Law, Probability and Risk (2003) **2**, 25–46

Expressing scientific uncertainty

CHARLES WEISS†
Walsh School of Foreign Service, Georgetown University, 37th and O Streets, NW, Washington, D.C. 20057, USA

[Received on 12 August 2002; revision received on 27 October 2002; accepted on 5 November 2002]

This paper proposes a subjective scale of scientific uncertainty that allows a source of scientific information to express to a lay audience the subjective level of certainty or uncertainty that it associates with a particular assertion of scientific fact, or to represent the range of expert opinion regarding that certainty or uncertainty. The scale is intended as a tool to help increase the precision and rationality of discourse in controversies in which generalists untrained in natural science must judge the merits of opposing arguments in disputes among scientific experts. It complements the quantitative scale of uncertainty, based on Bayesian statistics, used in the recent report of the Inter-Governmental Panel on Climate Change. Both of these scales are designed for use in situations where the risk probabilities are not precisely known.

The scale takes advantage of the fact that there are many more standards of proof recognized in the US legal system beyond the familiar 'criminal' and 'civil' standards of 'beyond a reasonable doubt' and 'preponderance of the evidence', respectively, and that these standards correspond to levels of certainty or uncertainty that constitute acceptable bases for legal decisions in a variety of practical contexts. The levels of certainty or uncertainty corresponding to these standards of proof correspond rather well to the informal scale of certainty used by research scientists in the course of their everyday work, and indeed by ordinary people as they estimate the likelihood of one or another proposition.

Keywords: risk uncertainty; standards of proof; scale of uncertainty; levels of uncertainty; scientific uncertainty.

1. Introduction

Scientific uncertainty is often a major factor in legal disputes involving large sums of money and in political controversies with major environmental or social consequences. In such situations, generalists untrained in natural science—judges, juries, government officials, managers in private industry, diplomats, and increasingly, members of the general public—must often judge the merits of arguments made on both sides of a public controversy among scientific experts. A critical input to this judgment is a reasonably precise understanding of the degree of uncertainty associated with particular assertions of scientific fact, or of a chain of evidence based on such assertions. For example, the advisability of possibly expensive and difficult interventions to minimize emissions of carbon dioxide into the atmosphere depends in large part on the level of uncertainty connected with scientific predictions of global warming.

† E-mail: weissc@georgetown.edu

In this paper we present a subjective, user-friendly scale of scientific certainty. This calibrated, numerical, 11-point scale is based on standards of proof applied in different situations by American criminal, civil and administrative law. These legal standards can be used as a standard vocabulary, somewhat analogous to the Richter scale for the strength of earthquakes, to express the degree of certainty or uncertainty associated with a given scientific assertion or chain of scientific evidence. We anticipate that the proposed scale will be useful both in communicating with policy makers, and in increasing the understanding of scientific uncertainty among the general public.

The scale proposed in this paper is a complement to quantitative scales based on subjective or Bayesian probability, which measure the odds that an informed better would accept that a given proposition is true. They allow a source of scientific information to express the subjective level of certainty or uncertainty that it associates with a particular assertion of scientific fact, and can also be used to represent the range of expert opinion regarding that certainty or uncertainty.[1] No scale can do away with disagreements over scientific uncertainty, but they can serve to make these disagreements clearer and more precise.

The need for a scale of uncertainty was clearly demonstrated by the reaction to the Second Report of the Inter-Governmental Panel on Climate Change (IPCC), issued in 1995, which was widely criticized for its internally inconsistent treatment of scientific uncertainty (IPCC, 1995). In response, the authors of the recently published third IPCC report have adopted and implemented a seven-point scale based on the numerical probability they assigned to the various assertions contained in the report (IPCC, 2001).

2. A proposed scale of scientific certainty

The standards of proof that are used to calibrate the proposed scale set forth the different levels of certainty or uncertainty that the law deems consistent with intervention in a wide variety of circumstances, based on situations that are more or less familiar to the public and are expressed in terms that it can readily understand. They range in stringency from the familiar standard of the criminal justice system, that guilt must be proven 'beyond a reasonable doubt', to the less familiar standard of 'reasonable suspicion' that suffices to justify a brief 'stop and frisk' by a policeman in order to ensure that a person is not carrying a concealed weapon.[2]

We correlate these standards with the measures of certainty used informally by working scientists in gauging the likelihood that a given scientific proposition will 'turn out to be true' upon further research. In separate papers, we show how this and other similar

[1] By uncertainty we mean any situation where the odds of an unfortunate consequence are unknown, whether because of inadequate data (second-order risk, in the terminology of Einhorn & Hogarth (1985)), or because of incomplete scientific understanding or an indeterminate chain of causality (ignorance and indeterminacy, respectively, in the terminology of Wynne (1992), which in turn would result in vagueness, from the point of view of the decision maker, in the terminology of Wallsten (1990)). We shall use the terms 'certainty' and 'uncertainty' as simple inverses. We shall use the term 'risk' to denote situations where the objective probability of unfortunate consequences is known from previous experience. This usage is different from that of decision theory, in which the term uncertainty is often used to denote situations in which risk is characterized by a known probability. See also footnote 20.

[2] We thank Professors Paul Rothstein and Samuel Dash of the Georgetown University Law Centre, and Kathleen Beaufait for introducing the author to the intricacies of legal standards of proof.

scales can be used to introduce explicitly the level of scientific uncertainty into the mix of factors that enter into scientific advice and scientific advocacy, and into the implementation of the Precautionary Principle in international environmental law and in other aspects of international and domestic regulation (Weiss, 2002).

The proposed scale is subjective, in that it is meant to enable advocates, advisers, lawyers, historians of science, researchers on technology assessment and policy makers to express themselves explicitly and with reasonable precision regarding the degree of certainty or uncertainty that they themselves associate with a given scientific assertion or chain of evidence. This will enable them, if they so wish, to clearly relate this degree of certainty or uncertainty to the other factors bearing on the particular decision. It is calibrated, in the sense that each step in the hierarchy of increased certainty is correlated with a reasonably well-defined criterion corresponding to a set of situations defined in the law.

The proposed scale is not intended as a replacement for quantitative scales, based on so-called frequentist statistics (see footnote 20) and used by epidemiologists and risk assessors to characterize risks with well known objective probabilities. It complements probabilistic scales, based on so-called Bayesian statistics, that have been used by advisory bodies such as the IPCC to characterize the subjective uncertainty they associate with statements whose scientific basis is uncertain. The qualitative nature of the proposed scale should make it more accessible to members of the public that are not accustomed to thinking in quantitative terms.

In addition to facilitating expressions of opinion regarding the certainty or uncertainty of a given scientific assertion, the scale provides a way to express the range of opinion among experts regarding that assertion at any given time, and to locate one's view of the certainty or uncertainty of a particular assertion on the spectrum of scientific opinion: for example as 'conventional wisdom', as an opinion held by a minority of qualified scientists, as an iconoclastic view requiring a substantial paradigm shift, or as a view contradicting well-established scientific principles.

We anticipate that this scale will be useful to a wide range of potential users: scientific advisers, journalists, non-governmental organizations, regulators, technology assessors, historians of science, science policy researchers, and drafters of legislation. It provides a new way for researchers or technology assessors, for example, to characterize the level of certainty or uncertainty that they associate with a given assertion as a function of time. In the policy realm, it enables expert scientific advisers to convey in an authoritative manner to policy makers and the public the level of scientific certainty they associate with a given assertion. It also enables both decision makers and the general public to match the degree of certainty associated with the seriousness of a given danger (for example, the hazards associated with a given pesticide) with their willingness to accept or desire to avoid the possibility of unfortunate consequences, and hence to judge the scientific arguments being presented by non-governmental organizations, trade organizations, and other advocates. While we do not consider this possibility in this paper, the scale can be extended to cover expressions of uncertainty outside the realm of science, as for example in assessments made in the course of policy, management or intelligence work.

