Laudan's Error. Reasonable Doubt and Acquittals of Guilty People

Journal:	*The International Journal of Evidence and Proof*
Manuscript ID	EPJ-20-0014
Manuscript Type:	Original Article
Keywords:	False Acquittals, Proof Beyond Reasonable Doubt, Standard of Proof, Larry Laudan, Rate of Error
Abstract:	Proof beyond a reasonable doubt (BARD) is one of the most fundamental requirements of American criminal law and other legal systems. Professor Larry Laudan has criticized this requirement for several reasons. His main contention is that the BARD formula converts evidential support into subjective confidence, and is therefore not a genuine standard of proof. At the same time, Laudan holds that BARD produces a large number of guilty defendant's acquittals due to its excessive demand for evidence. The aims of this article are twofold: firstly, to show the inconsistency between these theses; and secondly, independent of this inconsistency, Laudan's argument is unacceptable regarding the number of guilty defendant's acquittals. Perhaps the real ratio of false negatives to false positives were what Laudan holds them to be, yet he fails to provide any suitable argument to support his claim, or to attribute the alleged frequency of errors to a particular standard of proof – BARD or otherwise.

Laudan's Error. Reasonable Doubt and Acquittals of Guilty People

Diego Dei Vecchi*

Abstract: Proof beyond a reasonable doubt (BARD) is one of the most fundamental requirements of American criminal law and other legal systems. Professor Larry Laudan has criticized this requirement for several reasons. His main contention is that the BARD formula converts evidential support into subjective confidence, and is therefore not a genuine standard *of proof*. At the same time, Laudan holds that BARD produces a large number of guilty defendant's acquittals due to its excessive demand for evidence. The aims of this article are twofold: firstly, to show the inconsistency between these theses; and secondly, independent of this inconsistency, Laudan's argument is unacceptable regarding the number of guilty defendant's acquittals. Perhaps the real *ratio* of false negatives to false positives were what Laudan holds them to be, yet he fails to provide any suitable argument to support his claim, or to attribute the alleged frequency of errors to a particular standard of proof – BARD or otherwise.

1 Introduction

For almost two decades Larry Laudan has been directing his philosophical efforts to analyse how decisions concerning the sufficiency of evidence are made in our criminal justice systems. Errors, their frequency, and consequences are at the core of his worries; particularly, the kinds of errors he calls *false verdicts*, to wit: verdicts whose factual premise is wrong. He distinguishes two types of false verdicts. On the one hand, convictions against innocent people also known as false positives or false convictions. On the other, acquittals in favour of guilty defendants also known as false negatives or false acquittals. False negatives may emerge from a prosecutor's decision to drop charges or can stem from a false verdict after the trial has ended. This paper focuses on Laudan's analysis of false negatives as a consequence of the trial.

* For very helpful comments, suggestions and corrections about earlier drafts, I am grateful to Edgar Aguilera García, Jorge Baquerizo, Hernán G. Bouvier, Carolina Fernández Blanco, Jordi Ferrer Beltrán, Daniel González Lagier, Diego Papayannis, Pablo Rapetti, and Cármen Vázquez. Different versions of the paper also benefited much from friendly discussion and comments from audience members on three occasions: the first, at the University of Girona, Cátedra de Cultura Jurídica, October 2019; the second, at the University of Alicante, February 2019; the third one, at the BIAP Workshop on Evidence, February 2020.

To completely avoid errors of every kind is incompatible with any course of action. The absolute avoidance of false positives and false negatives entails not making any decision at all regarding citizens' guilt. Yet, not deciding also seems to be an error: no community seems willing to resign judgment and the punishment of crimes.[1] Under these circumstances, legal systems resort to certain tools, the apparent function of which is to satisfy two needs: first, to set the threshold of evidence for accepting certain factual propositions as true and act as a result of this acceptance; and second, to distribute the errors involved when deciding and acting in this way. One of the most important tools to these effects is the standard of proof (SoP).

