REVIEW ARTICLE

The Truth Will Out? Incoherence and Scepticism in *Foundations of Evidence Law*

David Hamer*

Alex Stein, **Foundations of Evidence Law** Oxford: Oxford University Press, 2006, 264 pp, hb £50

INTRODUCTION

A trial is an enquiry into the facts of past events.[1] There will always be a separation between the evidence and the material facts. This may generate uncertainty and disagreement about the facts of the case, and provoke controversy about the objectives of evidence law and the rationality of the fact-finding enterprise. This is the terrain covered by Alex Stein's recent *Foundations of Evidence Law.*

Stein has previously theorised many rules and principles of evidence law.[2] Much of this work is incorporated into his new book, recast as supposedly coherent elements of a larger theory. Stein defends the edifice of evidence law against free proof abolitionists who view it as ramshackle, sprawling, and crying out for demolition.[3] Stein's normative agenda is heavily grounded in existing institutions and doctrine, and he often makes substantial descriptive claims.[4]

Stein's project is ambitious, moving between high abstraction and minute detail, and is frequently insightful. Ultimately, however, Stein fails to persuade. His vision contains major structural flaws. He delimits the domain of evidence law too narrowly, and frames the objectives of evidence law awkwardly. His treatment of related rules of evidence, such as standards of proof, appears ad hoc and inconsistent. More importantly, his understanding of the nature of inference, as revealed by his discussion of evidential weight and resilience, is seriously flawed. Ultimately, rather than making a persuasive case against free proof, Stein may have increased its attraction. *Foundations* provides a telling demonstration of the difficulty of bringing coherence to evidence law. Discretion being the better part of valour, this is an area from which the law may continue to remove itself.

*Senior Lecturer, TC Beirne School of Law, University of Queensland. I am grateful for the comments of the anonymous reviewer.

1 And, less commonly and more problematically, future and hypothetical events: A. Stein, *Foundations of Evidence Law* (Oxford: Oxford University Press, 2005) 34 (hereinafter referred to as Stein); D. Hamer '"Chance would be a fine thing": Proof of causation and quantum in an unpredictable world' (1999) 23 Melb ULR 557.

2 For Stein's bibliography, see http://www.professoralexstein.com/pages/2/index.htm (last visited 8 June 2006).

3 Stein 110, 138.

4 *ibid* xii, 110, 139.

 (2007) 70(2) MLR 318–338

Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA

THE EVIDENTIAL DOMAIN

Stein's first misjudgement is in the over-narrow demarcation of his topic. He limits the domain of evidence law to rules which he views as 'genuinely evidential'[5] – those serving the objectives of factual accuracy, efficiency and risk-apportionment.[6] He excludes rules which 'promote objectives extraneous to fact-finding',[7] even those whose 'design and operation are evidence-related'.[8] But any rule with an evidential operation will inevitably impact upon the operation of fact-finding and should be considered genuinely evidential. And many rules serve both evidential and extraneous objectives, defeating Stein's dichotomy. Both points argue for a less procrustean and more inclusive notion of the evidence law. Preferable is the approach of Weinstein who includes as an objective of evidence law, 'supporting independent social policies'.[9]

Many principles in the evidential heartland reflect concerns extraneous to fact-finding. The hearsay rule, for example, may be justified by reference to the factual accuracy goal. An out-of-court statement is viewed as untrustworthy as its maker has not taken the oath and is not available for cross-examination. Hearsay statements may exceptionally be admitted where accompanied by 'indicia of reliability' or 'circumstantial guarantees of trustworthiness'.[10] For example, because businesses and their customers constantly rely upon the accuracy of business records, courts may also consider them reliable.[11] But the hearsay rule is not only concerned with factual accuracy. As Weinstein points out, 'some of its vitality is due to its psychic value to litigants, who feel that those giving evidence against them should do it publicly and face to face'.[12] In the United States, the hearsay rule is identified with the accused's constitutional right of confrontation, which Tribe describes as an 'affirmation of respect for the accused as a human being'.[13]

The 'similar fact' rule, excluding evidence of the defendant's other misconduct, serves all three of Stein's genuinely evidential goals. The fact-finder may over-estimate the value of the evidence, diminishing factual accuracy.[14] The evidence may cause the fact-finder to not give the defendant the benefit of a reasonable doubt, and apportion the risk of error illegitimately.[15] Finally, other misconduct evidence can diminish the efficiency of proceedings by proliferating issues,[16] surprising the

5 *ibid* x.
6 n 58 below.
7 Stein 26. Given this it is not clear why he works so hard to demonstrate that 'there is no such thing as a free-standing rule of evidence, unassociated with any of the law's three fact-finding objectives': at 11–12, and see at 31–33. This is an immediate corollary of his definition of evidence law.
8 *ibid* 27.
9 J. Weinstein, 'Some difficulties in devising rules for determining truth in judicial trials' (1966) 66 Columbia LR 223, 241.
10 Stein 156. Stein appears to endorse the conventional rationale at this point, but he later sets up a competing theory, at 227–234.
11 eg *Albrighton* v *Royal Prince Alfred Hospital* [1980] 2 NSWLR 542, 548–549.
12 Weinstein, n 9 above, 245, and see R. Friedman, 'Infinite strands, infinitesimally thin: Storytelling, Bayesianism, hearsay and other evidence' (1992) 14 Cardozo LR 79, 99–100.
13 L. Tribe, 'Trial by mathematics: Precision and ritual in the legal process' (1971) 84 Harv LR 1329, 1392; see eg *Pointer* v *Texas* 380 US 400 (1965).
14 eg *DPP* v *Boardman* [1975] AC 421, 456; *Pfennig* (1995) 182 CLR 461, 478, 488.
15 eg R. Lempert (1977) 'Modelling Relevance' (1977) 75 Mich LR 1021 at 1034.
16 eg *Perry* (1982) 150 CLR 591, 586.

defendant,[17] and confusing the fact-finder.[18] The exclusion also serves extraneous objectives. To allow a defendant's prior offences to be used against him appears contrary to the notion of rehabilitation and human autonomy.[19] And use of prior offences could be self-fulfilling – the more the police focus on known offenders, the more difficult it is for them to rejoin mainstream society, leading to the creation of 'an underclass of "usual suspects".'[20]

Stein places both exclusionary rules within the domain of evidence law. He suggests that the rights they provide 'ultimately derive from epistemic fallibility, not from moral virtuousness'.[21] In support of this claim he suggests that, if fact-finders were infallible, there would be no room for 'evidential rights that are valuable intrinsically, rather than instrumentally'.[22] Epistemic fallibility is an inherent feature of our existence, so Stein's argument is of dubious value. But there could be room for these principles in a factually certain world. The defendant might still demand the right to confront his accusers as to the moral basis of their claims. And the defendant could still argue, for example on sentencing, that he should be judged on the basis of his present actions without reference to past misdeeds.

But Stein is correct in suggesting that hearsay and similar fact rules do not operate without regard for their costs. They do not provide 'free-standing' rights.[23] 'In law, balancing and trade-offs are unavoidable.'[24] While these exclusionary rules may be partly justified by reference to the defendant's moral rights, they often serve fact-finding goals at the same time. And where the moral goal and the fact-finding goal are in conflict, the former may give way. Evidence may be brought within an exception to the exclusionary rule so as not to inflict too great a cost on fact-finding. The most damning evidence of a defendant's past misconduct will get past the similar fact rule if it is sufficiently probative.[25]

Stein's 'anti-deontological point'[26] draws the hearsay and prior misconduct rules clearly within the evidential domain despite their moral content. But it also throws doubt on Stein's attempt to exclude rules that serve extraneous objectives through an evidential instrumentality. Given their inevitable impact upon fact-finding objectives, they should be drawn firmly within the evidential domain. An examination of two of Stein's examples will demonstrate the impossibility of his position.