3. Uncertainty and public controversy

In public controversies involving scientific uncertainty, advocates of both sides typically make use of whatever scientific arguments favour their cause, so that issues of scientific uncertainty become inextricably intertwined with differences in policy and philosophy. The forum for such controversies may be a court of law, a confidential discussion between decision makers and their advisers and associates, an international negotiation or, increasingly, the mass media and the Internet. The source of scientific information may be the Internet, the mass media, the professional or semi-popular scientific literature, one or more trusted advisers or consultants, a non-governmental organization or other opinion maker, or in the case of a government agency or international organization, the reports of the National Research Council of the US National Academies of Science, or a formal scientific advisory committee. These last are obligated to reflect the result of a careful effort to balance and integrate a wide range of views, and have traditionally been regarded as the authoritative voice of the scientific community on policy matters involving a high degree of technical content.

In practice, however, most policy controversies involving scientific uncertainty are no longer resolved in private dialogue between policy makers and experts. Rather, they involve wide-ranging debate, in which the general public is likely to express strong opinions, even when the underlying science is difficult and complex. Indeed, the participation of the general public in the environmental assessment process has been identified as an essential part of 'social learning' (Social Learning Group, 2001). In the age of the Internet, the public often draws its information, not only from a mainstream scientific consensus, but also from opinions held by only a minority of reputable scientists (Rowland, 1993). This phenomenon has been especially pronounced during the debate over climate change, in which minority opinions have been heavily publicized. To the despair of the scientific community, moreover, some segments of the public are influenced by arguments that are supported by no scientific evidence whatsoever, or that even contravene well-established scientific principles (Park, 2000; Sagan, 1996).

In public discussion, these diverse opinions are frequently presented as having equal status. As a result, the opinions of scientific advisers, even when expressed as the consensus of distinguished bodies of experts such as those convened by the National Research Council, the IPCC and similar bodies, become only one of many inputs into the public debate. In effect, and despite their best efforts to control the discussion, the function of these advisory groups becomes that of setting the framework for public debate, rather than that of rendering authoritative judgments, as has traditionally been the case (Weiss, 2002). This constitutes a major change in the function and, in consequence, in the self-image of these groups.

4. Legal standards of proof as equivalent to levels of uncertainty

Controversies involving scientific uncertainty typically involve a disagreement over whether or not a given problem is sufficiently important, and sufficiently well understood, to justify a legal remedy or some form of national or international policy or regulatory intervention. The underlying question concerns the level of proof required to justify the proposed remedy or intervention. On one side are typically found those who argue that a regulatory or policy intervention, or an award of damages or other legal remedy, is unjustified because the evidence for action is insufficient to satisfy rigorous criteria of scientific proof. To act on the basis of the evidence available, they argue, would be to base decisions on ‘bad science’. On the other side would be those that argue that available evidence and understanding, although falling short of rigorous scientific proof, is sufficient to justify the proposed intervention.

Different branches of US law have evolved a rich and nuanced menu of standards of proof to deal with the variety of human experience that requires decision makers to weigh the relative probabilities of different interpretations of the facts before them. Some of these standards, like the standards of ‘beyond a reasonable doubt’ required for criminal conviction, or of ‘preponderance of the evidence’ (sometimes rendered ‘more likely than not’) required for a decision in a civil case, are well known to the public through its exposure to the mass media. Others, like the requirement for ‘reasonably articulable suspicion’ in certain situations related to search and seizure, are less well known or are generally regarded as remote from the policy arena.

The various standards of proof used in different branches of the law are designed to give proper relative weight to the rights of the different stakeholders under different sets of circumstances. A given standard of proof embodies a societal judgment regarding the desired balance between false positives and false negatives (e.g. convicting the innocent and letting the guilty go free, respectively), and hence on the balance between conflicting rights (Strong, 1999, p. 517). To cite two common examples, the criterion of ‘beyond a reasonable doubt’ used in criminal cases is intended to assure that innocent people are not convicted, even if some guilty people go free. In contrast, the criterion of ‘preponderance of the evidence’, used in civil cases, is intended to assure a level playing field in which neither the plaintiff nor the defendant is to be shown particular preference.

We have assembled a set of standards of proof, of differing stringency and drawn from diverse branches of the law, into a hierarchy of levels of increasing certainty (or decreasing uncertainty).[3] A legal standard of proof is defined as ‘the level of certainty and the degree of evidence necessary to establish proof in a criminal or civil proceeding’ (Merriam-Webster, 1996).[4] In this paper, we propose that the levels of this hierarchy may be used to convey different levels of scientific uncertainty.

This hierarchy is set forth in Table 1 and is expressed as a numerical scale ranging from zero to ten. The standards of proof discussed in the table are set in italics when they appear for the first time in the discussion below. Each of these standards of proof has been clearly defined in the US courts and refined by being applied to actual cases. Most are the law of the land. For our purposes, however, it is not important whether or not they are actually current legal precedent. For this reason, the author does not assert that the cases cited in the footnotes to this article constitute currently valid precedent (although this is true in nearly all cases), but only that they define the standards that are being used as benchmarks in the proposed scale.

[3] A somewhat similar but less complete discussion of standards of proof, as they apply to cases involving scientific evidence, is found in Loevinger (1992). For a general treatment of standards of proof, see McCauliff (1982).

[4] In more formal legal language, a standard of proof is defined as the criterion by which the finder of fact (a jury, judge or administrator) is to judge whether the burden of persuasion has been met in a particular case, i.e. whether the evidence and arguments presented are sufficiently convincing (Strong, 1999, p. 409 and 508).

TABLE 1 *A proposed scale of scientific certainty based on legally defined standards of proof* [a]

Level	Legal standard	Other language	Legal action
10	'Beyond any doubt' (not a legal standard)		Exceeds criminal standard; implicit in some critiques of the death penalty (Ryan, 2000)
9	'Beyond a reasonable doubt' (Criminal Law)	'So convincing that a reasonable person would not hesitate to act' (Mueller & Kirkpatrick, 1997, p. 145); 'proof that leaves you firmly convinced . . . [no] real possibility that he is not guilty . . . ' (Federal Judicial Centre, 1987, pp. 17–18)	Criminal conviction (Strong, 1999, sec 341, p. 428)
8	'Clear and convincing evidence'[b] (Civil Law)	'Clear, unequivocal and convincing' (Schwartz, 1991, p. 387); leading to 'a firm belief or conviction that the allegation is true' (Park *et al.*, 1991, p. 91; footnotes 6 and 8)	Quasi-penal civil actions, such as termination of parental rights, denaturalization or deportation (Strong, 1999, sec. 340, p. 425); Criminal sentencing hearings (*U.S. v. Fatico*, 1978)
7	'Clear showing' (Civil Law)	'Clear likelihood of success' (*Bristol v. Microsoft*, 1998); 'Reasonable probability' (Gotanda, 1993)	Granting preliminary injunction (Yeazell, 2000, p. 365; [c]Wright *et al.*, 1995, pp. 129–130; Gotanda, 1993); [d]overturning consent decree[e]
6	'Substantial and credible evidence'	'Such evidence as a reasonable mind might accept as adequate to support a conclusion' (*ConEdison v. NLRB*, 1938)	Referring evidence for impeachment (US Code, 2001)
5	'Preponderance of the evidence' (Civil Law)	'Existence of a contested fact more probable than not' (Mueller & Kirkpatrick, 1997, p. 121); 'preponderance of probability' (Strong, 1999, p. 423)	Most civil cases (Mueller & Kirkpatrick, 1997, p. 121); Administrative and regulatory rulings (Schwartz, 1991, p. 387)
4	'Clear indication' [f]		Proposed as criterion for night-time, X-Ray or Body cavity searches (Lafave & Israel, 1992, p. 111 and 224)
3	'Probable cause'[g] (Criminal Law)	'Would warrant a belief by a reasonable man' (Lafave *et al.*, 2000, p. 149); 'More than bare suspicion . . . less than evidence that would justify conviction' (Lafave, 1968, p. 113)	Field arrest or search incident to arrest; search warrant (Lafave *et al.*, 2000, p. 149); arraignment and indictment (*Illinois v. Wardlow*, 1999)

TABLE 1 *Continued.*

Level	Legal standard	Other language	Legal action
2	'Reasonable grounds for suspicion' (Criminal Law)	'Reasonable, articulable suspicion' (*Illinois v. Wardlow*, 1999); 'substantial possibility' (Lafave, 1968, p. 40 and 87)	'Terry Stop and Frisk' (*Terry v. Ohio*, 1968)
1	'No reasonable grounds for suspicion'[h] (Criminal Law)	'Inchoate and unparticularized suspicion or hunch' (*Illinois v. Wardlow*, 1999); 'fanciful conjecture'[i] (*Victor v. Nebraska* (1994); *Sandoval v. California*, 1994)	Does not justify Terry stop (*Terry v. Ohio*, 1968)
0	Impossible (Criminal Law)	Action taken could not possibly have resulted in the crime being charged	A possible defence, but not a standard of proof