The phrase 'Beyond a Reasonable Doubt' (hereafter, BARD for short) has been accepted in many different criminal justice systems as the proper formula to set out the SoP. The hope is that the formula successfully sets a threshold of evidence, which when applied, reflects a value judgment according to which convicting an innocent person is worse than acquitting a guilty one. It is generally accepted that BARD instantiates the so-called Blackstone's *ratio*: ten false acquittals for every false conviction.

Larry Laudan has argued a great deal in favour of abandoning BARD, and gives two main reasons. We shall call Laudan's first claim against BARD *the confusion thesis*, as it sets out that BARD unacceptably confounds epistemic reasons (the right kinds of reasons to justify acceptance of factual propositions[2]) with subjective confidence. In this respect, he argues that BARD is not a genuine SoP as it is open to several different interpretations, the majority of which lead to confusion between the degree of proof and subjective confidence. The second claim against BARD is that it produces a very large number of false acquittals. The argument here is twofold. On one hand, Laudan claims that BARD admits more false acquittals than a rational-legal system would tolerate, given that (in his moral view) Blackstone's *ratio* is not only groundless, but pernicious. We shall call this claim the *consequentialist thesis*, since the pernicious character of BARD lies in its potentially harmful consequences;[3] however, this line of argument is not dealt with here. On the other hand, according to Laudan, BARD not only tolerates too many false acquittals in accordance with its underlying justification (i.e. the Blackstone *ratio*), the main problem is that when BARD is actually applied, it generates many more false

[1] See Forst, 2004, p. 22 ss. and Whitman, 2008, § 7 regarding the origins of BARD.
[2] The idea of *reasons of the right kind* evokes Strawson, 1962. More recently, Darwall, 2006, §§ 8 and 9; Darwall, 2013.
[3] Truthfully, these harmful consequences are the result of the Blackstone *ratio*, which Laudan seems to assume is the 'underlying reason' behind BARD. See Schauer, 1991, § 4.1.

acquittals than even Blackstone's *ratio* tolerates. We shall call this claim the *failure thesis*. I will focus on this facet of Laudan's "too-many-false-acquittals" claim.

I will argue that the *confusion thesis* and the *failure thesis* appear to be inconsistent, but the main goal of this paper is not to make this inconsistency explicit. Laudan's apparent inconsistency is just a touchstone to argue that the *failure thesis* must be rejected. By revealing what the failure thesis requires in order not to be inconsistent with the *confusion thesis* opens the door to showing that Laudan's reasoning in favour of the *failure thesis* is absolutely unacceptable. He provides no suitable ground to support his claim, nor to attribute the outcome of the alleged high number of false negatives to a particular SoP – BARD or otherwise. Paradoxically, some of the arguments Laudan himself develops in favour of the *confusion thesis* will be helpful here.

The paper proceeds in four steps. First, I will briefly refer to the minimal conditions a standard must satisfy in order to be an authentic standard *of proof* (§1). I will then recreate Laudan's arguments to ground the claim that BARD cannot function as an authentic SoP. This is the *confusion thesis* (§2). Third, I will re-depict Laudan's arguments to estimate the *ratio* between false positives and false negatives produced in the U.S. criminal justice system (for violent crimes) and to (partially) blame BARD for these errors. This is his *failure thesis* (§3). Finally, I will demonstrate that Laudan's argument is either highly inconsistent (§4.1) or groundless (§4.2).

1 BARD Under Suspicion

One common assumption in Laudan's writings is that standards of proof are the most appropriate tools for distributing errors in criminal cases.[4] He thinks that the greater the exigency of proof, the lower the number of false convictions; but at the same time, this increases the number of false acquittals. Symmetrically, the lower the standard, the higher the number of false convictions; however, the number of false acquittals would diminish and punishment for the guilty would increase. The distribution of error by means of a SoP consists precisely in setting a threshold of evidence, the outcomes of which should reflect the *moral* judgment on the importance of both types of errors.