A child, born to a mother living with her husband, is presumed to be legitimate.[27] According to Stein, this is a 'substantive rule of family law', rather than a part of the law of evidence, because it 'protects the stability of the family . . . at the

17 eg *R* v *Makin* (1983) 14 NSWR 1, 39–40; Evidence Act 1995 (Cth), ss 97(1)(a) and 98(1)(a).
18 eg Z. Cowen and P. Carter, *Essays in the Law of Evidence* (Oxford: Clarendon Press, 1956) 145.
19 eg A. Zuckerman, *Principles of Criminal Evidence* (Oxford: Clarendon Press, 1989) 232; Stein 32.
20 D.T. Wasserman, 'The morality of statistical proof and the risk of mistaken liability' (1991) 13 Cardozo LR 935, 953.
21 Stein 33.
22 *ibid* 33.
23 *ibid* 32–33; see also above n 7. This discussion flows on from his discussion of the evidential domain, although I found the connection between the two a little obscure.
24 Stein 32.
25 eg the abduction and sexual assault of a child in *Pfennig* n 14 above.
26 Stein 31.
27 J. Bray, 'The Increasing Vulnerability of the Presumption of Legitimacy', in E. Campbell and L. Waller (eds), *Well and Truly Tried: Essays on Evidence* (Sydney: Law Book Company, 1982) 10.

expense of accuracy in fact-finding.'[28] But in most of its manifestations, the presumption has not been wholly indifferent to factual accuracy. As Stein points out, the presumption 'favours erroneous decisions upholding the child's legitimacy over erroneous decisions that hold the child illegitimate'.[29] The degree of preference has varied over time and between jurisdictions, depending upon policy preferences. At common law the presumption of legitimacy had to be rebutted beyond reasonable doubt.[30] Under legislation in certain jurisdictions, the presumption of legitimacy was made irrebuttable.[31] In this form, exceptionally, the presumption was purely substantive,[32] showing no regard to contrary evidence or objective reality.[33] More recently, however, as the social and legal stigma of illegitimacy has waned, the objective facts have been given greater weight, and legislation has made the presumption rebuttable on the balance of probabilities.[34] Contrary to Stein's categorization, the presumption of legitimacy, except in its irrebuttable form, is concerned with fact-finding objectives and should not be excluded from the evidential domain.[35]

Also non-evidential, according to Stein, are rules excluding evidence that has been obtained unlawfully.[36] Again Stein suggests that such exclusionary rules 'are properly classified as extraneous to fact-finding'[37] since they 'override accuracy in fact-finding for the sake of other objectives and values',[38] namely 'the integrity of the criminal justice system and the acceptability of criminal verdicts'.[39] But it is doubtful whether these principles can be so neatly pigeonholed. Some of their value may derive from doubts that arise about the reliability of evidence that has been unlawfully obtained,[40] for example, such as involuntary confessions. Even where the reliability of improperly obtained evidence is confirmed – the murder

28 Stein 27, citing *Michael H* v *Gerald D* 491 US 110 (1989), 119–121, 124–32.
29 Stein 152.
30 *Cocks* v *Juncken* (1947) 74 CLR 277, 293–308 (Dixon J); *Morris* v *Davies* (1837) 5 Cl & Fin 163, 275; 7 ER 365, 404.
31 eg, Children (Equality of Status) Act 1976 (NSW), s 10(1), cited by P. Byrne and J. D. Heydon, *Cross on Evidence* (Sydney: Butterworths, 3rd Australian ed, 1986) 221.
32 K. Broun, 'The Unfulfillable Promise of One Rule for All Presumptions' (1984) 62 North Carolina LR 697, 700; L. Cohen, 'Presumptions According to Purpose: A Functional Approach' (1981) 45 Albany LR 1079, 1083.
33 E. Landowski, 'Truth and Veridiction in Law' (1989) II *International Journal for the Semiotics of Law* (1989), 29, 35–36; *MacCormick* v *Federal Commissioner of Taxation* (1984) 158 CLR 622, 646.
The irrebuttable presumption would arguably still give way to the presumption of innocence. For example, if a defendant charged with incest was seeking to prove that the complainant, although born to the defendant's wife, was illegitimate, the law would require him to do no more than raise a reasonable doubt. N. Bridge, 'Presumptions and Burdens' (1949) 12 MLR 273–284; Zuckerman n 19 above, 116.
34 eg Status of Children Act 1996 (NSW), ss 9, 15; Family Law Reform Act 1969, s 26; *Serio* v *Serio* (1983) 4 FLR 756, discussed in M. Redmayne 'Standards of Proof in Civil Litigation' (1999) 62 MLR 167, 180.
35 Although, for practical purposes it may be more sensible to deal with the variety of presumptions in the relevant substantive law textbooks rather than in an evidence law text: Byrne and Heydon, n 31 above, 167; J. D. Heydon, *Cross on Evidence* (Sydney: LexisNexis, 7th Australian ed, 2004) 285.
36 Stein 110.
37 *ibid* 26.
38 *ibid* 110.
39 *ibid* 26.
40 eg PACE, s 76(2); Evidence Act 1995 (Cth), s 85.

weapon is found where stated in the defendant's confession[41] – the application of the exclusionary rule should not be considered to lie outside the evidential domain. Impacting upon the operation of the rule will be the question whether procedural integrity 'can be served by means less corrosive of the judicial system's ability to ascertain the truth.'[42]

Under Australia's uniform evidence law, the trial judge has a discretion to admit improperly and illegally obtained evidence.[43] A factor influencing this decision is probative value,[44] since 'exclusion of an item of evidence is more likely to endanger accurate fact finding if the evidence is highly probative than if it is of minimal relevance'.[45] In the United States, the exclusionary rules have Constitutional force and operate far more stringently.[46] But concern has been expressed about freeing guilty defendants,[47] and exceptions have been created. Where the illegality is of a minor kind, for example an inadvertent technical defect in a warrant, the 'substantial social costs'[48] of the exclusionary rule may be considered too high. The impact of the exclusionary rule on factually accurate convictions is a live issue even in the United States, and the rule should be viewed as lying within the evidential domain.

Stein ill-advisedly narrows the evidential domain in one direction. But he sensibly extends it in another direction, to include rules that, although substantive in form, serve fact-finding objectives.[49] He suggests, for example, that the objective mental element of the provocation defence is substantively problematic,[50] and should be understood as evidential. It addresses the supposed ease of concocting a subjective mental state,[51] the risk that the defence would be misused, and resultant increases in costs of error and error avoidance. Given these objectives, Stein places the rule in 'the domain of evidence law, not in the criminal law domain. It advances evidential objectives rather than the goals of the substantive criminal law, such as deterrence, desert, and retribution.'[52] But again, Stein's approach is too strictly dichotomous.

The importation of objective mental states into the criminal law has certainly been influenced by proof difficulties. The difficulty of mounting successful prosecutions for sexual assault has become one of the most pressing issues in criminal justice. One response has been to make the defence of mistake objective – the

41 eg *R* v *Wray* (1970) 11 DLR (3d) 673.
42 Weinstein n 9 above, 228.
43 Evidence Act 1995 (Cth), s 138(1).
44 Evidence Act 1995 (Cth), s 138(3)(a).
45 *Evidence (Interim)* (Australian Law Reform Commission 26, 1985), vol 1, [964]; quoted in S. Odgers, *Uniform Evidence Law* (Sydney: Lawbook Co, 6th ed, 2004) [1.3.15140].
46 4th and 5th Amendments; *Weeks* v *US* 232 US 383 (1914), 393; *Silverthorne Lumber Co* v *United States* 251 US 385 (1920), 392 .
47 *Bivens* v *Six Unknown Federal Narcotics Agents*, 403 US 388 (1971), 416.
48 *US* v *Leon* 468 US 897 (1984), 907.
49 Stein 3.
50 *ibid* 4.
51 It may be questioned how difficult it really is for the prosecution to prove a subjective mental state. Mens rea is traditionally viewed an essential element to be proved by the prosecution to the usual criminal standard of proof: *Lambert* [2002] 2 AC 545 at [35]; *Edwards* [1975] 1 QB 27 (Lawton LJ); G. Williams, *Criminal Law, The General Part* (London: Stevens, 2nd ed, 1961) 903.
52 Stein 4.

defendant's mistaken belief in the complainant's consent must be honest *and reasonable*[53] – a significant retreat from the strong subjectivism of *DPP* v *Morgan*.[54] But such reforms do not stem solely from perceived proof difficulties. They are also motivated by the substantive view that, in some situations at least, a failure to meet a reasonable standard of behaviour *is criminal*.[55] Consider Arbour J's concern about a mistake defence in a sexual assault case being based 'exclusively on . . . unacceptable myths and stereotypes'[56] – 'for example . . . on [the defendant's] prior belief that the complainant certainly would consent to sexual relations with him, or on the complainant's passivity or lack of resistance or, worse, on his conception that "no means yes", or that the complainant is in fact consenting in her own mind even though she is expressing lack of consent.'[57]

Stein raises some interesting issues about the scope of evidence law. However, he defines evidence law too narrowly and seeks to draw too sharp a distinction between the evidential and the non-evidential. Consequently he risks missing crucial interactions between fact-finding, its context and the substantive law that it serves.