[a] Including a few not accepted by US courts.
[b] Equivalent to 'moral certainty'. See main text.
[c] The proposition to be 'clearly shown' is that the case will probably be won by the plaintiff seeking the injunction.
[d] The proposition to be 'clearly shown' is that grievous harm will result from new, changed conditions unforeseen at the time of the consent decree. Some jurisdictions require only a 'fair chance of success' (*Mazurek v. Armstrong*, 1996).
[e] See: *U. S. v. Swift & Co*., 286 U.S.106, 119, 76 L. Ed. 999, 52 S. Ct. 460, codified in *Federal Rules of Civil Procedure*, Rule 60 (b)(5).
[f] This standard is well defined in the legal literature, but has not been accepted by the courts as a valid legal standard of proof.
[g] Some other well-defined legal standards of proof, taken from the civil and administrative law, have approximately the same force as 'probable cause'. These include (i) the requirement of the Federal Administrative Procedures Act (APA) that regulatory decisions have a 'rational basis' in science, lest they be held 'arbitrary and capricious' (*People v. Cecil Todd*, 1992); and (ii) the requirement that evidence considered in connection with the setting of penalties in a civil anti-trust suit be 'credible' (*Bristol v. Microsoft*, 1998).
[h] Roughly equivalent to 'clearly erroneous', the criterion for rejection by an appellate court of a lower court's findings of fact (*U.S. v. United Gypsum Co*, 1948).
[i] This refers to a degree of doubt less than that necessary to acquit a criminal defendant.

We begin with the standards of proof most familiar to the public. The most rigorous standard of legal proof, the one used in criminal cases, is that of '*beyond a reasonable doubt*'. According to standard legal texts, evidence meeting this criterion must be 'so convincing that a reasonable person would not hesitate to rely and act upon it in the most important of his own affairs' (Mueller & Kirkpatrick, 1997, p. 145).[5] Proof 'beyond a reasonable doubt' is not proof to absolute certainty. But a 'reasonable doubt' must be more than a 'mere possible doubt' or 'fanciful conjecture'. The standard of 'beyond a reasonable doubt' is assigned to level 9 of the proposed scale.

[5] Shapiro (1991) provides an interesting account of the origins of 'beyond a reasonable doubt' as a legal formula in Anglo-American law.

In contrast, '*preponderance of the evidence*', the familiar standard of proof used in civil contests and most administrative proceedings, is defined as 'the greater weight', or 'better' evidence, that indicates 'a preponderance of probability' so that 'the existence of a contested fact is more probable than its nonexistence'.[6] This standard is assigned to level 5 of the proposed scale. A standard explanation to a lay jury of the difference between the civil and criminal standards of proof is to hold a pencil at a slight tilt from the horizontal, and to ask the jury to imagine that the evidence for each side of a case is represented by a weight at equal distances from and on opposite sides of an imaginary balance point at the mid-point of the pencil. The lawyer then explains that even a slight preponderance of the evidence is sufficient to decide a civil case. In a criminal case, by contrast, the necessary weight of evidence corresponds to the same pencil being held slightly off the vertical.[7]

As is evident from Table 1, several levels of certainty fall within the substantial gap that lies between the 'criminal' and the 'civil' standards of proof. In a number of quasi-penal situations falling under the jurisdiction of the civil courts but involving 'a level of deprivation of individual rights less than would result from a criminal prosecution', the civil standard of proof is raised from 'preponderance of the evidence' to the stronger criterion of '*clear and convincing evidence*'. This has been defined as evidence that 'leads to a firm belief or conviction that the allegations are true' (Park *et al.*, 1991, p. 91). This standard is applied to such quasi-penal civil cases as deportations, civil commitment to a mental hospital, denaturalization and deportation, termination of parental rights, establishing the terms of a lost will, illegitimacy of a child born to a married woman, or disciplining of a lawyer (Strong, 1999, p. 515).

The standard of 'clear and convincing evidence' approximates that of 'moral certainty', defined in the dictionary as 'likelihood so great as to be safely acted upon, although not capable of certain proof' (Merriam-Webster's, 1996). The Supreme Court has held that 'moral certainty' is regarded in law as a weaker standard than 'beyond a reasonable doubt'.[8] This is consistent with the dictionary definition cited herein, since one may act on

[6] The Federal Administrative Procedures Act (APA) establishes the standard of 'preponderance of the evidence' in administrative proceedings. The Supreme Court considered and rejected the alternative standard of 'clear and convincing evidence' in the case of *Steadman v. SEC* (1981), on the grounds that the 'preponderance of the evidence' was the standard intended by Congress (Schwartz, 1991, sec. 7.9, p. 366 and 386).

[7] Professor Sam Dash of the Georgetown University Law Center, personal communication. This explanation, while effective before a jury, is unsatisfying to a scientist. The problem is as follows. If we represent the evidence on the two sides of the case as weights suspended from points located at equal distances from a fulcrum at the midpoint of a uniform, eraser-less, unsharpened pencil, or the more real-life situation of children on either end of a see-saw sitting at equal distances from the balance point, even a small imbalance of the weights on either end would result (in the absence of friction) in a torque that would send the seesaw spinning around the balance point until it met some obstacle, such as the ground.

If the balance point is not shifted towards the weightier evidence or the heavier child, a restoring force of some sort (say, a spring) would be needed in order to enable the pencil or the seesaw to remain stable at an angle from the horizontal. In other words, in the absence of a restoring force, the seesaw would end up with the heavier end on the ground, no matter how small the weight differential. Similarly, the pencil would end up in a vertical position with the heavier weight hanging from its bottom end. If a restoring force from a spring were introduced into the demonstration, the deviation of the pencil or the seesaw from the horizontal would depend both on the weight differential and the stiffness of the spring.

[8] "To equate 'beyond a reasonable doubt' with 'moral certainty' is to overstate the degree of doubt needed for acquittal" (*Cage v. Louisiana*, 1990). 'Moral certainty' is the standard in the state courts of New York for excluding alternative explanations in cases entirely dependent on circumstantial evidence (*People v. Edgar Bearden*, 1943).

'moral certainty' even in the presence of 'reasonable doubts' that make one hesitate before doing so.[9] The standard of 'clear and convincing evidence' is assigned to level 8 of the proposed scale.

A second standard of proof between the criminal and civil standards is defined by the requirement, imposed by some courts, that requests for stays of execution of lower court decisions pending appeal, or for temporary injunctions pending trial, be backed by a '*clear showing*', or (equivalently) a showing of 'reasonable probability' that the action will succeed on its merits in cases where damage to the applicant would be especially severe if the stay or injunction is not granted (Gotanda, 1993). The 'clear showing' standard is also applied in cases in which it is claimed that a consent decree should be overturned on the grounds that grievous harm will result from new and unforeseen conditions (*U.S. v. Swift & Co*, 1932).[10] This standard is assigned to level 7 of the proposed scale.

A third standard of proof falling between 'beyond a reasonable doubt' and 'preponderance of the evidence' stems from a seemingly unlikely source, namely the provisions of the Independent Counsel Act that govern the transmission of a report from the Independent Counsel to the House of Representatives regarding the impeachment of the President of the United States or other government officials.[11] According to the act, such evidence is to be '*substantial and credible*'. This standard is assigned to level 6 of the proposed scale.

We now consider standards of proof that fall short of the 'preponderance of evidence' required for victory in a civil court proceeding. For this purpose, we turn to the rich menu provided by US constitutional provisions and court decisions governing search and seizure. Each of these standards of proof strikes a different balance among the rights of the subject to be free from intrusion, the interests of the public in detecting and apprehending lawbreakers, and the safety and security of officers of the law.

