2011.html [Accessed 28 August 2020].
211 See, e.g. Groscup et al, "The Effects of Daubert on the Admissibility of Expert Testimony in State and Federal Criminal Cases" (2002) 8 Psychology, Public Policy, and Law 339; S. Gatowski et al, "Asking the Gatekeepers: A National Survey of Judges on Judging Expert Evidence in a Post-Daubert World" (2001) 25 Law & Human Behaviour 433; M.B. Kovera and B.D. McAuliff, "The Effects of Peer Review and Evidence Quality on Judge Evaluations of Psychological Science: Are Judges Effective Gatekeepers?" (2000) 85 Journal of Applied Psychology 574; M.B. Kovera, M.B. Russano, and B.D. McAuliff, "Assessment of the Commonsense Psychology Underlying Daubert: Legal Decision Makers' Abilities to Evaluate Expert Evidence in Hostile Work Environment Cases" (2002) 8 Psychology, Public Policy, and Law 180.
212 See D.M. Risinger, "Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?" (2000) 64(1) Albany L. Rev. 99, 143.
213 Faigman, "Anecdotal Forensics, Phrenology, and Other Abject Lessons From the History of Science" (2008) 59 Hastings L.J. 979, 992; see also at 992–993, 999.
214 Martire and Edmond, "Rethinking Expert Opinion Evidence" (2017) 40 Melbourne University Law Review 967, 996.
215 Faigman, "Anecdotal Forensics, Phrenology, and Other Abject Lessons From the History of Science" (2008) 59 Hastings L.J. 979, 994–995; Cole, "Grandfathering Evidence: Fingerprint Admissibility Rulings from Jennings to Llera Plaza and Back Again" (2004) 41 American Criminal Law Review 1189, 1211.

216 Faigman, "Anecdotal Forensics, Phrenology, and Other Abject Lessons From the History of Science" (2008) 59 Hastings L.J. 979, 992– 993; Beecher-Monas, Evaluating Scientific Evidence (2007), p.95.
217 See A. Sanders, R. Young and M. Burton, Criminal Justice, 4th edn (Oxford: Oxford University Press, 2010), p.21ff, discussing Herbert Packer's famous two models of criminal justice outlined in The Limits of the Criminal Sanction (Stanford, California: Stanford University Press, 1968).
218 See fnn.198 and 199, above.
219 See, e.g. Edmond et al, "Admissibility Compared: The Reception of Incriminating Expert Evidence (i.e., Forensic Science) in Four Adversarial Jurisdictions" (2014) 3 University of Denver Criminal Law Review 31, 102; Edmond, "Is Reliability Sufficient—The Law Commission and Expert Evidence in International and Interdisciplinary Perspective (Part 1)" (2012) 16 International Journal of Evidence and Proof 30, 45; K.A. Findlay, "Innocents at Risk: Adversary Imbalance, Forensic Science and the Search for Truth" (2008) 38 Seton Hall L. Rev. 893, 896; Risinger, "Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?" (2000) 64(1) Albany L. Rev. 99, 99; Edmond and Roach, "A Contextual Approach To the Admissibility of The State's Forensic Science and Medical Evidence" (2011) 61 University of Toronto Law Journal 343, especially 358, 396, 398.
220 See Duff and Ferguson (eds), Current Developments in Scottish Criminal Evidence Law (2018), passim.
221 This may be partly due to the fact that defence lawyers tend to rely on experts from the "soft" rather than "hard" sciences which as noted above and at fnn.11–12 have traditionally attracted greater judicial scepticism.
222 Most comprehensively detailed by G. Edmond and M. San Roque, "The Cool Crucible: Forensic Science and the Frailty of the Criminal Trial" (2012) 24 Current Issues in Criminal Justice 51. See also Edmond, "Re-assessing Reliability" in Forensic Science and Expert Testimony (2018), pp.94–95, 101–103 and Edmond, "Is Reliability Sufficient—The Law Commission and Expert Evidence in International and Interdisciplinary Perspective (Part 1)" (2012) 16 International Journal of Evidence and Proof 30, 49–51; Kovera,