FACT-FINDING OBJECTIVES

As well as delimiting the evidential domain, Stein's fact-finding objectives are central to his larger argument against free proof. However, Stein's objectives and his vision of how evidence law pursues them are problematic. This weakens his defence of evidence law against its free proof critics.

Stein's objectives are: '(1) enhancement of accuracy in fact-finding . . .; (2) minimization of the expenses that fact-finding procedures and decisions incur; and (3) apportionment of the risk of error . . . between the parties to litigation'.[58] It is instructive to compare them with Weinstein's 'variety of ends served by our rules of evidence': [59] 'Among the goals – in addition to truth finding – . . . are economizing of resources, inspiring confidence, supporting independent social policies, permitting ease in prediction and application, adding to the efficiency of the entire legal system, and tranquilizing disputants.' By comparison with those of Weinstein, Stein's objectives are few, abstract and potentially overlapping. As will be seen later, these features give rise to some of the difficulties with Stein's theory.

53 Sexual Offences Act 2003, s 1(c); Home Office, *Protecting the Public* Cm 5668 (2002), 9, 17; see also Victorian Law Reform Commission, *Sexual Offences: Final Report* (2004), 413–415.

54 [1976] AC 182.

55 M. Weinberg, 'Moral Blameworthiness – The "objective test" dilemma' (2003) 24 Aust Bar Rev 173, 175; O. W. Holmes, *The Common Law* (Boston: Little Brown & Co, 1881) 49–51; consider also the 'rise and fall of inadvertent recklessness': Dori Kimel, 'Inadvertent Recklessness in Criminal Law' (2004) 120 LQR 548, 554, a case-note on *R* v *G* [2004] 1 AC 1034; and the growing body of literature questioning any attempt to sharply distinguish subjective and objective mental states: A. Candeub, 'Consciousness and Culpability' (2002) 54 Alabama LR 113; K. Ferzan, 'Opaque Recklessness' (2001) 91 *Journal of Criminal Law and Criminology* 597.

56 *R* v *Cinous* [2002] 2 SCR 3 [167]; see also Victorian Law Reform Commission n 53 above, 409–413.

57 *R* v *Cinous* [2002] 2 SCR 3 [167].

58 Stein 1.

59 Weinstein n 9 above, 241; quoted in Stein 36.

Stein and Weinstein both mention the factual accuracy goal first. Weinstein expressly identifies 'truth finding' as the 'central purpose',[60] noting that '[u]nless reasonably accurate fact finding is *assumed*, there does not appear to be any sound basis for our judicial system.'[61] Stein similarly observes that '[g]etting the facts right is a prerequisite to proper determination of the litigated entitlements and liabilities.'[62] Weinstein recognises that factual accuracy is a matter of assumption. As Stein points out, there is no way to definitively establish the correspondence between the facts as found and what actually occurred.[63] It is the difficulty of proving actuality that evidence law is designed to address. Stein rejects scepticism in favour of the more straightforward 'common-sense view' that humans have a 'well-integrated reasoning apparatus';[64] although, elsewhere his commitment to the rationality of the enterprise waivers. And Stein displays a willingness to sacrifice factual accuracy in order to achieve other goals.

Weinstein and Stein both recognise efficiency as a goal of evidence law. However, their conceptualizations of efficiency differ significantly. For Weisenstein, efficiency means minimizing the resources expended on litigation, particularly time and money. He recognises that this imperative may clash with the desire for factual accuracy: 'The hardest and most important job of a procedural system is to keep striking a wise balance throughout the various points of conflict.'[65] Stein has a far broader notion of efficiency. The costs to be minimised include the cost of error as well as the cost of error-avoidance.[66] And the cost of error is not limited to the immediate damage to the losing party as a result of the error,[67] but extends to social benefits such as deterrence and corrective justice.[68] Thus Stein's notion of efficiency incorporates many of the other goals specifically mentioned by Weinstein.

The breadth of Stein's conception of efficiency raises a concern. Although listed separately, factual accuracy, for Stein, is also a component of efficiency – a gain in factual accuracy is a reduction in the cost of error. This creates the risk that factual accuracy may be devalued; it may be sacrificed too readily for broader efficiency gains. Consider Stein's justification of the corroboration requirement in 'word against word' civil cases.[69] Stein suggests this requirement is a 'decisional shortcut' with a 'sound economic rationale'.[70] 'The number of . . . cases in which the fact-finders can *justifiably* assess the claimant's testimony as more probable than not is relatively small. Expending adjudication costs on the entire pool of "word against word" cases in order to identify these outliers is therefore economically

60 Weinstein n 9 above, 243; see also Zuckerman n 19 above, 7; J. Koehler and D. Shaviro, 'Veridical Verdicts: Increasing Verdict Accuracy through the use of Overtly Probabilistic Evidence and Methods' (1990) 75 Cornell LR 247, 250; W. Twining, 'Rationality and scepticism in judicial proof: some signposts' (1989) II *International Journal for the Semiotics of Law* 69, 72.
61 Weinstein n 9 above, 243(emphasis added).
62 Stein 10.
63 *ibid* 56.
64 *ibid* 57.
65 Weinstein n 9 above, 241; quoting from James, *Civil Procedure* 2 (Boston: Little Brown, 1965).
66 Stein 141.
67 eg *ibid* 144–146.
68 eg *ibid* 142, 148.
69 See authorities cited at *ibid* 242, fn 103.
70 *ibid* 137.

imprudent.'[71] Stein provides no empirical support for his statement that few uncorroborated claimants would be found sufficiently credible to succeed, and this may be questioned. Of more concern, however, is his suggestion that the court should consider such claimants 'outliers', avoid the cost of identifying them and disregard their just claims.

Stein's reasoning would have greater application to 'word against word' criminal cases. Given that the criminal standard is higher than the civil standard, there would presumably be even fewer cases where the complainant's uncorroborated testimony would have sufficient strength.[72] Sexual assault cases, for example, frequently turn into a credibility battle between complainant and defendant, particularly where there is no forensic evidence or consent is in issue. In many jurisdictions the law has required that a warning be given of the dangers of relying on uncorroborated complainant testimony. However, these requirements have increasingly been abolished.[73] One might have thought that Stein would argue for strengthened corroboration requirements, on economic grounds. Instead he endorses Roberts and Zuckerman's view that the requirements 'reflected sexist stereotyping of – predominantly, female – sexual assault complainants, rather than well-founded assessments of complainants' testimonial unreliability'.[74] Stein fails to reconcile this with his view of 'word against word' civil cases.