The least demanding of these standards of proof is the criterion for the so-called 'Terry stop and frisk', defined as a brief detention 'so strictly limited that it is difficult to conceive of a less intrusive means that would accomplish the purpose of the stop'.[12] The doctrine of the Terry stop allows the police officer to pat someone down to be sure that (s)he does not have a weapon before the officer asks questions. A police officer may carry out a Terry stop on '*reasonable grounds for suspicion*', defined as a suspicion based on 'objective, articulable facts, leading an experienced, prudent officer to suspect that the individual

[9] The definition of 'beyond a reasonable doubt' used in Mueller & Kirkpatrick (1997, p. 146) includes the statement that this standard requires an 'abiding conviction to moral certainty'. This definition does not conflict with the one given in Table 1 because the use of the term 'moral certainty' in the Mueller–Kirkpatrick definition refers to an older usage, namely 'certainty based on empirical evidence'. It is therefore not in conflict with the distinction made in this paper between 'beyond a reasonable doubt' and the current meaning of 'moral certainty'. (*Victor v. Nebraska* (1994); *Sandoval v. California*, 1994, quoting Shapiro, 1986, 1991). 'Beyond a reasonable doubt', in contrast, implies that there will be no hesitation before acting, a more stringent standard.

[10] Also codified in *Federal Rules of Civil Procedure*, Rule 60 (b)(5). Some jurisdictions require only a 'fair chance of success' (*Mazurek v. Armstrong*, 1996).

[11] We thank Professor Samuel Dash of Georgetown University Law Centre for bringing this standard to our attention. This standard is unlikely to be further interpreted by the courts, given the fact that the impeachment and conviction of a federal official for 'high crimes and misdemeanors' is inherently a political process in which the standard of proof is a matter for each senator and representative to decide for himself or herself. (Rovella, 1998, quoting Gerald E. Lynch).

[12] Concurring opinion by Justice Brenner in *Florida v. Royer*, 1983, pp. 510–11.

is concealing something on his/her person contrary to law'.[13] The test is intended to balance the interest in preventing flight, minimizing the risk of harm to the officer, and the desirability of the orderly completion of the search. This is the least stringent standard of proof in US criminal law, and is assigned to level 2 of the proposed scale. Evidence that does not reach even this standard is referred to as a 'mere hunch', and is insufficient to justify even a Terry stop. This is assigned to level 1 of the proposed scale.

An actual arrest, or a search of a person or a building or other area where a person has a reasonable expectation of privacy, requires a higher standard of proof, namely '*probable cause*', a standard derived from court interpretations of the provisions of the Fourth Amendment to the US Constitution that citizens be protected from 'unreasonable search and seizure' and defined as 'reasonable grounds to believe' or 'reasonable, articulable suspicion'. The evidence must 'warrant a belief by a reasonable man, taking into account his or her experience'. This is assigned to level 3 of the proposed scale.

The standard of '*clear indication*' lies above 'reasonable cause' but well short of 'preponderance of the evidence'. It has been applied in a number of dissenting federal opinions and decisions by state courts, which have argued that certain situations, such as night-time searches, x-ray searches, and searches involving 'intrusions below the body's surface', demand a more stringent standard of proof than 'probable cause'.[14] This well-defined standard was rejected by the Supreme Court, on the practical grounds that 'a single familiar standard is essential to guide police officers' in situations in which rapid decisions must be made under difficult field conditions (*Dunaway v. NY*, 1979). It is nevertheless useful for the purposes of this paper, and is assigned to level 4 of the proposed scale.

We now come to the question of whether it is possible to define a standard of proof more rigorous than the criminal standard of 'beyond a reasonable doubt'. This issue presents special philosophical and practical problems. In law, no witness can be absolutely certain. In science, any theory can in principle be disproved. For that matter, in the strictest interpretation of inductive logic, we cannot be absolutely certain that the sun will rise tomorrow morning.[15] This paper takes the pragmatic position that this top standard should be '*beyond any doubt*', notwithstanding the fact that such a standard is not recognized by US law and indeed comes perilously close to the requirement of absolute certainty, a

[13] Lafave & Israel, 1992, p. 224; Lafave *et al.*, 2000, 215 ff.

[14] According to Professor Dash (personal communication), no Supreme Court decision has defined any such separate standard of proof applying to a 'reasonable' search, the only binding criterion of reasonableness being that a search not endanger the health or safety of the person(s) being searched. For night-time searches, see the dissent by Justice Marshall in *Gooding v. United States*, 416 U.S. 430. For x-ray searches at the border, see *U.S. v. Castrillon*, CDCalif. 82-1722, 9-27. For intrusive searches, see *Schmerber v. California*, 384 U.S. 757 (1966). For general discussion, see Lafave *et al.*, 2000, pp. 163–4. Lafave (1968, pp. 111, 112, and 224) cite cases that invoke a separate standard of proof ('real suspicion' or 'clear indication') to justify peculiarly intrusive interventions.

[15] Klee, 1997. In principle, an infinite number of observations are needed to prove by scientific induction—or, what is logically equivalent, induction tends towards proof as the number of observations tends towards infinity.

standard unattainable both in law and in the philosophy of science.[16] This is assigned to level 10 of the proposed scale.

Finally, at the bottom of the scale, we come to the legal concept of *impossibility*. This is a defence in criminal law, not a standard of proof, and constitutes the argument that the actions alleged, even if proven, could not possibly have resulted in the fulfillment of the elements of the criminal violation being charged, or the civil wrong alleged. In a classic example, supposedly derived from a Dick Tracy cartoon, a defendant charged with murder could raise an impossibility defence if (s)he could show that (s)he had shot and hit a cardboard silhouette of the intended target, rather than the target himself, and therefore could not have killed him whatever the shooter's intent.[17] In a recent, actual case, a murder conviction was overturned on appeal because the murder victim, the defendant's supposed child, could not have existed because the defendant had been sterilized before the 'murdered' baby was supposed to have been conceived (*Banks v. Alabama*, 2002). Impossibility is assigned to level zero of the proposed scale.

5. Uncertainty and the scientific researcher

To the working research scientist, scientific uncertainty is at the same time a subject of considerable ambivalence and a fact of daily professional life. In principle, a scientific assertion is either proven or unproven. If the former, it is a fact that may be entered into textbooks and taught to students. If the latter, it is a conjecture, whose validity may be 'suggested' but never asserted as fact. This binary concept of scientific truth is implicit in the language used in the scientific literature, is part of the socialization of every young research scientist, and is the philosophical justification of the reluctance of many researchers to express public judgments on scientific issues that have not been subjected to definitive proof. It also colours the common public expectation that science will provide unambiguous truth, free of uncertainty, and the frustration of the public, as well as that of political and judicial institutions, when such statements are not forthcoming.

In workaday practice, however, the distinction between fact and conjecture is not always so clear, and the attitude of research scientists towards uncertainty is considerably more nuanced and pragmatic. While they might be reluctant to say so in a published paper or in a public or formal setting, research scientists are constantly assessing the degree of certainty or uncertainty of scientific assertions, and extrapolating from recent trends as the basis for making personal estimates of the likely outcome of future research. Such

[16] The discovery that a number of residents of American death rows had been wrongly convicted gave rise to the argument that a standard more rigorous than 'beyond a reasonable doubt' should be applied to capital cases. This discussion has occasioned some confusion in terminology. For example, the Governor of Illinois, on learning that 13 inmates on death row had been exonerated by DNA testing, declared, 'Until I can be sure *with moral certainty* (italics added) that no innocent man or woman is facing a lethal injection, no one will meet that fate' (Ryan, 2000).

As we have seen, the standard of 'moral certainty', as the term is currently used, is actually *less* rigorous than that of 'beyond a reasonable doubt', so that Governor Ryan's stated standard of 'moral certainty' would presumably have been met if the standard of 'beyond a reasonable doubt' had been correctly applied in these cases. Governor Ryan presumably meant to urge that capital cases be resolved beyond any doubt at all.

[17] Samuel Dash, personal communication. This argument is valid only if the substance of the crime, rather than the motivation, is an essential element in the offence. It would not apply to attempted murder, or to other crimes where only the perpetrators' acts and intent are relevant, as for example, stealing false documents they believed to be trade secrets or buying fake drugs from an undercover agent.

assessments form the basis of strategic decisions regarding what topics to pursue or to ignore, and even in what order to do a series of planned experiments. In informal conversation, a working scientist might very well say to colleagues that a given conjecture, while still unproven, is supported by an increasing amount of evidence and is likely to 'turn out to be true' or even that it is 'reasonably certain'. Still another conjecture might be described as unlikely, not because it is impossible in principle but because the available evidence fails to support it or (in the opinion of the describer) is more consistent with a rival explanation. Still other assertions are dismissed as being without scientific validity, either because they are not subject to scientific verification, or because they contravene known and accepted scientific principles or observations. The greater the extent to which a given proposition contravenes accepted paradigms, the more likely it is to be met with scepticism or even resistance. In the scientist's aphorism, 'extraordinary claims require extraordinary evidence'.