Russano, and McAuliff, "Assessment of the Commonsense Psychology Underlying Daubert: Legal Decision Makers' Abilities to Evaluate Expert Evidence in Hostile Work Environment Cases" (2002) 8 Psychology, Public Policy, and Law 180; M. Kovera, B. McAuliff and K. Hebert, "Reasoning About Scientific Evidence: Effects of Juror Gender and Evidence Quality on Juror Decisions in a Hostile Work Environment Case" (1999) 84 Journal of Applied Psychology 362.
223 In addition to references at fn.222 above, see, e.g. D. McQuiston-Surrett and M.J. Saks, "The Testimony of Forensic Identification Science: What Expert Witnesses Say and What Factfinders Hear" (2009) 33 Law and Human Behavior 436; J. Sanders, "The Merits of Paternalistic Justification for Restrictions on the Admissibility of Expert Evidence" (2003) Seton Hall L. Rev. 881, 932–936.
224 See, e.g. Findlay "Innocents at Risk: Adversary Imbalance, Forensic Science and the Search for Truth" (2008) 38 Seton Hall L. Rev. 893; Roberts, "Science, Experts, and Criminal Justice" in The Handbook of the Criminal Justice Process (2002), pp.274–276 and "Science in the Criminal Process" (1994) 14 O.J.L.S. 469, 489–495.
225 Roberts, "Science in the Criminal Process" (1994) 14 O.J.L.S. 469, 491
226 See, e.g. R. v Weller [2010] EWCA Crim 1085 at [23]–[25].
227 Nicolson, Evidence and Proof in Scotland (2019), pp.345–353,
228 Developments in the Law, "Confronting the New Challenges of Scientific Evidence" (1994–95) 108 Harv. L. Rev. 1481, 1587
229 cf. Edmond, "Forensic Science Evidence and the Conditions for Rational (Jury) Evaluation" (2015) 39 Melbourne University Law Review 77, 104.

230 In addition to references in fn.222 above, see McQuiston-Surrett and Saks, "The Testimony of Forensic Identification Science: What Expert Witnesses Say and What Factfinders Hear" (2009) 33 Law and Human Behavior 436; Martire and Edmond, "Rethinking Expert Opinion Evidence" (2017) 40 Melbourne University Law Review 967, 989.
231 cf. Edmond, "Forensic Science Evidence and the Conditions for Rational (Jury) Evaluation" (2015) 39 Melbourne University Law
Review 77, 94.
232 See, e.g. V.G. Rose et al, "Evaluating the Comprehensibility of Jury Instructions: A Method and an Example" (2001) 25 Law and Human Behavior 409.
233 Edmond et al, "Admissibility Compared: The Reception of Incriminating Expert Evidence (i.e., Forensic Science) in Four Adversarial Jurisdictions" (2014) 3 University of Denver Criminal Law Review 31, 100.
234 See Nicolson, Evidence and Proof in Scotland (2019), pp.218–223 regarding experts in the adversarial system and at pp.133–151 regarding problems with the adversarial system more generally
235 See Edmond, "Forensic Science Evidence and the Conditions for Rational (Jury) Evaluation" (2015) 39 Melbourne University Law
Review 77, especially 82.
236 See references cited in fnn.13 and 14 above.
237 See, e.g. Edmond, "Forensic Science Evidence and the Conditions for Rational (Jury) Evaluation" (2015) 39 Melbourne University Law Review 77, 110–119; K. Ivkovic and V.P. Hans, "Jurors Evaluations of Expert Testimony: Judging the Messenger and the Message" (2003) 28 Law & Social Inquiry 441; N. Vidmar and S.S. Diamond, "Juries and Expert Evidence" (2001) 66 Brooklyn L. Rev. 1121; B.D. McAuliff et al, "Juror Decision-making in the Twenty-First Century: Confronting Science and Technology in Court" in Handbook of Psychology in Legal Contexts (2003).
238 See e.g. Vidmar and Diamond, "Juries and Expert Evidence" (2001) 66 Brooklyn L. Rev. 1121, 1149–11158, 1163–1164, 1170–1171; Schklar and Diamond, "Juror Reactions to DNA Evidence: Errors and Expectancies" (1999) 23 Law and Human Behavior 159; K. Martire et al, "The Expression and Interpretation of Uncertain Forensic Science Evidence: Verbal Equivalence, Evidence Strength, and the Weak Evidence Effect" (2013) 37 Law and Human Behavior 197; K. Martire, R. Kemp and B. Newell, "The Psychology of Interpreting Expert Evaluative Opinions" (2013) 45 Australian Journal of Forensic Sciences 305; B. Robertson, G.A. Vignaux and
C.E.H. Berger, Interpreting Evidence: Evaluating Forensic Science in the Courtroom, 2nd edn (Chichester: Wiley, 2016), especially Ch.9.
239 See P. Hawkins and A. Hawkins, "Lawyers' Probability Misconceptions and the Implications for Legal Education" (1998) 18 L.S. 316; E. Greene and L. Ellis, "Decision Making in Criminal Justice"in D. Carson et al (eds), Applying Psychology to Criminal Justice (Chichester: Wiley, 2007), pp.184–185; C. Guthrie, J.J. Rachlinski and A.J. Wistrich, "Inside the Judicial Mind" (2000) 86 Cornell L. Rev. 777, noting, however, at 816–818 that in some respects judges performed better than other subjects.
240 Nicolson, Evidence and Proof in Scotland (2019), p.20.
241 See fnn.16–17 above.
242 On expert ethics, see T. Ward, "Ethics and the Role of the Expert" in The Expert Witness, Forensic Science, and the Criminal Justice Systems of the UK (2019).
243 cf. Lady Scott, "Submissions for the Ninth Programme of Law Reform" (16 August 2014), https://www.scotlawcom.gov.uk/ files/7114/3161/1713/Lady_Scott.pdf [Accessed 9 November 2020], who expresses support for many of the suggestions below.
244 cf. Edmond and San Roque, "The Cool Crucible: Forensic Science and the Frailty of the Criminal Trial" (2012) 24 Current Issues in Criminal Justice 51, 53 and also Edmond, "Re-assessing Reliability" in Forensic Science and Expert Testimony (2018), p.72.
245 cf. T. Ward et al, "Forensic Science, Scientific Validity and Reliability: Advice from America" [2017] Crim. L.R. 357, 365