In Laudan's view, the role of the criminal law system is to accommodate the part of the social contract regarding the correct distribution between false positives and false negatives.[5] The confidence he has in how productive the standards of proof are in this

[4] See Laudan, 2006; Laudan, 2007; Laudan, 2011; Laudan, 2016; Laudan & Saunders, 2009.
[5] Laudan,2011; Laudan & Saunders, 2009.

respect contrast with his vehement repudiation of almost every exclusionary rule.[6] In any case, for a criterion to be a genuine SoP, it should be defined in epistemic terms, and be a threshold of the right kinds of reasons, to wit: epistemic reasons, that is, evidence.

It is common to recognize that at best, pure epistemology only provides a finite set of criteria to determine the best among several concurrent hypotheses or explanations. This amounts to deploying an ordinal comparison.[7] However, 'hard-core' epistemology is not the appropriate discipline to evaluate the moral significance of potential errors, nor can it provide an adequate *ratio* of distribution between them. In other words, epistemology, fails to provide reasons for action. Instead, it seems to be restricted to epistemic reasons and to reducing global error.[8] The thesis defending the *existence* of a hard-core in epistemology is quite controversial. It supposes the possibility of a kind of *epistemic view from nowhere*, which one can check objectively to see whether a set of (exclusively) epistemic reasons are sufficient to justify a belief.[9] Therefore, there would be a purely epistemic criterion of completeness for epistemic justification.

The controversial character of this thesis lies in that it is not plausible to conceive a criterion of completeness for epistemic justification which is independent of action. On one hand, BELIEF itself is conceptually related to action, at least according to some philosophical conceptions. Every belief entails a disposition for certain actions,[10] so completeness of epistemic justification would be a function of what actions every particular belief is a disposition for, as well as of its possible consequences. Consequently, there would be no place for a purely epistemic criterion of complete epistemic justification: completeness depends at least in part on *practical* considerations. On the other hand, even if understanding BELIEF as being inert (i.e. a propositional attitude independent of every possible action) were plausible, in most cases accepting factual propositions as being true blazes a trail for action: ranging from assertions about things one believes, to any other non-linguistic action undertaken under the assumption that the proposition believed is true. Again, completeness of epistemic justification would be a function of the action every particular belief is blazing a trail for. Consequently, there is no place for a purely epistemic criterion of complete epistemic justification. On the contrary, having sufficient epistemic reasons (i.e. evidence) depends at least in part on

[6] Laudan, 2006.
[7] See Harman, 1965; Lipton, 1991; Pardo & Allen, 2008. Cfr. Laudan, 2007.
[8] Laudan, 2006, § 5. See also Stein, 2005, p. 13.
[9] See Alston, 1989 [1985], pp. 82-83; Montmarquet, 2007.
[10] See Haack, 2010.

the practical context and ultimately entails a practical decision.[11] To decide if *p* is proven in a certain context in which if a person accepts *p* he or she will do φ is the upshot of a practical judgment regarding the importance of the possible negative consequences of doing φ mistakenly because *p* is false.

> Our ordinary talk of someone's being 'completely justified' in believing something is highly context-dependent; it means something like: 'in the circumstances –including such matters as how important it is to be right about whether p, whether it is A's particular business to know whether p, etc., etc.– A's evidence is good enough (supportive enough, comprehensive enough, secure enough) that he doesn't count as having been epistemically negligent, or as epistemically blameworthy, in believing that p'. This may be represented by 'A is completely justified in believing that p', which will refer to a context-dependent area somewhere vaguely in the upper range of the scale of justification. Its vagueness and context-dependence is what makes this ordinary conception useful for practical purposes (and for the statement of Gettier-type paradoxes).[12]

Nevertheless, we have seen that in the legal procedural context, judges or jurors (depending on the system) must make decisions regarding the sufficiency of the evidence in order to accept a concrete proposition, and then on this basis, *to act* by either convicting or acquitting the defendant. Thus, in the legal context, hard-core epistemology would be ineffective in guiding triers of fact when making their decisions pertaining to the sufficiency of the evidence. There are two possible options: the first one would be allowing every trier of fact to decide by him or herself whether a set of evidence is good enough to undertake the action at hand in each single (*token*) case. In other words, evaluating the moral acceptance of convicting (or not) in a certain case *in the light of a particular body of evidence* would be a legal power conferred to every single trier of fact; the second option would be to set general and objective standards of proof for different kinds of cases (*type*-cases), which would imply making the epistemic exigency of every SoP consistent with (or a reflection of) an *a priori* evaluation regarding the risk of error involved in every different *type*-action considered. Here, the moral judgment of whether to convict or not in certain cases *in the light of a singular objective threshold of evidence* would be the underlying justification of a general rule: the SoP.