What is more, scientists intuitively understand that the formation of a scientific consensus is a social process, and that a new hypothesis will naturally be accepted by some scientists before it is accepted by others. This is the basis of (Kuhn's 1970, p. 151) seminal finding that a major 'paradigm' shift in scientific thinking gains acceptance as much by the aging of its human opponents as by the accumulation of evidence in its favour.[18] To be sure, the new paradigm is not a purely social construct, but is constrained and protected from error by continuous empirical checks.

At different stages on the uneven path to acceptance or rejection, a scientific assertion may be regarded by one group of scientists as 'impossible' or 'improbable', by another as 'possible but still unproven', and by another as 'probable', depending on their scientific discipline, their personal tendency toward scepticism, their political or religious views, and perhaps also on their economic interests. Indeed, from the strictly anthropological point of view, scientific research may be viewed as an elaborate and expensive effort to change 'discourse' from 'may be' to 'is' (Latour & Woolgar, 1986, pp. 81–88).

Given this dynamic, a respected minority of scientists may retain its scepticism regarding a new hypothesis for many years, and may form a coherent and recognized subset of the relevant scientific community. This phenomenon of **minority science** is normal in research science, and takes an especially interesting and important form when scientific arguments are used in support of policy positions, a phenomenon we shall call **advocacy science**, or are used in support or opposition to environmental or other regulations (Atik, 1996/97; Weiss, 2002). If the minority scientific position on a particular policy issue has favourable consequences for a well-funded interest group, as has been the case for the so-called 'climate sceptics' in the current debate over climate change, the position may receive attention well out of proportion to the level of support it receives within the scientific community. It then becomes important for neutral parties to represent to the public the level of support that this hypothesis commands, without at the same time dismissing it out of hand.

In Table 2, we set forth a hierarchy of scientific uncertainty, derived from the practical workings of research science, that may be summarized in 11 subjective levels of scientific certainty: fundamental, rigorously proven, substantially proven, very probable, probable,

[18] See also Planck (1949, 33–34): 'A new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up familiar with it'.

more likely than not, attractive but unproven, plausible, possible, unlikely, and impossible. The categories are equally applicable to specific technical assertions (e.g. ‘global warming increases the atmospheric load of water vapour’.) and to broad assertions based on the aggregate assessment of numerous individual assertions (e.g. ‘the Earth will warm 1–3°C if atmospheric carbon dioxide levels double in the next century’). The assertions have to be sufficiently specific that their truth could in principle be verified if sufficient evidence were to become available. For example, the illustrative assertion for ‘if I have to choose’ refers to ‘the past 100 million years’, not to a less sharply defined time period such as the ‘recent past’.

TABLE 2 *Levels of certainty: relating ‘legal’ and ‘scientific’ criteria*

Level	Legal standards of proof	Informal scientific levels of certainty	Scientific assertions (author’s subjective opinion)
10	‘Beyond any doubt’	Fundamental theory on conclusion from experiment that explains a wide variety of observations, within well-understood limits of validity.[a]	Law of gravitation; Maxwell’s equations of electromagnetism; theory of relativity; theory of evolution; quantum electrodynamics; plate tectonics.
9	‘Beyond a reasonable doubt’	Rigorously proven. Critical experiment(s) give(s) a clear and unambiguous result, excluding alternative explanations.	CFCs cause the stratospheric ozone hole; smoking and asbestos cause cancer; DDT exposure leads to the thinning of eggshells of birds; Earth’s early atmosphere contained no oxygen; AIDS is caused by HIV.
8	‘Clear and Convincing Evidence’	Substantially proven. A few details remain to be worked out. ‘Reasonably certain’.	Dinosaurs went extinct due to a large meteor or comet impact; breast-feeding boosts infants’ immune systems; growth of plankton in the equatorial Pacific is limited by the availability of iron; phosphorus in US detergents caused lake eutrophication.
7	‘Clear Showing’	Very probable.	Half of all the molecules in interstellar space (other than H_2) are organic; cod stocks in the Grand Banks declined from over-fishing; zebra mussels succeed in US because they have no natural predators.
6	‘Substantial and credible evidence’	Probable. Evidence points in this direction, but not fully proven.	Neutrinos have non-zero rest mass; dust mites cause asthma.

TABLE 2 *Continued.*

Level	Legal standards of proof	Informal scientific levels of certainty	Scientific assertions (author's subjective opinion)
5	'Preponderance of the Evidence'	More likely than not. If I have to choose, this seems more likely to be true than untrue.	There has been liquid water on the surface of Mars at some time within the past 100 million years; recent increases in ground-level ultra-violet light have increased rates of skin cancer.
4	'Clear indication'	Attractive but unproven. Evidence is beginning to accumulate in this direction.	About half of all stars have at least one planet.
3	'Probable cause: reasonable grounds for belief'	Plausible hypothesis, supported by some evidence.	SO_2 emissions from power plants are the major cause of European tree damage; global warming will lead to the expansion of tropical diseases.
2	'Reasonable, articulable grounds for suspicion'	Possible, worth researching.	Synthetic chemicals have caused decrease in human sperm counts; traces of mercury in infant vaccines have led to increased rates of autism.
1	'No reasonable grounds for suspicion'	Unlikely: available evidence is against it, or violates existing paradigms, but not entirely ruled out.	Nuclear reactions can be initiated by electrochemical means (cold fusion); cell phones or high voltage power lines cause cancer.
0	'Impossible'	Against the known laws of physics or other science.	Perpetual motion machines; traits acquired during an individual's lifetime from environmental factors are passed on genetically to the next generation.

Note: developed jointly with Dr Robert Kandel of CNRS, Professor Wesley Matthews of the Georgetown Department of Physics and Dr Jennine Cavendar-Bares, then of the Smithsonian Institution and now of the University of Minnesota.

[a] The validity of these theories within their well-understood range of applicability is incontrovertible. Subsequent research might possibly establish that these principles are special cases of some larger law. Classical mechanics, for example, precisely predicts the motions of objects of the size common to human experience. At the atomic or molecular level, on the other hand, the laws of quantum mechanics govern, whereas at speeds comparable to that of light, the laws of relativity govern. The unification of these three theories is one of the great unsolved problems of physics.

We assign these levels numerical values from zero to ten, and compare them to the levels derived in the previous section from legal standards of proof. Considering the difference in their origins, the correspondence between the two sets of categories is remarkably good. In the right-hand column of Table 2, the author cites scientific assertions that he and some of his colleagues associate with each of the benchmark levels of uncertainty. These are expressions of opinion, and have no validity in and of themselves, other than to show that the scale provides a practical and user-friendly way for anyone, expert or not, to indicate the level of uncertainty (s)he assigns to a given assertion.

The systematic treatment of scientific uncertainty incorporated into the third report of the IPCC built on a major effort by a concerned group of scientists (informally known as the 'uncertainty police'), who ensured that the IPCC carefully assessed and assigned a value to the uncertainty associated with many of its statements of scientific fact (Inter-Governmental Panel on Climate Change, 2001). The well defined scale of scientific uncertainty adopted and implemented by this authoritative body is to our knowledge the first effort by an international advisory panel to treat such uncertainty in a systematic manner. The IPCC scale is keyed to Bayesian probability, and consists of seven levels: $>$99%, or 'virtually certain'; 90–99%, or 'very likely'; 67–90%, or 'likely'; 33–67%, or 'medium likelihood'; 10–33%, or 'unlikely'; $1-10$%, or 'very unlikely'; and $<$ 1%, or 'exceptionally unlikely'.[19]

The scale of numerical probability used by the IPCC is a Bayesian scale, in the sense that it expresses a subjective probability reflecting someone's opinion regarding the probability that a given assertion is true. Bayesian probability is the mathematical expression of the observation that many people are accustomed to estimating the odds at which a rational better would be willing to bet on a specified future event, on the basis of their present understanding of the circumstances and the intensity of their belief that the event will or will not happen.[20]

Table 3 summarizes the comparison between the IPCC scale and the scale proposed in this paper. The fourth column of Table 3 shows the author's best efforts to assign probabilities to the legal standards of proof set forth in Tables 1 and 2, and to correlate each level of the IPCC scale with one or more levels of the scale proposed in this paper.[21]

[19] The numerical values assigned to the verbal formulations of the IPCC scale compensate for the fact that different individuals, and even different technical experts in a particular field, have been found to assign a wide range of different probabilities to such words as 'probable' and 'likely' (Wallsten *et al.*, 1986).