Stein's third and final fact-finding objective, apportionment of the risk of error, bears no resemblance to any of Weinstein's goals. Stein's objective is extremely broad in two dimensions. First, according to Stein, the pursuit of this objective is ubiquitous; 'evidential rules and principles . . . have a single all-important function: allocation of the risk of error'.[75] Second, this allocation may be made on utilitarian or deontological grounds,[76] in order to maximise efficiency, or to achieve fairness at the expense of efficiency. This is not a necessary opposition – a utilitarian apportionment may also be fair.[77] But the question can arise whether, in apportioning risk, individual rights should be allowed to trump utility.[78]

Stein considers risk-apportionment as the 'fundamental'[79] and 'key function'[80] of evidence law. This forms a major plank in Stein's case against free proof. Stein accepts the 'epistemic confidence doctrine',[81] according to which regulation of empirical inquiry is unnecessary and ill-advised. But this is insufficient to invalidate

71 *ibid* (emphasis in original).
72 Jeremy Gans, '"Whom do you believe?" Criminal Appeals, Conflicting Testimony and the Burden of Proof' (2000) 22 Syd LR 220.
73 Scots law still requires corroboration: *McNairn* v *HM Advocate* [2005] HCJAC 112; 2005 SLT 1071. This is currently under review: Scottish Law Commission, *Rape and Other Sexual Offences* DP 131 (2006), [7.26].
74 Stein 25, fn 76, discussing *Carmell* v *Texas* 529 US 513 (2000), quoting from P. Roberts and A. Zuckerman, *Criminal Evidence* (Oxford: Oxford University Press, 2004) 479; see also *Longman* (1989) 168 CLR 79, 85–86.
75 Stein 138.
76 *ibid* 2.
77 *ibid* 214.
78 *ibid* 2, 17; Ronald Dworkin, 'Rights as Trumps', in Jeremy Waldron (ed), *Theories of Rights* (1984), 153.
79 Stein xi.
80 *ibid* x.
81 *ibid* 113.

evidence law. '"[N]atural fact-finding" ideas could . . . only be plausible if adjudicative fact-finding were "natural" in some unadulterated epistemological sense. But it is not. . . . [V]alue-preferences – or, more precisely, preferences with respect to the allocation of the risk of error – permeate adjudicative fact-finding.'[82] And the prevalent 'moral scepticism' indicates that '[e]nforceable value-preferences can only be formed by social consensus mechanisms, such as law'.[83] Stein's argument against free proof is sophisticated and complex, and there is insufficient space to examine it fully here. Here I focus on Stein's claim that risk allocation is 'a pervasive phenomenon',[84] which fact-finders carry out 'not only in their ultimate findings – as they ought to under the applicable burdens and standards of proof – but also in selecting evidence (through admission and exclusion mechanisms) and in relying upon particular evidence as a basis for their findings'.[85]

Before assessing Stein's claim, his meaning should be clarified. On one level, Stein makes a point which is both inarguable and insightful. A decision on the admissibility of evidence may be wrong, just as the fact-finder's ultimate decision may be wrong. Exclusionary rules have the effect of allocating the risk of error just as burdens and standards of proof allocate the risk of error. To this extent, Stein's point cannot be denied. But Stein makes a far stronger claim. According to Stein, risk-apportionment is not only an inevitable by-product of the operation of exclusionary rules. Risk-apportionment is '*the* key function';[86] it is a principle of design. This claim is far more difficult to maintain.

To evaluate Stein's claim it is helpful to first consider the conventional risk-allocation rationale for the criminal standard of proof. The application of the criminal standard determines the ultimate issue – whether the defendant is convicted or acquitted. There are two possible erroneous outcomes – wrongful conviction (false positive) and mistaken acquittal (false negative). The former is viewed far more seriously than the latter.[87] Setting the criminal standard at a high level favours the defendant over the prosecution. It reduces the risk of the worse error, although at the same time it increases the risk of the less harmful error and the overall expected error rate.[88]

Compare this with Stein's risk-allocation analysis of a rule restricting the admission of evidence impeaching the criminal defendant.

> [B]road admission of prior-conviction might excessively intensify the fact-finder's distrust of criminal defendants as witnesses. Suppression of such evidence, however,

82 *ibid* 113.
83 *ibid* 112.
84 *ibid* 64.
85 *ibid* 64. This claim, if correct, would defeat Stein's use of fact-finding objectives to constrain the evidential domain. He suggests, eg, that the rule excluding illegally obtained evidence lies outside the evidential domain because it serves an objective 'extraneous to fact-finding': at 26, 110. But why would the rule not fall within the risk-apportionment objective? The exclusion would apportion the risk against the state, whose officers had acted illegally in obtaining the evidence, and in favour of the defendant, whose rights had thus been infringed.
86 *ibid* x (emphasis added).
87 *Re Winship* 397 US 358 (1970); *Van der Meer* (1988) 82 ALR 10, 31.
88 D. Hamer 'Probabilistic standards of proof, their complements and the errors that are expected to flow from them' (2004) 1 University of New England LJ 71, 87–96; http://tlc.une.edu.au/lawjournal/pdf/UNELJ_1–1_Hamer.pdf (last visited 29 Sept 2006).

> can also produce fact-finding errors. Unaware of the defendant's criminal past, the fact-finders may credit his or her testimony with greater credibility than it deserves. Both types of error would distort the ultimate probabilities of guilt and innocence. Because false positives and false negatives do not inflict similar harms, the distortions that are likely to occur do not cancel each other out. There is, therefore, a sound utilitarian reason for excluding evidence that reveals the defendant's criminal record.[89]

Here, the risk-allocation reasoning is strained. Unlike the standard of proof, the exclusionary rule does not have a direct impact on the outcome of the case. The immediate error will be either mistaken exclusion, with the consequence that the fact-finder gives the defendant too much credibility, or mistaken admission, with the defendant being given too little credibility. But, at this stage, no binding decision is made, and the cost of error cannot be brought to account. Stein notes that erroneous exclusion or admission would 'distort the ultimate probabilities of guilt and innocence',[90] and talks in terms of the relative costs of ultimate errors – false positives and false negatives. Erroneous admission might bring about a wrongful conviction. But it is misleading to equate the two. A wrongful conviction might have occurred in any event. And there are other possibilities. The erroneous admission of impeachment evidence may bring about or happen to coincide with a correct conviction. And despite erroneous admission, there may be a correct or erroneous acquittal. A similar set of outcomes can be associated with the erroneous exclusion of impeachment evidence. The uncertain impact of erroneous application makes it difficult to view the exclusionary rule as an exercise in risk-allocation. The motivations for the rule are more likely to be along the lines of those for the related similar fact rule discussed above – a concern that the prejudicial evidence will be misused by the fact-finder, interfering with the pursuit of factual accuracy, the proper operation of the criminal standard of proof, and the policy goal of rehabilitation.

Stein provides a novel and interesting view of the objectives of fact-finding. However, its normative or descriptive superiority to the orthodox view is not established. Despite his arguments to the contrary, factual accuracy should be recognised as the paramount concern of evidence law and should not lightly be sacrificed for the sake of economic efficiency. And although the risk of error is a pervasive phenomenon, there is no call for viewing risk-allocation as the 'principal objective of evidence law'.[91] Stein fails to make out his defence of evidence law against the free proof movement.

QUANTITATIVE STANDARDS OF PROOF

So far I have examined Stein's theory at the macro level – his view of the domain of evidence law and its defining objectives. I now consider a set of Stein's more detailed analyses – those relating to standards of proof. These principles govern the fact-finder's ultimate decisions, apportioning the risk of error between the par-

89 Stein 16, but see his argument favouring the ready admission of this evidence at 164–165.
90 *ibid* 16.
91 *ibid* 133.

ties. Many commentators have used decision theory to model the risk-apportionment process, prescribing probabilistic standards of proof at which the expected harm of erroneous decisions is minimised.[92] At times Stein appears to approve of this utilitarian approach, but at other times he rejects it without adequate explanation. Stein's theory of standards of proof appears incoherent and inconsistent, which undermines his criticism of those who would describe evidence law as 'largely ununified and scattered, existing for disparate and sometimes conflicting reasons'.[93]

The decision-theory model of standards of proof is simple and powerful. Suppose that the harm of an erroneous finding of liability is L, and the harm of an erroneous rejection of liability is R. If liability is proven to probability level *p*, then the *expected* harm from a finding of liability is:[94]

$$(1 - p) \cdot \mathrm{L}$$

That is, the probability that the finding of liability is wrong (the complement of the probability of liability), multiplied by the harm of a wrong finding of liability. Similarly, the expected harm from a rejection of liability is:

$$p \,.\, \mathrm{R}$$

To minimise the expected harm of a decision, liability should be found when this would generate less expected harm than rejecting liability; that is, when the first quantity noted above is less than the second quantity noted above. This will occur when,

$$p > \mathrm{L}/(\mathrm{R} + \mathrm{L})$$

When the probability lies below this value, no finding of liability should be made.

Stein uses this model to contrast the civil and criminal standards of proof.[95]

> [T]he P > 0.5 rule should apply . . . when false positives and false negatives are equally harmful, which normally is the case in civil litigation. When one type of error is more harmful than the other, fact-finders should follow a different rule. In criminal adjudication, for example, false positives (wrongful convictions) are generally considered more harmful than false negatives (wrongful acquittals). . . . [W]hen the disutility differential [L/R] equals 9/1 (convicting an innocent is nine

92 J. Kaplan, 'Decision Theory and the Factfinding Process' (1968), 20 Stanford LR 1065; M. L. DeKay, 'The Difference between Blackstone-Like Error Ratios and Probabilistic Standards of Proof' (1996) 21 *Law and Social Inquiry* 95; Hamer, above n 88. Nance described this as the 'conventional' model of standards of proof: D Nance, 'Evidential Completeness and the Burden of Proof' (1998) 49 Hastings LJ 621, 622.