[11] See Walen, 2015.
[12] Haack, 2009 [1993], pp. 133-134.

Laudan is a faithful defender of the second strategy and a fierce critic of the first in equal measure. He believes that in the legal context, every decision on the sufficiency of evidence should be preceded, and guided, by a previous enactment of an objective SoP. This brings us to what Laudan calls the 'soft-core' of epistemology, where an *a priori* moral judgment underlies the choice of how demanding the SoP should be in order to decide if *p* is proven in a certain legal context, and to convict on that basis. This judgement recognises the importance of the possible negative consequences of convicting (or acquitting) mistakenly due to *p's* being false while taken as proven (or true while taken as not proven). According to this judgment, the particular threshold of evidence morally fits making this decision in certain *type* cases.

Nowadays it is usual to assume that the criminal SoP must reflect the moral judgment that convicting the innocent is worse than acquitting the guilty.[13] As Laudan asserts, the "collective decision" indicates a clear preference for false acquittals over false convictions. As mentioned earlier, one of his greater worries in this respect is the lack of any foundation for a particular *ratio* of distribution, especially for Blackstone's *ratio*. How many false acquittals is a society really willing to accept? How many is it morally acceptable or desirable to tolerate?

For Laudan, a well-founded SoP should have certain necessary features: to wit, its height should be settled in a nonarbitrary way; it should incorporate *all* the benefit of the doubt that we consider appropriate to give the defendant; it should guarantee the acceptable outcome of errors over the long run; and most importantly for present purposes, it should "unlike BARD, be cashed out in non-subjectivist terms."[14]

For a SoP to be a genuine standard *of proof*, the threshold should be set exclusively in *epistemic terms*. It should not depend on the jurors' subjective evaluation of guilt, "but on the establishment by the prosecutor of a powerful inferential link between the evidence presented and the guilt of the accused."[15] In other words, even if a value judgment on a *fair* distribution of error underlies the rule, the rule will count as an authentic SoP if, and only if, that value judgment is translated into an objective threshold of epistemic reasons. In order to have a genuine SoP we would need to make

[13] What must truly be reflected is that judgement is the criminal procedural system as a whole, and that the SoP is just one of many tools used. However, because of the way Laudan sets out the argument, we can reduce it here to the SoP.
[14] Laudan, 2006, p. 64. See also Ferrer Beltrán, 2018.
[15] Laudan, 2006, p. 81

a moral judgment regarding the desirability of a particular distribution of error which would then need to be *expressed* using epistemic criteria alone.[16]

2 BARD's first conviction: nuances of subjectivism and the confusion thesis

In 2003, Larry Laudan published an illuminating essay,[17] many of the conclusions of which were included in the second chapter of the book he published three years later.[18] In both works Laudan points out that BARD constitutes an empty formula, both as an epistemic threshold of evidence and as an expression of a particular moral *ratio* of error distribution. This is especially because of its radical subjectivism.

Although Laudan does not formulate the point in exactly this way, I believe we can distinguish a few nuances in this subjectivism. Indeed, Laudan begins his battle against BARD by setting down a repertory of explanations and interpretations of the formula made by judges and jurists in the United States throughout history, registering the following interpretations:

a. BARD as that security of belief appropriate to important decisions in one' s life;
b. Reasonable doubt as a doubt that would make a prudent person hesitate to act;
c. BARD as an abiding conviction of guilt;
d. Reasonable doubt as a doubt for which a reason could be given;
e. BARD as high probability;
f. BARD as a self-evident notion.