[20] The centuries-old controversy over whether people really do make intuitive estimates of probabilities is reviewed by Gigerenzer & Hoffrage (1995) and by Gigerenzer & Murray (1987). Bayesian statistics contrast with so-called frequentist statistics, used in epidemiological research. These assign objective probabilities, based on experience, to risky events (Morgan & Henrion, 1990, 126–137; Rosenfeld, 1975; Henrion & Fischhoff, 1986). These probabilities are derived from empirical correlations between effect and presumed cause, for example from the statistical relationship between auto accidents and miles driven. Calman & Royston (1997) proposed a logarithmic 'Richter-type' scale for making such risks clearer to the public.

The results of such analyses are often expressed as a statistical correlation, usually expressed as a confidence level of 95%, meaning that there is only a 5% probability that the observed correlation in the data between effect and postulated cause could have been due to chance when the correct correlation was zero. Such a statistical correlation cannot in and of itself be taken as proof, either in law or in science. The authoritative Henle–Koch–Evans (HKE) rules governing epidemiological research, for example, specify that a statistical correlation, even if it satisfies a confidence test, is still only a correlation. Similarly, in legal situations that hinge on statistics, such as paternity and DNA tests, such an objective statistical correlation must be accompanied by a plausible biological mechanism or other evidence of causality (Evans, 1976; Kaye, 1989). In such cases, subjective scales of uncertainty like the author's and the IPCC's may be applied to the evidence (statistical correlation plus accompanying evidence) underlying the overall judgement, but not to the statistical evidence taken by itself.

[21] While the legal profession by and large shies away from assigning numerical probabilities to the various standards of proof, a number of attempts to do so do exist in the legal literature (McCauliff, 1982; *U.S. v. Fatico*, 1978). The many difficulties encountered by judges and legal scholars in assigning quantitative probabilities to the various standards of proof will be discussed in a separate paper.

It is apparent from this column that the proposed scale is non-linear, in that the difference between adjacent levels of the scale does not always correspond to the same difference in percentage probability. (Unlike the Richter scale, it is not logarithmic.)

The last column of Table 3 illustrates the IPCC's use of the Bayesian scale by citing assertions, drawn from the Summary for Policy Makers, of scientific assertions corresponding to each level of uncertainty, as estimated by the Working Group. Some of these assertions are broad generalizations; others are narrow and specific. The IPCC assertions cluster in the 'likely' and 'very likely' range; indeed, only one assertion in the summary is rated as 'virtually certain', and none are of 'medium certainty'. ('Beyond all doubt' and 'impossible' do not appear on the IPCC scale.)

6. Uncertainty in the Courtroom

The application of the proposed scale of scientific uncertainty—or indeed, of any such scale—in the courtroom presents a variety of tricky questions that will be discussed in a subsequent paper. We confine ourselves here to comments on a few issues of particular importance.

The distinction between 'rigorously proved' and 'substantially proved' in the scientific column of Table 2 deserves additional discussion, since it corresponds to the difference between success and failure in the prosecutor's efforts to provide proof 'beyond a reasonable doubt'—i.e. the difference between conviction and acquittal in a criminal trial. In both the legal and the scientific situations, this distinction rests on success or failure in excluding alternative explanations. As we have previously discussed, proof 'beyond a reasonable doubt' takes place in law when there is no hesitation in acting on a conclusion, and is consistent only with doubts based on guesswork, speculation, imagination, or fanciful conjecture. This can take place only after alternative explanations—in this case, that the defendant did not commit the crime—have been excluded.[22] Similarly, 'rigorous proof'—the equivalent of proof 'beyond a reasonable doubt'—takes place in science when alternative explanations have been excluded.

In contrast, 'substantial proof', the next lesser standard, takes place in science when a critical experiment has given a definitive answer, but sufficient details remain to be cleared up as to allow alternative explanations still to have a chance of turning out to be correct. The analogous legal standard, 'clear and convincing' evidence, is defined as evidence that results in a 'clear conviction', but that would be consistent with the possibility of an alternative conclusion.

What, then, are we to make of it when, as frequently happens, an expert witness testifies in court that (s)he is 'reasonably certain' that a particular assertion is true? To a lay person, this phrase might seem consistent with 'clear and convincing evidence' (level 8), 'clear

[22] Because of the importance of this distinction, attempts to define the legal criterion of 'beyond a reasonable doubt' have been subject to close criticism and debate (see Mulrine, 1997; Kenney, 1995; Corwin, 2001). In a biting dissent to the decision of the Supreme Court in *Victor v. Nebraska* (1994); *Sandoval v. California* (1994), Justice Ruth Bader Ginsberg expresses a strong preference for the alternative formulation by the Federal Judicial Centre, to the effect that proof 'beyond a reasonable doubt' takes place when the jury is 'firmly convinced' that there is 'no real possibility' of innocence, i.e. when the alternative of innocence has been excluded . This alternative formulation is thus consistent with the distinction between 'rigorous' and 'substantial' proof as we have defined it in the main text.

TABLE 3 *Comparison of the 'legal', 'scientific', 'Bayesian' and 'IPCC' scales of scientific uncertainty*

Level	Legal standards of proof	Informal scientific levels of certainty	Bayesian probability	Level in IPCC scale	IPCC assertions (2001)
10	'Beyond any doubt'	Fundamental theory that explains a wide range of observations	100%	(absent)	(none)
9	'Beyond a reasonable doubt'	Rigorously proven; critical experiment(s) give(s) a clear result	>99%	'Virtually certain'	CO_2 emissions from fossil fuel burning will be the dominant influence on trends in atomospheric concentrations of CO_2 in the 21st Century.
8	'Clear and convincing evidence'	Substantially proven; a few details remain to be worked out 'Reasonably certain'. (see main text)	90–99%	'Very likely'	The projected rate of global warming has no precedent in the last 10 000 years; globally, the 1990s were the warmest decade since 1861; forests took up more CO_2 in the 1990s than was lost to deforestation.
7	'Clear showing'	Very probable	80–90%	'Likely'	1998 was the warmest year, and 1990 was the warmest decade in the Northern Hemisphere in the last 1000 years; The present atmospheric CO_2 concentration has not been exceeded in the past 10 000 years; there is increased future risk of drying and drought in mid-latitude continental interiors; the Greenland Antarctic ice sheet will lose mass and the Antarctic ice sheet will gain mass.
6	'Substantial and Credible Evidence'	Probable; evidence points in this direction, but not fully proven	67–80%		
5	'Preponderance of the Evidence'	If I have to choose, this seems more likely true than untrue	50–67%	'Medium Likelihood'	(None)
4	'Clear Indication'	Attractive but unproven; evidence is beginning to accumulate in this direction	33–50%		
3	'Probable cause: reasonable Grounds for Belief'	Plausible hypothesis, supported by some evidence	10–33%	'Unlikely'	Warming over the past 1000 years is entirely of natural origin, according to reconstructions of climatic data; changes in natural forcing [i.e. causes that are not human-induced] are sufficient to explain global warming during the last 50 years.
2	'Reasonable, articulable grounds for suspicion'	Possible	1–10%	'Very unlikely'	Warming over the past 100 years is due to internal variability, as estimated by current models.The loss of grounded ice will lead to a substantial rise in sea level during the 21st century.
1	'No reasonable grounds for suspicion'	Unlikely: available evidence is against it, or violates existing paradigms, but not entirely ruled out	<1%	'Exceptionally unlikely'	(None)
0	'Impossible'	Against the known laws of physics or other science	0%		(None)

showing', (level 7), or even 'substantial and credible evidence' (level 6). However, given the scientist's habitual and ingrained caution, it is likely that (s)he would not pronounce him/herself to be 'reasonably certain' unless (s)he thought that the evidence had reached the level that a lay person would consider 'clear and convincing', i.e., sufficient for a civil trial or even a quasi-penal civil proceeding but not enough for a criminal conviction. Clearly, a witness using this formulation may need to be pressed for greater precision in defining the intended degree of uncertainty, a situation in which the proposed or some similar scale should be useful.