246 See, e.g. O. Sallavaci, The Impact of Scientific Evidence on the Criminal Trial (Abingdon: Routledge, 2014), p.190; L. Hewson and J. Goodman-Delahunty, "Using Multimedia to Support Jury Understanding of DNA Profiling Evidence" (2008) 40 Australian Journal of Forensic Sciences 55
247 Vidmar and Diamond, "Juries and Expert Evidence" (2001) 66 Brooklyn L. Rev. 1121, 1136–1137, arguing that, even brief training can improve reasoning.
248 Sallavaci, The Impact of Scientific Evidence on the Criminal Trial (2014), p.135.
249 See Ross and Chalmers, Walker and Walker: The Law of Evidence (2015), p.304.
250 But see G. Edmond, "Judicial Representations of Scientific Evidence" (2000) 63 M.L.R. 216, 234, especially 242ff.
251 Vidmar and Diamond, "Juries and Expert Evidence" (2001) 66 Brooklyn L. Rev. 1121, 1179; Damaška, Evidence Law Adrift (1997), p.145
252 cf. Roberts, "Science, Experts, and Criminal Justice" in The Handbook of the Criminal Justice Process (2002), p.269.
253 See Martire and Edmond, "Rethinking Expert Opinion Evidence" (2017) 40 Melbourne University Law Review 967, 995; G. Edmond, "Merton and the Hot Tub: Scientific Conventions and Expert Evidence in Australian Civil Procedure" (2009) 72 Law and Contemporary Problems 159.
254 Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, para.5.96–5.98.