This shows that before deciding the case, there is a main type of subjectivism related to BARD when selecting one interpretation or explanation among those available. He states that "beyond reasonable doubt, is abysmally unclear to all those – jurors, judges, and attorneys – whose task is to see that those standards are honoured."[19]

[16] I am sceptical of this. I think that expressing (and carrying out) a value judgment of distribution of errors by means of epistemic criteria of justification is conceptually impossible. This is because the saturation of epistemic criteria of justification itself depends on a context-dependent practical judgment, as already mentioned. If this is so, then, to use epistemic terminology in general rules would be useless. In each single case, the trier of fact needs to evaluate if the epistemic criteria employed by the allegedly SoP is sufficiently fulfill in the light either of the evidence of the case and of the action to be undertaken, and any possible consequences.
[17] Laudan, 2003.
[18] Laudan, 2006.
[19] Laudan, 2006, p. 4

In addition, there is no reason to think that these alternative explanations are exhaustive. Thus,

> [w]hat we face are not different glosses of the same notion but different conceptions of the level of proof necessary to convict someone of a crime. To make matters worse, courts have faulted all these definitions as wrong, misleading, or unintelligible.[20]

However, interpretation is not the only, nor the most important problem. The main problem is not the coexistence of different explanations of BARD in legal criminal practice, nor the coexistence of different conceptions of the level of proof, but rather that they are not genuine conceptions of *proof*, given that the threshold is not defined in terms of epistemic reasons at all, but merely as the particular state of mind the triers of fact must find themselves experiencing. For this reason:

> ... the criteria they offer are purely subjective, they cannot address the question about the rationality of the jurors' confidence or lack of confidence in the guilt of the accused. That discrimination between rational and irrational doubts can never be resolved if one remains focused exclusively at the level of the jurors' degree of conviction.[21]

There could have been *good* interpretations of BARD, making it a real standard of proof, and a genuine epistemic threshold of evidence; but Laudan argues that the U.S. system has been historically out of order in this respect. Every explanation of BARD leads to an uncontrollable subjectivism: Which decisions are *important* in human life? When *should a person hesitate?* How do we determine what makes a person *prudent*? What does it mean to *be certain* in an abiding way?

In light of these questions, one of the above interpretations is particularly relevant to our aim: that which interprets BARD as a high degree of confidence in terms of probability. According to Laudan, this interpretation is much more popular among legal scholars and the public than among judges.[22] Nevertheless, it is crucial to our ends for two reasons: firstly, according to Laudan's lessons, it is an interpretation that leads BARD to an unavoidable subjectivism, or rather, arbitrariness regarding the sufficiency of evidence. On the other hand, Laudan adopts probabilistic understanding of the SoP as an "heuristic hypothesis" to develop a semi-independent line of argument, namely, the

[20] Laudan, 2003, p. 313. Also, cfr. Lillquist, 2002.
[21] Laudan, 2003, p. 320.
[22] Laudan, 2003, p. 310

argument trying to elucidate the actual *ratio* of false positives and false negatives the U.S. criminal system generates by applying BARD. I will now discuss the first point, leaving the second for §3.

Although it is not the only way to understand 'probability', Laudan takes for granted that when this locution is used to explain a SoP, it is understood as a specific degree of confidence where complete subjective certainty is tantamount to the top of the scale: a probability of 1.0.[23] This should be enough to conclude that, if understood as such, BARD – or any other SoP – is not a threshold of genuine epistemic reasons, but rather a distortion of one.

In his first piece on BARD, Laudan only skimmed the surface of the 'subjectivity argument' regarding the probabilistic interpretation of the formula. He stops right after conjecturing that the refusal of this interpretation among judges and the "shabbiness" of the arguments supporting that refusal are the result of their worries about not making "an explicit admission that wrongful convictions will inevitably occur."[24] To recognize that evidence sufficiency could be reached after achieving a degree of confidence less than complete certainty would be to recognize that innocent people can be convicted, and this would delegitimize the whole legal system. Even if complete certainty is impossible in any human affair, which makes errors unavoidable, hypocrisy seems to call for concealment.