Neither law nor science demands absolute certainty. In particular, to say, as we do in Table 2, that the Theory of Relativity, for example, is established 'beyond any doubt', is not to question the validity of scientific propositions established to a lesser standard, but only to say that this theory has attained a higher status as an over-arching explanation of a wide range of phenomena.

The application of the proposed scale—or indeed, of any scale of scientific uncertainty—runs into difficulty when it is applied to long-established empirical methods. Many of these—like DNA testing and fingerprinting—are biometrics thought to be unique to each individual on the basis of wide experience. These are of crucial importance to forensic science, and indeed critical to a variety of technologies for personal identification thought to be critical to national and business security.

Are these technologies—fingerprinting, say—properly considered 'rigorously proven' within the meaning of Table 2? This is not a theoretical question. A recent case briefly called into question the validity of fingerprinting—a biometric hallowed by long usage—on the grounds that the uniqueness of the fingerprints of each individual had never been subject to adequate scientific proof.[23] The proposed distinction between 'rigorous' and 'substantial' proof does not settle this question, but does provide a criterion by which it can be considered. If some future court were to decide that the uniqueness of fingerprints has indeed not been established to the exclusion of the alternative explanations (i.e. that two different individuals may on occasion have identical fingerprints), then a coincidence between fingerprints found at a crime scene and those of a particular person will amount to a statistical correlation rather than an absolute identification, and would require additional evidence before they are sufficient for a conviction.

7. A scale worth trying

In summary, then, we have proposed an 11-point scale of scientific uncertainty ranging from 'beyond all doubt' (scale value of 10) at one extreme, to 'impossible' (scale value of zero) on the other. In between are scale values from nine to one, corresponding to 'beyond a reasonable doubt' (to a lawyer) or 'rigorously proven' (to a scientist), down to 'no reasonable grounds for suspicion' (to a lawyer) or 'unlikely, against available evidence' (to a scientist), respectively.

The scale takes advantage of the fact that there are many more standards of proof recognized in the US legal system beyond the familiar 'criminal' and 'civil' standards of 'beyond a reasonable doubt' and 'preponderance of the evidence', respectively, and

[23] *U.S. v. Plaza, Acosta & Rodriguez*, 2001, and 2002. The judge in this case changed his own first decision that had questioned the uniqueness of fingerprints.

that these correspond to levels of certainty or uncertainty that constitute acceptable bases for legal decisions in a variety of practical contexts. Furthermore, the levels of certainty or uncertainty corresponding to these standards of proof correspond rather well to the informal scale of certainty used by research scientists in the course of their everyday work, and indeed by ordinary people as they estimate the likelihood of one or another proposition. We have used these legal standards of proof, and this informal scale of scientific uncertainty, as the bases of a scale that should help to bring some order to discussions of uncertainty in a wide variety of policy fora. As with the standards themselves, the practical meaning of the proposed scale of scientific certainty will no doubt evolve with experience once it has begun to be applied.

Like any verbal scale of probability, the proposed scale is subject to the criticism that different people may assign a wide range of different meanings to the same word. Indeed, psychologists have found that even expert meteorologists—who might be expected to have a strong interest in the precise communication of degrees of uncertainty—impute a wide range of meanings to such 'vague' qualitative terms as 'probable' or 'possible' (Wallsten, 1990).

Nevertheless, we would suggest that the proposed scale should prove substantially superior to the vague terms tested in the psychological literature, and indeed will offer a valid alternative to the apparent objective precision of quantitative Bayesian scales, because the legal contexts surrounding the standards of proof on which the proposed scale is based should provide a fixed calibration point on which the public may anchor its interpretation. This would be consistent with the experimental finding that people deal with uncertainty by anchoring their estimate of uncertainty in an initial value, and then adjusting this estimate depending on the perceived degree of vagueness and their personal attitude towards risk (Einhorn & Hogarth, 1985; Wallsten, 1990).[24]

Like any subjective scale, the proposed scale can allow a person only to characterize his or her own judgements regarding uncertainty in a particular situation. Such a scale does not provide a means of resolving disagreements, but only of making them more precise. Nor does it stop advocates from using the categories for strategic purposes: i.e. to tailor the statement of uncertainty to the standard of proof required in a particular situation. In other words, a scale is only a tool; it is not self-policing.

The proposed scale seems on its face to have sufficient advantages that it is worth testing on a substantial scale as a complement to the estimation methods now in common use based on decision theory. The claims made for the scale do lend themselves to a variety of empirical tests. The claim that the scale is user-friendly can be tested with appropriate focus groups. The claim that it is useful by scientific advisory panels lends itself to pilot testing by the National Research Council, the IPCC, or other similar organizations. The claim that it is useful to scientific advocacy groups can be tested by such groups as the Union of Concerned Scientists or the Committee on Responsible

[24] Objective scales of probability, too, have their limitations. Indeed, the classic study of risk assessment by the National Research Council specifically counsels against using numbers that convey the impression of precision when such use is unjustified (National Research Council, 1982). Wallsten (1990) concludes from psychological experiments that 'there is no clear advantage in communicating numerically or linguistically', and makes no objection to 'a vague probabilistic estimate if that is all the information allows'.

Genetics. Historians, sociologists and anthropologists of science, as well as practitioners of technology assessment, can test its applicability to research in their respective professions.

We do not expect that the proposed scale of scientific certainty will succeed in separating science from values, or in ending disagreements over the status and validity of scientific assertions. On the contrary, it is unreasonable to expect any scale to be able to impose order on a freewheeling, high-stakes policy debate such as those on climate change or the possible destruction by synthetic chemicals ('endocrine disruptors') of humanity's ability to reproduce itself. Nor can we expect to force the multifarious aspects of science and public policy to fit into a single mould.

What we can expect is that the proposed scale will complement the quantitative scale already exploited by the IPCC, and that together the two scales will help to make the treatment of scientific uncertainty in public controversies involving science and technology more explicit, more precise, and more amenable to rational argument. One may hope that in the long run, this will help to improve risk assessment for policy makers, to increase public understanding of the role of scientific uncertainty as it affects public policy, to undermine some of the more absurd arguments now being presented to the public in the guise of scientific fact, and to increase the level of honesty in the presentation of scientific advice, scientific advocacy, and expert testimony.

Acknowledgements

The author thanks Tyler Vogel for excellent research assistance, and Jonathan Koehler for helpful comments.

REFERENCES

ADMINISTRATIVE PROCEDURES ACT (APA), 1946 Section 7(c), (5U.S.C.S. Section 556(d)).

ATIK, J. 1996/97 Science and regulatory convergence. *International Law and Business*, **17**, 736–758.

Banks v. State of Alabama 2002 WL1822104 (Ala. Crim. App.).

Bristol v. Microsoft 1998 US Dist, 42 F. Supp. 2d 153.

Cage v. Louisiana 1990 498 U.S. 39, 41.

CALMAN, K. C. & ROYSTON, G. H. D. 1997 Risk language and dialects. *British Medical Journal*, **315**, 939–942.

Commonwealth of Massachusetts v. John W. Webster 1850 59 Mass. 295; 1850 Mass. LEXIS 8; 5 Cush. 295.

Consolidated Edison v. National Labor Relations Board 1938 305 U.S. 197, 59 S. Ct. 206.

CORWIN, M. 2001 Issues on the third circuit: Defining 'Beyond a Reasonable Doubt' for the Criminal Jury. *Villanova L. Rev.*, **46**, 829–853.

Dunaway v. New York 1979 442 U.S. 200.

EINHORN, H. J. & HOGARTH, R. M. 1985 Ambiguity and uncertainty in probabilistic inference. *Psychological Review*, **92**, 433–459.

EVANS, A. S. 1976 Causation and disease: the Henle–Koch postulates revisited. *Yale Journal of Biology and Medicine*, **175**, 175–195.