93 Stein 110, quoting from J. D. Heydon, *Evidence: Cases and Materials* (London: Butterworths, 2nd ed, 1984) 3.

94 Terminology varies. A distinction could be drawn between the expected utility of correct decisions and the expected disutility of incorrect decisions. However, these two measures can be reduced into a simpler notion of 'cost' of error: DeKay n 92 above, 99–100. Conventions for mathematical notation also vary, but Stein unhelpfully adopts inconsistent approaches at different places in the book: 15, 149, 172. Here I use lower-case italic to indicate that the probability of liability, *p*, varies between different cases, and upper-case roman to indicate that the harms of error, L and R, are (assumed to be) constant for a given class of case.

95 Stein 14–16.

> times more harmful than acquitting a guilty criminal), adjudicators should convict the accused if the probability of his or her guilt is greater than 0.9.[96]

Stein applies the model more broadly. As mentioned above he explains the presumption of legitimacy in these terms.[97] Given that a mistaken finding that a child is illegitimate (L) is more harmful than a mistaken finding of legitimacy (R), a high standard of proof is imposed on the party seeking to prove illegitimacy. Similar reasoning applies in those serious civil cases where the defendant's liberty is at stake, through committal to a mental institution, deportation or denaturalization. To mistakenly deprive a person of her liberty is 'considerably more harmful' than to mistakenly continue her liberty.[98] The risks of error are 'manifestly asymmetrical', and a higher standard of proof is imposed on the party that would deprive the other of her liberty.[99]

But Stein rejects this explanation for a higher standard in civil cases involving criminal allegations such as fraud. Stein acknowledges that the defendant has more at stake than in a typical civil action – '[i]dentified as fraudulent, the person suffers a reputation loss'.[100] However, he suggests the higher standard is not to be explained by decision theory: 'This . . . does not aim at attaining the socially desirable ratio of false positives versus false negatives'.[101] Instead he argues that the higher standard attenuates overdeterrence.[102] Regardless of the value of this explanation – which is not necessarily inconsistent with the decision-theory approach – Stein's reasoning appears inconsistent. He fails to explain why the decision-theory analysis is applicable to serious civil cases involving a finding of illegitimacy and loss of liberty, but not serious civil cases involving allegations of criminal conduct.

Stein also expresses ambivalence about the application of decision-theory analysis to the criminal standard of proof. As noted above, at one point he uses decision theory to contrast it with the civil standard. At other points he strongly advocates the utilitarian approach to proof of crime. 'In law, balancing and tradeoffs are unavoidable'.[103] The rights claimed by a criminal defendant could 'seriously dilute deterrence'.[104] 'More innocent victims . . . would . . . become exposed to and, inevitably, suffer from crime.'[105] Consideration of the harm to victims from erroneous acquittals throws doubt on the notion that a wrongful

96 *ibid* 149.
97 n 29 above.
98 Stein 153.
99 In the United States, a third intermediate standard is imposed, requiring 'clear and convincing evidence'. In England and Australia, it is said that the ordinary civil standard applies, but its requirements take account of 'the gravity of the consequences flowing from a particular finding': *Briginshaw* v *Briginshaw* (1938) 60 CLR 336, 361–362. Another consideration mentioned by Dixon J in *Briginshaw* is the 'inherent unlikelihood' of more serious allegations, which may provide an alternative explanation for more stringent proof requirements: *Neat Holdings Pty Ltd* v *Karajan Holdings Pty Ltd* (1992) 110 ALR 449, 450; *Re H* [1996] AC 563, 586; see generally Redmayne n 34 above.
100 Stein 153.
101 *ibid*.
102 *ibid*, citing his article R. Bierschbach and A. Stein, 'Overenforcement' (2005) 93 *Georgetown LJ* 1743.
103 Stein 31.
104 *ibid*.
105 *ibid*, see also at 173–174.

conviction constitutes a 'special moral injustice'.[106] But then Stein does an about-face, favours deontology over utilitarianism, and endorses a 'next-to-certainty'[107] standard, which cannot be represented by any point on the probability scale 'except for the unrealistic 1'.[108] A standard set at any lower probability level would suggest that wrongful convictions were tolerable. Suppose, for example, the standard were set at 0.9, and in 20 cases guilt is proven to 0.95, all resulting in convictions. The expectation is that one these convictions was wrongful. That defendant would be entitled to ask, 'Why should I be sacrificed? Why me?'[109] Stein does not just empathise with the law's squeamishness about the inevitability of mistaken convictions.[110] He gives the 'next-to-certainty' standard his full normative endorsement. In Stein's view, this is the only way that the law can 'treat its citizens with equal concern and respect'.[111] There is insufficient space in this review to explore Stein's arguments in full. Here I highlight inconsistencies in his position, and the unworkability of his near-absolute non-probabilistic standard.

At one point Stein indicates that his criminal standard of proof 'should apply indiscriminately across the board', and that 'no trade off should be allowed between these requirements and the severity of the offence on trial'.[112] Stein's unyielding attitude is contrary to authorities which recognise that the 'beyond reasonable doubt' formulation is inherently flexible. 'Jurymen themselves set the standard of what is reasonable in the circumstances.'[113] '[T]here is no absolute standard . . . there may be degrees of proof within that standard. . . . [I]n proportion as the crime is enormous, so ought the proof to be clear.'[114] In a capital case, 'the prospects of the death penalty may affect [the jurors'] honest judgment of the facts . . . or what they may deem to be a reasonable doubt'.[115] And Stein himself does not hold the line, recognizing two situations where trade-offs lead to a relaxation in the criminal standard.

The first concerns criminal defences, such as insanity and diminished responsibility, which merely excuse the defendant's conduct[116] rather than fully justifying it, as with self-defence. The erroneous denial of a justificatory defence would be just as harmful as any other wrongful conviction, and such defences should be disproved by the prosecution beyond reasonable doubt.[117] However, an excusatory defence does not fully exonerate the defendant, and the harm of a mistaken

106 *ibid* 174.
107 *ibid* 178.
108 *ibid* 178.
109 Or, as Stein poetically puts it: 'why I?': *ibid* 214.
110 D. Kaye, 'The Laws of Probability and the Laws of the Land' (1979) 47 U Chicago LR 34, 40; Tribe n 13 above, 1372; C. Nesson, 'Reasonable Doubt and Permissive Inferences: The Value of Complexity' (1979) 92 Harv LR 1187, 1225; but see D. Shaviro, 'Statistical-Probability Evidence and the Appearance of Justice' (1989) 103 Harv LR 530, 544–5.
111 Stein 175.
112 *ibid* 179.
113 *Green* (1971) 126 CLR 28, 32.
114 *Bater* v *Bater* [1951] P 35, 36–37; endorsed in *Khawaja* v *Secretary of State for the Home Office* [1984] AC 74, 112–113.
115 *Adams* v *Texas* 448 US 38 (1980), 50 (White J). See also R. E. Lillquist, 'Recasting Reasonable Doubt: Decision Theory and the Virtues of Variability' (2002) 36 UC Davis LR 85.
116 Stein 180–181, also at 149–150.
117 For further discussion of the distinction between excuses and justifications, see eg, G. Fletcher, 'Two Kinds of Legal Rules: A Comparative Study of Burden-of-Persuasion Practices in Criminal

denial of the defence would not outweigh the harm to law enforcement flowing from mistakenly allowing the defence. There is a greater symmetry between errors than is usual in the criminal context, and the defendant should be required to prove an excusatory defence on the balance of probabilities.[118] Stein approves of this trade-off and the reduced standard of criminal proof it produces.