255 cf. Edmond, "Is Reliability Sufficient—The Law Commission and Expert Evidence in International and Interdisciplinary Perspective (Part 1)" (2012) 16 International Journal of Evidence and Proof 30, 52.
256 Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, paras 5.89–5.98.
257 For overviews, see e.g. Developments in the Law, "Confronting the New Challenges of Scientific Evidence" (1994–95) 108 Harv. L. Rev. 1481, 1590–1604; D.L. Burk, "When Scientists Act Like Lawyers: The Problem of Adversary Science" (1993) 33 Jurimetrics 363, 371–375; E. Di Lello, "Fighting Fire with Firefighters: A Proposal for Expert judges at the Trial Level" (1993) 93 Columbia Law Review 473.
258 As argued for by J. Hartshorne and J. Miola, "Expert Evidence: Difficulties and Solutions in Prosecutions for Infant Harm" (2010) 30 L.S. 279, 293.
259 On the relative merits of judges and jurors in this regard see, e.g. Edmond, "Forensic Science Evidence and the Conditions for Rational (Jury) Evaluation" (2015) 39 Melbourne University Law Review 77, 124; R. Lempert, "The Jury and Scientific Evidence" (1999) 9 The Kansas Journal of Law and Public Policy 22, 24ff; S. Brewer, "Scientific Expert Testimony and Intellectual Due Process" (1998) 107 Yale L.J. 1535, 1678, 1680; Black, Ayala, and Saffran-Brinks, "Science and the Law in the Wake of Daubert: A New Search for Scientific Knowledge" (1994) 72 Texas L. Rev. 715, 787–788.
260 See, e.g. McQuiston-Surrett and Saks, "Communicating Opinion Evidence in the Forensic Identification Sciences: Accuracy and Impact" (2008) 59 Hastings L.J. 1159, 1180; Vidmar and Diamond, "Juries and Expert Evidence" (2001) 66 Brooklyn L. Rev. 1121, especially 1173–1176.
261 G. Edmond, "The Next Step or Moonwalking? Expert Evidence, The Public Understanding of Science and The Case against Imwinkelried's Didactic Trial Procedures" (1998) 2 International Journal of Evidence and Proof 13, 22.
262 But cf. Damaška, Evidence Law Adrift (1997), p.144.
263 Lady Scott, "Submissions for the Ninth Programme of Law Reform" (16 August 2014), p.10, https://www.scotlawcom.gov.uk/ files/7114/3161/1713/Lady_Scott.pdf [Accessed 9 November 2020].
264 See Roberts, "Drawing on Expertise: Legal Decision-making and the Reception of Expert Evidence" [2008] Crim. L.R. 443, 459– 461, discussing the proposals by the English and Welsh Law Commission: Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, Pt 6. In addition to the references at fn.257 above, the merits and demerits of using neutral

experts are canvassed by, e.g. Jones, Expert Witnesses: Science, Medicine and the Practice of Law (1994), pp.37ff; Roberts and Zuckerman, Criminal Evidence (2010), pp.505–509; Roberts, "Science, Experts, and Criminal Justice" in The Handbook of the Criminal Justice Process (2002), pp.277–280; Faigman, "Expert Evidence: The Rules and the Rationality the Law Applies (or Should Apply) to Psychological Expertise" in Handbook of Psychology in Legal Contexts, (2003), pp.394–398.

265 cf. Nicolson, Evidence and Proof in Scotland (2019), pp.218–223.

266 Though this has not been confirmed by mock jury studies: Burk, "When Scientists Act Like Lawyers: The Problem of Adversary Science" (1993) 33 Jurimetrics 363, 374.

267 cf. the "rather cumbersome and labour-intensive" (Edmond, "Is Reliability Sufficient—The Law Commission and Expert Evidence in International and Interdisciplinary Perspective (Part 1)" (2012) 16 International Journal of Evidence and Proof 30, 38), procedure of a panel of experienced lawyers liaising with professional bodies to provide judges in particular cases with a list of suitable experts proposed by the English and Welsh Law Commission: Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, para.6.44ff.

268 Sallavaci, The Impact of Scientific Evidence on the Criminal Trial (2014), pp.91, 135, 188, but see the more optimistic recommendation of annual training by the Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, pp.181–182.

269 As raised by commentary on Daubert and in its minority judgement itself: see fn.129 above.

270 See, e.g. Findlay "Innocents at Risk: Adversary Imbalance, Forensic Science and the Search for Truth" (2008) 38 Seton Hall L. Rev. 893, 955–964; P.J. Neufeld, "The (Near) Irrelevance of Daubert to Criminal Justice and Some Suggestions for Reform" (2005) 95 American Journal of Public Health Supplement S107, S113; G. Edmond and A. Roberts. "Procedural Fairness, the Criminal Trial and Forensic Science and Medicine" (2011) 33 Sydney Law Review 359, 389–392; Sallavaci, The Impact of Scientific Evidence on the Criminal Trial (2014), pp.89–90, 188 (suggesting that in England and Wales this role should be played by the Forensic Science Regulator).

271 i.e. not where the expertise is truly and solely experience-based or capable of testing such as the knowledge of an art historian or comparative lawyer—see section II above.

272 See Edmond, "Forensic Science Evidence and the Conditions for Rational (Jury) Evaluation" (2015) 39 Melbourne University Law Review 77, especially at 96–98, 106; Edmond and San Roque, "The Cool Crucible: Forensic Science and the Frailty of the Criminal Trial" (2012) 24 Current Issues in Criminal Justice 51, 62.