FEDERAL JUDICIAL CENTRE 1987 *Pattern Criminal Jury Instructions: Report of the Subcommittee on Pattern Jury Instructions, Committee on the Operation of the Jury System,* Judicial Conference of the United States.

Florida v. Royer 1983 460 U.S. 491.

GIGERENZER, G. & HOFFRAGE, U. 1995 How to improve Bayesian reasoning without instruction: frequency formats. *Psychological Review*, **102**, 684–704.

GIGERENZER, G. & MURRAY, D. J. 1987 *Cognition as Intuitive Statistics*. Hillsdale NJ: Lawrence Erlbaum Associates.

GOTANDA, J. Y. 1993 The emerging standards for issuing appellate stays. *Baylor L. Rev.*, **45**, 809.

HENRION, M. & FISCHHOFF, B. 1986 Assessing uncertainty in physical constants. *American Journal of Physics*, **54**, 791–798.

Illinois v. Wardlow 1999 119 S Ct. 1573.

INTER-GOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC) 1995 *Report of Working Group I*, Available: http://www.ipcc.ch/pub/sarsum1.htm.

INTER-GOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC) 2001 *Report of Working Group I*, Available: http://www.ipcc.ch/pub/spm22-01.pdf.

JASANOFF, S. 1995 *Science at the Bar*. Cambridge MA: Harvard University Press.

KAYE, D. H. 1989 The probability of an ultimate issue: the strange case of paternity testing. *Iowa L. Rev.*, **75**, 109.

KENNEY, S. 1995 Supreme Court Review: fifth amendment—upholding the constitutional merit of misleading reasonable doubt jury instructions. *J. Criminal L. and Criminology*, **85**, 989–1027.

KLEE, R. 1997 *Introduction to the Philosophy of Science*. New York: Oxford University Press.

KUHN, T. S. 1970 *The Structure of Scientific Revolutions*, 2nd Edition. Chicago: University of Chicago Press.

LAFAVE, W. R. 1968 'Street Encounters' and the constitution: Terry, Sibron, Peters and beyond. *Mich. L. Rev.*, **67**, 40–87.

LAFAVE, W. R. & ISRAEL, J. H. 1992 *Criminal Procedure 2nd Edition*. St. Paul MN: West Publishers.

LAFAVE, W. R., ISRAEL, J. H. & KING, N. J. 2000 *Criminal Procedure 3rd Edition*. St. Paul MN: West Publishers.

LATOUR, B. & WOOLGAR, S. 1986 *Laboratory Life: the Construction of Scientific Facts*. Princeton: Princeton University Press.

LOEVINGER, L. 1992 Standards of proof in science and law. *Jurimetrics Journal*, **32**, 323–344.

LUO, M. 2002 Justice in a Small Town (3-part series). Associated Press, June 25, July 3 and July 7. Available on Lexis/Nexis.

Mazurek v. Armstrong 1996 94 F.3d 566, 567-568.

MCCAULIFF, C. M. A. 1982 Burdens of proof: degrees of belief, quanta of evidence, or constitutional guarantees? *Vanderbilt Law Review*, **35**, 1293–1335.

MERRIAM-WEBSTER, 1996 *Merriam-Webster Law Dictionary*, Available: http://www.lawyers.com/lawyers-com/content/glossary/glossary.html

MORGAN, M. G. & HENRION, M. 1990 *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge, UK: Cambridge University Press.

MUELLER, C. B. & KIRKPATRICK, L. C. 1997 *Evidence*. New York: Aspen Publication Co.

MULRINE, T. V. 1997 Reasonable doubt: how in the world is it defined? *American University L. J.*, **12**, 195–225.

NATIONAL RESEARCH COUNCIL 1982 *Committee on Risk and Decision Making*, Risk and Decision Making: Perspectives and Research. Washington: National Academy Press.

NESSON, C. 1985 The evidence or the event? on judicial proof and the acceptability of verdicts. *Harvard Law Review*, **98**, 1357–1392.

NEWMAN, J. O. 1994 Beyond 'Reasonable Doubt'. *New York University L. Rev.*, **68**, 201–204.

NEW YORK TIMES 2002 *Court Dismisses Guilty Plea in Baby's Death.* August 12.

PARK, R. L. 2000 *Voodoo Science: The Road from Foolishness to Fraud*. New York: Oxford University Press.

PARK, R. C., LEONARD, D. P. & GOLDBERG, S. 1991 *Evidence Law (hornbook)*. St. Paul, MN: West Group.

People of the State of New York v. Cecil Todd 1992 153 Misc. 2d 579; 582 N.Y.S.2d 352: Criminal Court Of The City Of New York, Bronx County.

People of the State of New York v. Edgar Bearden 1943 290 N.Y. 478; 49 N.E.2d 785.

PLANCK, M. 1949 *Scientific Autobiography and Other Papers*. (F. Gaynor, ed.). New York: Philosophical Library.

ROSENFELD, ARTHUR H. 1975 *Annual Review of Nuclear Science*, **25**, 555–98.

ROVELLA, D. 1998 Standard for Impeachment Report is Untested. *Washington Brief* 3 (August).

ROWLAND, F. S. 1993 The need for scientific communication with the public (AAAS Presidential address). *Science*, **260**, 1571–1576.

RYAN, G. GOV-IL 2000 Governor Ryan Declares Moratorium on Executions. Press release January 31.Available: http://www.state.il.us/gov/press/00/Jan/morat.htm.

SAGAN, C. 1996 *The Demon-Haunted World: Science as a Candle in the Dark*. New York: Ballantine Books.

Schmerber v. California 1966 384 U.S. 757.

SCHWARTZ, B. 1991 *Administrative Law*, 3rd Edition. Boston: Little Brown & Co..

SHAPIRO, B. J. 1986 To a moral certainty: theories of knowledge and Anglo–American juries, 1600–1850. *Hastings L. J.*, **38**, 153–193.

SHAPIRO, B. J. 1991 *'Beyond Reasonable Doubt' and 'Probable Cause': Historical Perspectives on the Anglo-American Law of Evidence*. Berkeley: University of California Press.

SOCIAL LEARNING GROUP 2001 *Learning to Manage Global Environmental Risks, 2 vols.*. Cambridge, MA: MIT Press.

Steadman v. SEC 1981 450 U.S. 91; 101 S. Ct. 999; 67 L. Ed. 2d 69: Supreme Court.

STRONG, J. W. (EDS) 1999 *McCormick on Evidence*, 5th Edition. St. Paul, MN: West Publishing.

Terry v. Ohio 1968 392 U.S.I.

US CODE 2001 Sec. 595(c) of Title 28.

U.S. v. Castrillon 1983 716 F.2d 1279; 1983 U.S. App.

U.S. v. Fatico 1978 458 F. Supp. 388.

U.S. v. Plaza, Acosta & Rodriguez 2001 179 F. Supp. 2d 444 and 464.

U.S. v. Plaza, Acosta & Rodriguez 2002 188 F.Supp.2d. 549.

U.S. v. Swift & Co. 1932 286 U.S.106, 119; 76 L. Ed. 999; 52 S. Ct.460.

U.S. v. United Gypsum Co 1948 333 U.S. 364.

Victor v. Nebraska; Sandoval v. California 1994 511 U.S. 1, 114 S.Ct. 1239.

WALLSTEN, T. S., FILLENBAUM, S. & COX, J. A. 1986 Base rate effects on the interpretations of probability and frequency expressions. *J. Memory and Language*, **25**, 571–587.

WALLSTEN, THOMAS S. 1990 The costs and benefits of vague information. *Insights in Decision Making. A Tribute to the Late Hillel Einhorn*. (R. Hogarth, ed.). Chicago: University of Chicago Press, pp. 28–43.

WEBSTER'S DICTIONARY, 1996 *New Edition*. W&R Chambers, Ltd.

WEISS, C. 2002 Scientific uncertainty in advising and advocacy. *Technology in Society*, **24**, 375–86.

WRIGHT, A. C., MILLER, A. & KANE, M. 1995 *Federal Practice and Procedure*, 2nd Edition.

WYNNE, B. 1992 Uncertainty and environmental learning. *Global Environmental Change* 2 (June), 111–127.

YEAZELL, S. C. 2000 *Civil Procedure*, 5th Edition. Gaithersburg, MD: Aspen.