The fact that the probabilistic interpretation of BARD leads to uncontrollable subjectivism becomes clear in Laudan's writings a few years later, not when specifically discussing BARD, but when discussing the possibility of fixing a SoP by means of the estimation of probabilities. Laudan rejects this possibility stating that "any probabilistic SoP will suffer from the same problems of subjectivity that (...) proved to be among the principal causes of BARD's undoing."[25]

[23] "When it is said that the degree of the probability relation is the same as the degree of belief which it justifies, it seems to presuppose that both probability relations and degrees of belief can be naturally expressed in terms of numbers, and then that the number expressing or measuring the probability relation is the same as that expressing the appropriate degree of belief" (Ramsey, 1954, pp. 160 [176-177]). Regarding this peculiar understanding of 'probability' see also J. L. Cohen, 1989; de Finetti, 1993 [1973], p. 215; von Wright, 2001 [1951], p. 169.

[24] Laudan, 2003, p. 311. Cfr. Tribe, 1971, pp. 1372-1373.

[25] Laudan, 2006, p. 77. Also see Clermont, 2015; Dahlman, 2017; DeKay, 1996; Dhami, 2008; Kaplow, 2011; Kaplow, 2012; Nance, 2001; Nance, 2016; Nance, 2018; Newman, 2007; Picinali, 2012; Tillers & Green, 1988. This discussion cannot be discussed even minimally here, and it is unnecessary. We can suppose Laudan is right on this point.

Laudan provides two reasons to support this claim. The first is the difficulties in assigning specific probabilities to beliefs where the events at stake are non-statistical such as the events to be decided, not only in judicial processes, but also in other areas of empirical investigation.[26]

> Indeed, if the criminal SoP were defined explicitly in terms of probabilities, I would be obliged to make such a determination. This, in general, is not something that I or most jurors could do with any degree of reliability.[27]

The second, and most important, reason is because a SoP formulated in terms of probabilities, including BARD, is not a standard *of proof* at all. Laudan emphasises that this is at the core of its *confusion thesis*:

> Not to mince words, this is a travesty system of proof. A SoP – in every area in which proof is called for outside the law (including natural science, clinical trials in medicine, mathematics, epidemiological studies, and so on) – is meant to tell the investigator or inquirer when she is entitled to regard something as proved, that is, when the relation of the evidence or the premises to the sought conclusion warrants the acceptance of the conclusion as proven for the purposes at hand. In the criminal law, by contrast, that issue is either wholly ignored or shamelessly finessed. Instead of specifying that the juror's level of confidence in guilt should depend on whether a robust proof has been offered, the criminal law makes the SoP parasitic on the inquirer's (that is, the juror's) level of confidence in the defendant's guilt. We have a proof, says the law, so long as jurors are strongly persuaded of the guilt of the accused (or so long as they will assign it a probability greater than x, where x is the probabilistic SoP). Never mind how they arrived at their high confidence, we have a proof. This gets things precisely backwards.[28]

Under this understanding of BARD, Laudan's argument is sound as there is no guarantee that the degree of confidence felt by triers of fact (or whoever else) covariates with the evidence he or she has at hand. To corroborate this covariation, which is always contingent,[29] it is essential to contrast the available evidence. However, the reverse is not valid: the degree of confidence in the truth of a proposition held by a trier of fact says nothing about the evidence he or she actually possessed. This kind of distortion swaps the threshold of evidence for a state of mind which should have been justified by

[26] See also L. J. Cohen, 1977, p. 74. Cfr. Haack, 2014a, pp. 17-18.
[27] Laudan, 2006, p. 77.
[28] Laudan, 2006, p. 79.
[29] See Burton, 2008.

an independent epistemic body of evidence. As a result, sufficiency of proof turns into an uncontrollable (and probably irrational) state of mind, and is compatible with the absolute absence of any real proof:

> The system offers the juror no neutral or objective standard of proof, saying instead that the intensity of her subjective confidence in guilt determines whether she should convict or acquit the defendant. Making matters worse, the system puts no checks on how the juror goes about arriving at that subjective level of confidence. She is given wholly free rein to make of the evidence what she will and is required at the end only to affirm, if she votes to convict, that she is genuinely persuaded that the accused committed the crime (...). What we face here is not a SoP but a pretext – and a flimsy one at that – for a conviction or for an acquittal.[30]

I am of the opinion that there is a more optimistic way of understanding BARD, to wit: as a methodological requirement for triers of fact. In my view, BARD (and similar formulas) can be understood as a requirement to (a) assess the evidence in an epistemically rational way, and (b) to determine if the evidence so assessed is good enough to convict the defendant in the case at hand. Surely, decisions about the completeness of epistemic justification would now fall within the province of every single trier of fact in each singular case. If one understood practical evaluation as being a necessarily subjective matter, then a subjective decision of the trier of fact would always be carried out by declaring certain propositions as proven.[31] But this would not mean that anything goes when it comes to decisions about the sufficiency of the evidence. At least the assessment of evidence, the first part of the whole affair, would be always rationally controllable according to epistemic criteria. The publicity of trials and maybe also the requirement to explicitly justify judicial decisions would be a way to verify if there actually was evidence, i.e. genuine epistemic reasons; and if inferences were correctly drawn. If this is sound in light of BARD, the concept of proof, i.e. the predicate 'proven' applied to a factual statement, is to be understood as a *thick ethical concept*.[32] To say that something is proven is to say that there is evidence that allows certain epistemic inferences, which is the descriptive side of the affair, and that this evidence is sufficient in order to take responsibility for convicting the defendant in that particular case, given

[30] Laudan, 2006.
[31] For additional nuances of subjectivity regarding BARD, see Lee, 2016; Lee, 2017.
[32] See Williams, 2011, pp. 143-144. See also Blackburn, 1992; Dancy, 1995; Eklund, 2011; Gibbard, 1992; Gibbard & Simon, 1992; Kirchin, 2013; A. W. Moore, 2006; Payne, 2005; Scheffler, 1987; Tappolet, 2004; Väyrynen, 2011; Väyrynen, 2012; Väyrynen, 2013. Indeed, according to what I have said before, I think that this is always the case when someone declares some proposition as *completely* justified, *proven*.

such and such potential consequences should the decision be wrong, which is the evaluative side of the affair. How to conceive the nature of this evaluative side of the predicate (i.e. the one about sufficiency of evidence), most certainly depends on a great deal of further meta-ethical assumptions.[33]

This is not discussed at this point, however. Laudan's *failure thesis* needs to be analysed in order to clearly identify its most problematic components. To do this, we can assume that BARD leads to radical subjectivism. In fact, according to Laudan, BARD is open to a number of different subjective translations. These translations transmute any trace of a genuine SoP into a parody of a system of proof. This particularly happens when there is an attempt to formulate the SoP in terms of probabilities:

> So long as definitions of BARD fail to address the robustness of the evidence and so long as they remain fixated purely on the strength of the juror' s belief, BARD will remain open to the devastating criticism that it confuses mere strength of belief (which may be wildly irrational) with warranted belief.[34]

Let us assume, for the sake of argument, that Laudan's *confusion thesis* regarding BARD is sound.

3 BARD's second conviction: invalid verdicts, the guarantee of false acquittals and the failures thesis

One of Laudan's most persistent (and wholly justified) worries during his years of intellectual work devoted to the epistemology of Criminal Law has been the absence of empirical evidence regarding the real *ratio* of errors the legal criminal system produces when deciding on facts. He points out that no serious empirical research method can be indifferent about its own reliability, or about how often it leads to error. Nevertheless, sadly, the legal system is both simultaneously ignorant and indifferent regarding these matters.[35]

Under these disconcerting circumstances, Laudan devised a way of waking citizens up by making the real rate of error explicit. As a result, he found that this rate deviated