273 See Beecher-Monas, Evaluating Scientific Evidence (2007), pp.10–11; D.E. Bernstein, "The Admissibility of Scientific Evidence after Daubert v. Merrell Dow Pharmacueticals, Inc." (1993) 15 Cardozo L. Rev. 2139 and the references cited in fnn.274 and 276 below.

274 Faigman, "Anecdotal Forensics, Phrenology, and Other Abject Lessons From the History of Science" (2008) 59 Hastings L.J. 979, 1000; Edmond and Roach, "A Contextual Approach To the Admissibility of The State's Forensic Science and Medical Evidence" (2011) 61 University of Toronto Law Journal 343, 408; T. Ward, "Expert Evidence and the Law Commission: Implementation without Legislation?" [2013] Crim. L.R. 561, 574.

275 E.K. Cheng and A.H. Yoon, "Does Frye or Daubert Matter? A Study of Scientific Admissibility Standards" (2005) 91 Va. L. Rev. 471, 472–473.

276 See E.I. Imwinkelreid, "Regulating Expert Evidence in US Courts: Measuring Daubert's Impact" in Forensic Science and Expert Testimony (2018).

277 L. Heffernan and M. Coen, "The Reliability of Expert Evidence: Reflections on the Law Commission's Proposals For Reform" (2009) 73 J. Crim. L. 488, 502–503.

278 Or perhaps only the prosecution—see text at fnn.280–281 below.

279 Edmond, "Forensic Science Evidence and the Conditions for Rational (Jury) Evaluation" (2015) 39 Melbourne University Law Review
77, 85ff.

280 Edmond, "Re-assessing Reliability" in Forensic Science and Expert Testimony (2018), pp.100–101.

281 See e.g. Developments in the Law, "Confronting the New Challenges of Scientific Evidence" (1994–95) 108 Harv. L. Rev. 1481, 1529– 1530, 1552; Redmayne, Expert Evidence and Criminal Justice (2001), p.136; Hartshorne and Miola, "Expert Evidence: Difficulties and Solutions in Prosecutions for Infant Harm" (2010) 30 L.S. 279, 295.

282 Edmond et al, "Admissibility Compared: The Reception of Incriminating Expert Evidence (i.e., Forensic Science) in Four Adversarial Jurisdictions" (2014) 3 University of Denver Criminal Law Review 31, 97; Lady Scott, "Submissions for the Ninth Programme of Law Reform" (16 August 2014), p.9, https://www.scotlawcom.gov.uk/files/7114/3161/1713/Lady_Scott.pdf [Accessed 9 November 2020].

283 cf. Law Commission, Report on Expert Evidence in Criminal Proceedings in England and Wales, paras 3.77, 5.42–5.61 which suggest a limited judicial power to disapply the reliability test so that it does not have to be applied routinely and unnecessarily.

284 Edmond, "Is Reliability Sufficient—The Law Commission and Expert Evidence in International and Interdisciplinary Perspective (Part 1)" (2012) 16 International Journal of Evidence and Proof 30, 69.

285 As implicitly recommended by Lady Scott, "Submissions for the Ninth Programme of Law Reform" (16 August 2014), p.10, https://
www.scotlawcom.gov.uk/files/7114/3161/1713/Lady_Scott.pdf [Accessed 9 November 2020].

286 See fn.206, above.

287 For an overview, see Nicolson, Evidence and Proof in Scotland (2019), pp.204–215 passim.

288 Edmond and Roach, "A Contextual Approach To the Admissibility of The State's Forensic Science and Medical Evidence" (2011) 61 University of Toronto Law Journal 343, 348.

289 Neufeld, "The (Near) Irrelevance of Daubert to Criminal Justice and Some Suggestions for Reform" (2005) 95 American Journal of Public Health Supplement S107, S111ff.

290 P. Kitcher, "Scientific Knowledge" in P.K. Moser (ed.), The Oxford Handbook of Epistemology (Oxford: Oxford University Press, 2002),
p.385. See also A.I. Goldman, Knowledge in a Social World (Oxford: Oxford University Press, 1999), Ch.8.

291 G. Edmond,"Just Truth? A Partial Archaeology of the Admissibility Revolution Associated With Daubert" (SSRN Electronic Journal, 10 July 2006), p.38,
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=916062 [Accessed 9 November 2020].

292 Nicolson, Evidence and Proof in Scotland (2019), Ch.6.

293 S. Haack, Manifesto of a Passionate Moderate (Chicago: University of Chicago Press, 1998), p.94 (emphasis in original); see also at pp.95, 1