Stein also advocates a lower criminal standard of proof for 'administrative violations punishable by fines and injunctions alone'.[119] He does not elaborate, but this relaxation of the criminal standard is presumably on the basis that a mistaken finding of 'quasi-criminal' liability would be less harmful than for 'truly criminal' offences.[120] Meanwhile, a failure of law enforcement may severely compromise public protection. 'Thalidomide, Bhopal, Chernobyl and the Exxon Valdez can leave no doubt as to the potential human and environmental devastation'[121] flowing from quasi-criminal infringements. This is a second lowering of the criminal standard that seemingly meets Stein's approval, despite his 'no trade-off' stance.

A further problem for Stein's near-absolute criminal standard of proof is its unworkability. 'In the real world of human actions we can never be absolutely certain of anything'.[122] Of course, Stein appreciates this,[123] and his standard of proof requires not absolute certainty, but 'next-to-certainty'.[124] Although his standard would not require the fact-finder to achieve certainty, Stein is reluctant to concede that it carries an explicit risk of wrongful conviction. Underlying Stein's attempt to have it both ways is a difficult distinction between, as he terms them, Risk I and Risk II errors. Stein claims that the criminal standard immunises the defendant from the former but exposes him to the latter.[125] The distinction is not quantitative – the size of the risk – but qualitative.[126] Stein describes this as a 'doctrinal differentiation'[127] but then, inconsistently, suggests that courts are 'reluctant to articulate' it.[128] Stein himself provides various articulations which are neither consistent nor persuasive.

Stein first suggests that 'any perceptible doubt [Risk I] . . . must work in favour of the accused' whereas '[d]oubts that remain . . . imperceptible [Risk II] . . . do not pass the threshold of reasonableness'.[129] But Stein's attempt to distinguish his

Cases' (1968) 77 Yale LJ 880, 919. For criticism, see eg E. Colvin, 'Exculpatory Defences in Criminal Law' (1990) 10 OJLS 381, 382, 386–7, 390.

118 eg *R* v *Jordan* [2002] QB 1112 holding that a reverse persuasive burden on diminished responsibility is consistent with the presumption of innocence under the Human Rights Act 1998.

119 Stein 180.

120 eg *Lambert* [2002] 2 AC 545 [154]; *Sherras* v *De Rutzen* [1895] 1 QB 918, 922, but see A. Stein, 'After *HUNT*: The Burden of Proof, Risk of Non-Persuasion and Judicial Pragmatism' (1991) 54 MLR 570–576, arguing against relaxation of standard of proof for defence to charge of unauthorised tree-felling.

121 *Wholesale Travel Group Inc* [1991] 3 SCR 154, 250.

122 B. Shapiro, '"To a Moral Certainty": Theories of Knowledge and Anglo-American Juries 1600–1850' (1986) 38 Hastings LJ 153, 193; see also *Re Winship* n 87 above, 370.

123 eg Stein 2, 56.

124 *ibid* 178.

125 *ibid* 173; Zuckerman n 19 above, 134–140.

126 Stein 178.

127 *ibid* 174.

128 *ibid* 173, fn 6.

129 *ibid* 173. Stein muddies the water by comparing a conviction in the face of a perceptible doubt to a 'deliberately erroneous conviction', at 174. This expression suggests that the fact-finder *knew* that the defendant was innocent, rather than merely doubting his guilt.

standard from certainty fails. If a fact-finder has a doubt, that is, 'a *feeling* of uncertainty',[130] then, by definition, the fact-finder will have perceived it. If the standard requires the fact-finder to have no perceptible doubt about the defendant's guilt, then the standard requires certainty.[131] But this is too demanding. If it is acknowledged that absolute certainty about real world events is unachievable, doubt will always be perceptible. There may be occasions where a fact-finder unjustifiably feels absolute certainty. But these will be rare. A standard of certainty would result in too few convictions and a failure of law enforcement.

At another point Stein takes quite a different tack. He suggests that the criminal standard 'immunizes the accused only from the evidentially confirmed risk of erroneous conviction (Risk I) . . . [but] exposes the accused to the risk of erroneous conviction when the risk lacks evidential confirmation (Risk II).'[132] This version has the opposite difficulty. Stein does create a standard that clearly lies below certainty, but it appears too far below. To require that there is specific evidence of the defendant's innocence for acquittal contravenes the presumption of innocence.[133] Stein varies his expression slightly, describing the distinction as one between 'evidenced and case-specific . . . scenarios in which the defendant is innocent' (Risk I) and 'abstract and theoretical . . . scenarios' (Risk II),[134] but the distinction remains problematic. On this view, for acquittal, there must be a definite and concrete account of the defendant's innocence consistent with the evidence. There must be a 'doubt based on a reason'[135] capable of clear articulation by the fact-finder. But to suggest that 'a generalized unease or skepticism about the prosecution's evidence is not a valid basis to resist entreaties to vote for conviction'[136] still infringes the presumption of innocence. It sets the criminal standard of proof too low, and, contrary to Stein's stated aim, would provide insufficient protection for the innocent defendant.

Stein provides an example of how his criminal standard would apply to a bank robbery. The defendant, D, is identified as a person seen near the bank shortly after the robbery carrying a sub-machine gun. Fingerprints matching D are found at the bank. D's alibi witness, W, is thoroughly discredited by a prosecution recording of a conversation in which D promised W £500,000 for providing false evidence. Stein recognises that the fact-finder may not have absolute certainty – 'D may have been framed by corrupt police officers . . . it is also possible that his alibi conspiracy with W was, in his mind, his only chance of escaping false conviction'.[137] However, he suggests that D would be convicted. This conclusion is plausible, but it does not illustrate the application of Stein's standard. The defence scenarios may be insufficient to acquit, but, if the fact-finder gave them any consideration, they would clearly not be imperceptible. Nor is Stein justified in dismissing these

130 Oxford Dictionary of English (2nd rev ed, 2005) (emphasis added).
131 Stein 174.
132 *ibid* 173.
133 Compare Zuckerman n 19 above, 139.
134 Stein 178.
135 T. Mulrine, 'Reasonable Doubt: How in the World is it Defined?' (1997) 12 *American University Journal of International Law and Policy* 195, 224; J. Newman, 'Beyond "Reasonable Doubt"' (1993) 68 NYULR 979, 983 at fn 17.
136 Newman, n 135 above, 983.
137 Stein 179.

scenarios as 'completely unevidenced [and] theoretical'.[138] The defendant may not be able to point to any case-specific evidence supporting the scenarios, but this does not mean they are of no account. Courts are aware that '"fabrication of evidence . . . does occur", and at times has even been 'a routine feature of "police culture"'.[139] They also appreciate that even the innocent, when under suspicion, have a 'tendency. . . to wish to distance themselves from the persons or events connected with the alleged crimes and to endeavour to improve their position by falsehood'.[140] The difficulty with the defence scenarios is not qualitative – that they present Risk II rather than Risk I. The problem for the defendant is quantitative.[141] The defence scenarios simply lack sufficient probabilistic strength.

Decision theory provides a compelling account of how standards of proof are set at probabilistic levels that minimise the expected cost of error. Stein applies this widely, though not uniformly, to civil standards. And he balks at a probabilistic criminal standard, as it involves an explicit acknowledgement of the risk of wrongful conviction. But unless the law enforcement goal is abandoned, this risk cannot be avoided, and decision theory appears the best way of managing it. Stein fails to establish his qualitative near-certainty standard as a viable alternative to the probabilistic standard.

WEIGHT OF EVIDENCE

Stein supports the decision-theory model of probabilistic standards of proof in some contexts. However, he also argues that the probabilistic measure of proof, by itself, is inadequate. It is insufficient for a fact-finder to assess the probability level of the version of facts, given the body of evidence. According to Stein, the probabilistic criterion needs to be supplemented by a second 'weight' criterion.[142]

Stein suggests that the familiar 'naked statistical evidence' cases illustrate his point.[143] The plaintiff in the Blue Bus case,[144] having been negligently hit by an unidentified bus, brings an action against the Blue Bus Co solely on the basis that it owns 80 per cent of the buses in town. The probability of the Blue Bus Co's liability appears to be 0.80. In the Prisoners in the Yard case, 999 of 1000 inmates participated in the killing of a prison officer, the other unidentified inmate

138 *ibid*.

139 *Kelly* (2004) 218 CLR 216, 249–250, quoting from Queensland, *Report of Committee of Inquiry into the Enforcement of Criminal Law in Queensland* (1977), [26].

140 *Harris* (1990) SASR 321, 323.

141 Stein 178.

142 eg *ibid* 48. Stein relies heavily on L. J. Cohen *The Probable and the Provable* (Oxford: Clarendon Press, 1977). However, Cohen presented weight as an alternative inductive measure of probability, not one that supplements the more familiar mathematical measure.

143 This legal issue has long been recognised: *Sargent* v *Massachusetts Accident Co* 307 Mass 246 (1940); *Briginshaw* n 99 above, 361–362; J. McBaine, 'Burden of Proof: Degrees of Belief' (1944) 32 *Calif LR* 242. The issue has also been long discussed by philosophers and psychologists: J. M. Keynes, *A Treatise on Probability* (London: Macmillan, 1921) Ch 6; C. Hartshorne and P. Weiss (eds), *Collected Papers of Charles Sanders Peirce*, (Cambridge Mass: Harvard University Press, 1932) *Volume 2, Elements of Logic*, 421; D. Ellsburg, 'Risk, Ambiguity and the Savage Axioms' (1961) 75 *Quarterly Journal of Economics* 643; Cohen n 142 above.

144 Stein 78; Tribe n 13 above, 1340–1341, 1346–1350. See also A. Tversky and D. Kahneman, 'Evidential Impact of Base Rates' in D. Kahneman, P. Slovic, and A. Tversky (eds), *Judgement under Uncertainty: Heuristics and Biases* (Cambridge: Cambridge University Press, 1982).

playing no role. Every inmate claims to be the non-participant. The probability of each inmate's guilt appears to be 0.999.[145] In both cases, the evidence appears strong enough in probabilistic terms to satisfy the applicable standard of proof. However, many commentators assume that the evidence would be insufficient to establish liability.[146] Some, including Stein, suggest that the difficulty is that the evidence lacks sufficient weight.[147] However, Stein's description of the weight concept and its supposed advantages is uneven and unconvincing.

Stein's strongest argument highlights the non-epistemological benefits of weight. These have also been appreciated by other commentators. A proof lacking weight would be lacking in detail and generalised. Weight, as Stein points out, 'individualizes adjudicative findings of fact that attach to an individual event'.[148] This is important in a system which is 'traditionally strongly attached to individualized justice and strive[s] to arrive at the just result in the light of concrete circumstances of the case: Justice . . . can hardly be separated from details'.[149] As well as this moral benefit, a detailed concrete proof may also 'enhance the . . . the power of the law's substantive message'.[150] These qualities might justify a weight requirement, despite its other costs. At times this appears to be Stein's position.[151] But he primarily advocates the weight criterion on epistemological grounds.

Stein is equivocal in addressing the epistemological nature of weight. He suggests that '[e]pistemology cannot decide whether fact-finders should base their adjudicative decisions upon probability estimates that do not carry much weight.'[152] 'Weighty and non-weighty probabilities are apples and oranges. There is no common denominator to which these probabilities can be reduced in order to be compared against each other. Weighty and non-weighty probabilities are incommensurable.'[153] But, as James Logue comments, to remove weight from epistemology is 'a very unhappy shift . . . [H]owever weight or reliability is to be explained, it is undoubtedly an epistemic concept, and attempts to account for it or measure it in terms of non-epistemic criteria must be suspect'.[154] And, on other occasions Stein expounds the epistemic benefits of increasing the weight of evidence: 'Some determinations of probability are better evidenced and, consequently, weightier than others, and rational fact-finders ought to take this factor into account.'[155]

145 Stein 78; Nesson n 110 above, 1192–3.
146 eg Tribe and Nesson, nn 144,145 above. Some commentators have questioned this assumption: eg J. Brook 'The Use of Statistical Evidence of Identification in Civil Litigation: Well-worn Hypotheticals, Real Cases, and Controversy' (1985) St Louis ULJ 293, 299.
147 Stein 85; Cohen n 142 above; N. Cohen, 'Confidence in probability: Burdens of persuasion in a world of imperfect knowledge' (1985) 60 NYULR 385.
148 Stein 70.
149 M. Damaška 'Presentation of evidence and fact-finding precision' (1975) 123 University of Pennsylvania LR 1083, 1103–1104.
150 C. Nesson, 'The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts Thesis' (1985) 98 Harv LR 1357, 1391.
151 Stein 226.
152 *ibid* 133, see also at 84.
153 *ibid* 84.
154 J. Logue, *Projective Probability* (Oxford: Oxford University Press, 1995) 84.
155 Stein 132.

The epistemic problem which Stein considers the weight criterion to solve is the 'unrealized potential of the missing evidence to produce a different factual conclusion'.[156] Propositions flowing from low weight probabilities are 'too risky to rely upon. Decisions relying on such propositions are epistemologically questionable, if not altogether illegitimate.'[157] Stein asserts that weighty probabilities avoid this risk because they are 'reliable', 'resilient', 'invariant' and 'robust',[158] 'less likely to be shaken by potential additions to its information base'.[159] But these benefits are merely asserted. They are not demonstrated, and may prove to be illusory.

Promoters of the resilience concept sometimes invoke the statistical concept of confidence.[160] Consider, for example, the probability that the next ball drawn at random from an urn will be red, assessed from the frequency of red in past draws.[161] Ten draws resulted in five red balls. Three hundred of 600 draws were red. Both frequency figures support a probability of 0.5 that the next ball drawn will be red. But the estimate based upon 600 draws would be far more reliable. Statistically, the 95 per cent confidence interval for ten draws is 0.5 ± 0.32.[162] For 600 draws it is 0.5 ± 0.04.[163] This is a situation in which a greater weight of evidence has measurably increased reliability. But this inference problem is far removed from that of the juridical trial. The statistical confidence interval is derived from frequency data based upon repeated independent trials,[164] which is very different from a subjective assessment stemming from the accumulation of diverse items of trial evidence.

Since trials are generally postdictive, the (inaccessible) correct probability, based on (unavailable) ideal evidence, will either be one – the event happened –

156 *ibid* 120.
157 *ibid* 120.
158 The first three terms appear at *ibid* 48, and the last three at 82.
159 *ibid* 88.
160 N. Cohen n 147 above; Logue n 154 above, 91. Stein cites both, eg at 48 fn 51, but he also cites an incisive critique of N. Cohen: at 84 fn 54: D. Kaye, 'Apples and Oranges: Confidence Coefficients and the Burden of Persuasion' (1987) 73 Cornell L Rev 54.
161 The urn is opaque and the balls are all identical to the touch. The ball is replaced after each draw (or alternatively the urn contains an infinite number of balls) and the urn is shaken.
162 The 95 per cent confidence interval is most commonly used. Briefly, this can be roughly calculated as $p \pm 2\sigma$. The symbol σ represents the standard deviation, which can be estimated by $\sqrt{(pq/n)}$, where n is the number of samples, p is the frequency of 'successes' (eg red balls) and q the frequency of 'failures' (eg non-red balls): W. Mendhall, *Introduction to probability and statistics* (Massachusetts: Duxbury Press, 1979) 239–240. This formula is based upon the mathematics of the normal function which can be taken as an approximation of the binomial probability distribution. The normal approximation is better for larger sample sizes, however Mendenhall suggests that where the frequency is centred, as it is in the present case with $p = 0.5$, the normal approximation of a binomial distribution of only ten samples is still 'reasonably good': at 203.
163 The technical meaning of the confidence interval is too complex to explore here. See eg I. Evett and B. Weir, *Interpreting DNA Evidence: Statistical Genetics for Forensic Scientists* (Sunderland Mass: Sinauer Associates, 1998) 65.
164 The concepts of frequency and repeated independent trials can be generalized to a degree by de Finetti's notion of exchangeability: de Finetti, La prévision: ses lois logiques, ses sources subjectives, (1937) 7 *Annales de l'institut Henri Poincaré* 1, (English translation: De Finetti, B., 'Foresight: its logical laws, its subjective sources' in H. Kyburg, and H. Smokler, *Studies in Subjective Probability* (New York: J. Wiley, 1964)); see also Logue n 154 above, 68–7.

or zero – it did not.[165] It follows that any assessment towards the middle of the probability scale, no matter how weighty, cannot be considered resilient; fresh clinching evidence could shift it significantly. This is not a problem for Stein since he assumes that an increase in the weight of evidence, as well as bringing greater resilience, also has the effect that '[t]he estimate would come close to certainty'.[166] There is, Stein suggests, a 'logical relationship between probability estimates and their weights'.[167] Stein illustrates his position by reference to the issue of a defence witness's trustworthiness.[168] With no evidence either way, and two alternatives to choose from – trustworthy or not – Stein suggests that the initial probability of trustworthiness is 0.5. Evidence then indicates that the witness is a career criminal, and the probability of trustworthiness drops to 0.2. Further evidence shows that the witness, delivering exculpatory testimony, is the defendant's brother, and the probability of trustworthiness drops further to 0.05. This, for Stein, demonstrates that 'adherence [to a rigid weight standard] excludes from consideration middle-range probabilities, far removed from both certainty . . . and impossibility. A fact-finder certifying that "This testimony is as likely as it is not to be truthful" must be either lazy or uninformed.'[169]

But Stein's example is contrived and his claim about a necessary correspondence between weight and probability is fallacious. The Blue Bus and Prisoners in the Yard cases demonstrate that evidence of low weight can establish high and even extreme probabilities; this is their point. In Stein's witness example, it is unclear why the fact-finder began at 0.5, the level of maximal uncertainty. The lack of any specific evidence about this witness does not necessarily render the two alternatives – trustworthy, untrustworthy – equally probable.[170] Given the solemnity of the proceedings, penalties for perjury, and experience with other witnesses, the fact-finder may have started with a generalization that witnesses generally tell the truth in court, ascribing a probability of, say, 0.8 to the witness's trustworthiness.[171] The increased weight of evidence – regarding the witness's criminal record and relationship with the defendant – may then have been accompanied by an increase in uncertainty, not certainty.

The two items of evidence in Stein's witness example pointed in the same direction. Stein's 'logical relationship between probability estimates and their weights'[172] appears premised upon the various items of evidence being consonant.[173] But it is easy to create a plausible scenario in which dissonant evidence accumulates, positive and negative evidence cancelling each other out, providing no net increase in

165 Determinists would say that the same applies to predictions, and any probability other than zero or one merely reflects our lack of knowledge: Stein 126; E. Jaynes, *Probability Theory: The Logic of Science* (Cambridge: Cambridge University Press, 2003) Ch 10.
166 Stein 89.
167 *ibid* 176.
168 *ibid* 89–90.
169 *ibid* 90.
170 The conditions would not justify such a basic application of the principle of insufficient reason: *ibid* 44; Evett and Weir n 163 above, 7; Jaynes n 165 above, 331–333.
171 As does Stein in another example a few pages earlier at 77.
172 *ibid* 176.
173 G. Shafer *A Mathematical Theory of Evidence* (Princeton: Princeton University Press, 1976), 225–226, notes that L. J. Cohen's theory of weight also makes an assumption of consonance.

certainty. The witness was a career criminal. But he converted to Christianity in prison and claims to have turned a new leaf. The witness is testifying in favour of his brother. But they have been estranged for many years. A fact-finder might be informed and motivated, but still have considerable uncertainty as to a witness's trustworthiness.[174]

Stein's commitment to the logical relationship between weight and certainty is not unwavering. At one point he decries the postulate of 'a linear progression relationship between the amount of information that fact-finders have and the accuracy of their decision' as 'unwarranted', 'fallacious' and 'plainly wrong'.[175] But the possibility that accumulating evidence will be more dissonant than consonant could undermine the entire fact-finding enterprise. Since Thayer it has been recognised as fundamental that probative evidence must be admitted unless there is good reason for exclusion.[176] But what is the point of admitting a probative item of evidence if its impact may immediately be undone by the succeeding item of evidence?[177] Whether probative evidence is admitted or excluded, either way there will be an 'informational void'.[178] Stein appreciates that to abandon notions of consonance and weight altogether is to engender a 'global epistemic scepticism'.[179]

However, scepticism can be averted without going to the other extreme and adopting unwarranted assumptions of consonance and resilience. A more modest solution would be to adopt some version of the principle of total evidence.[180] The fact-finder should make reasonable efforts to ensure that any probability assessment takes account of all available probative evidence. Questions as to the whether this principle is *a priori* may arise,[181] but it would appear to reflect human experience that decisions made on a greater quantity of evidence are generally better.[182] This principle provides the necessary motivation for a fact-finder to increase the weight of evidence. And it does so without the dubious assumption that the evidence will be consonant, or that weight will be accompanied by resilience or increased certainty. Provided the principle is complied with and relevant evidence is not ignored, a low weight probability assessment should not be viewed with any greater suspicion than a weighty probability assessment.

Stein expresses ambivalence about the notions of weight and resilience, and is irresolute in affirming the rationality of the fact-finding endeavour. This suggests an alternative counterargument to free proof. As shown above, Stein was unsuc-

174 Actually, this is a difficult example. Unlike a past event, certainty about a witness's trustworthiness may not be possible, even with ideal evidence.
175 Stein 122–123.
176 J. Thayer, *A Preliminary Treatise on Evidence at Common Law* (Boston: Little Brown, 1896) 265; J. Montrose, 'Basic concepts of the law of evidence' (1954) 70 LQR 527–532; compare Stein 15.
177 Logue n 154 above, 86; I. Good, 'On the Principle of Total Evidence' (1967) 17 *British Journal for the Philosophy of Science* 319, 319.
178 Stein 131.
179 *ibid* 123.
180 See R. Carnap, Logical Foundations of Probability (Chicago: University of Chicago Press, 1950) 211–213; Jaynes n 165 above, 338–339.
181 Good, n 177 above; B. Skyrms, *Choice and Chance* (Belmont: Wadsworth, 4th ed, 2000) 155.
182 See n 176 above; also consider the concept of 'discrimination' in scoring the accuracy of real world forecasters: J. Yates, 'Subjective Probability Accuracy Analysis', in G. Wright and P. Ayton (eds), Subjective Probability (Chichester: John Wiley & Sons, 1994) 381, 391.

cessful in finding a role for evidence law in the value-laden risk-apportionment decisions that supposedly pervade the fact-finding process. But he may have been too swift in upholding the epistemic confidence doctrine. Stein's analysis of weight demonstrates that there are major uncertainties about how fact-finders should best evaluate a body of evidence in their pursuit of factual accuracy. There may be a significant role for evidence law in authoritatively and advantageously resolving these issues.[183] Of course, this assumes such issues are capable of resolution, and that the law would be an appropriate vehicle for doing so, which may both be contentious propositions.

CONCLUSION

Stein's *Foundations* challenges the orthodox view of evidence law. According to Stein, factual accuracy is not the paramount goal. Evidence law's central function is to allocate the risk of error on the basis of cost-benefit analyses and moral rights. Stein views the increasing trend towards free proof as misguided and dangerous. Fact-finders may be trusted with purely epistemic enquiries, but risk-allocation is value-laden and should be settled authoritatively. Stein considers it appropriate that evidence law intervene widely, guiding every stage of proceedings where evidence is admitted, used and brought to account.

However, Stein's argument fails at a number of points. He does not establish the ubiquity of risk allocation or its ascendancy over factual accuracy and other concerns. And his vision of evidence law as a coherent and unified body of principle is not sustained. On the contrary, his analyses suggest that evidence law is multi-faceted, contestable, and frequently inconsistent.

But Stein does not leave the free proof agenda unscathed. On the contrary, although the central thrust of Stein's argument misses its mark, he raises important questions about the values that evidence law should serve, and how this is best achieved. Stein may not effectively defend the existing structure of evidence law, but he does demonstrate the need for us to understand it better before we tear it down.

183 Among other inferential issues requiring resolution is the impact of the right to silence, in particular, on correct and incorrect acquittals: Stein 8–9; 158–165, 200–204; D. Seidmann and A. Stein, 'The Right to Silence Helps the Innocent: A Game Theoretic Analysis of the Fifth Amendment Privilege' (2000) 114 Harv LR 430; G. Van Kessel, 'Quieting the Guilty and Acquitting the Innocent: A Close Look at a New Twist on the Right to Silence' (2002) 35 Ind LR 925; D. Hamer, 'The Privilege of Silence and the Persistent Risk of Self-Incrimination' (2004) 28 Crim LJ 160–178, 200–216.