[33] This triggers a large number of issues which cannot be dealt with in this framework. I have elaborated on this in Dei Vecchi, 2014. See also Walen, 2015. The idea of understanding BARD as requiring a method of reasoning has been recently defended by Picinali, 2015. See also Bayón Mohino, 2009; González Lagier, 2018; Haack, 2014b; Lillquist, 2002; Stein, 2005, pp. 117-119.
[34] Laudan, 2003, p. 330.
[35] "This ignorance of, and indifference to, error rates is especially bizarre given that more often hangs on the outcomes of criminal trials than on many other forms of empirical inquiry": Laudan, 2016, p. xiii.

completely in view of the enormous number of false acquittals the system produces. As stated, the Blackstone's *ratio* is hypothetically the underlying reason behind BARD. Laudan founds that the number of false acquittals the system actually produces enormously exceeds even the one the Blackstone's *ratio* would tolerate. This ratification induces him towards a fierce political agenda requiring the revision of several criminal and procedural rules he takes as responsible for those errors. He particularly seeks a decrease in the evidence exigence in order to convict.

In his 2006 book, Laudan took empirical data from a book by Kalven and Zeisel (K-Z)[36], pointing out that their data indicates that BARD generates a great many *invalid verdicts*. For Laudan, an invalid verdict is a significant type of error, and he differentiates this from a false verdict, which can occur in one, or both, of two ways:

> a) The trier of fact may give more or less weight to an item of evidence than it genuinely merits, or b) she may misconceive the height of the standard of proof. In either case, the verdict is inferentially flawed.[37]

Visibly, a verdict can be valid (i.e. be justified inferentially in light of the epistemic threshold) and yet, despite that, be false.

Data taken from K-Z's research with regard to a group of judges, indicate that –in the opinion of that group of judges– juries falsely acquitted defendants approximately 20 percent of the time. Laudan recognizes that this does not mean that 20 percent of acquitted defendants were guilty, but rather, that in the opinion of judges, 20 percent of acquittals were cases of invalid verdicts; cases in which conviction was actually justified according to the SoP; however, juries failed to realize this. In fact, Laudan suggested that: "[w]hile this is only anecdotal evidence, it is impressive for all that, suggesting that (if judges are right) one in five acquittals at trial is an invalid verdict."[38]

More recently, he has been trying to go further to verify the rate of false verdicts, especially the kind of false verdicts particularly despised among jurists and judges, namely, false (even if valid) acquittals. Until this attempt, he had recognized that, "[i]n any event, for now, the frequency of false acquittals is unknown (or better, barely known, since there are plausible conjectures one can make about it)."[39]

[36] Kalven & Zeisel, 1966.
[37] Laudan, 2006, p. 13.
[38] Laudan, 2006, p. 71.
[39] Laudan, 2011, p. 203.

Laudan developed those conjectures in his subsequent work in order to estimate the rate of false acquittals the U.S. legal system generates. To that aim, he developed a combination of conceptual analysis and empirical data analysis departing from a series of statistical research results concerning the violent crimes situation in the U.S. in 2008.[40] As mentioned earlier, despite the broad range of Laudan's work, I am centering on false acquittals *at trials*. In these cases, Laudan holds that the system produces between 75% and 80% of false acquittals within the total number of acquittals at trial. To justify this claim, he takes several argumentative steps.

He starts from a thesis of "conventional wisdom", according to which, the majority of defendants acquitted at trial are "probably factually guilty"[41]:

> ... these defendants wouldn't even be going to trial unless *the prosecutor believed* that he had a *strong chance* of persuading jurors that these defendants were guilty beyond a reasonable doubt.[42]

This conventional wisdom, Laudan says, enjoys a *prima facie* plausibility, even if it does not involve empirical evidence. This plausibility lies in a speculation about the prudential rationality of prosecutors, according to which,

> [e]ven if the prosecutor sometimes overestimates the strength of his case against the defendant, it seems reasonable to suppose that most defendants winning an acquittal at trial have an apparent guilt in the range from about 70% to 90%. In such circumstances, that means, at minimum, either that the prosecutor is an obscenely bad judge of the strength of his case or that the jury concluded that the defendant is probably guilty but that the evidence is too weak to warrant a conviction. One's initial inclination in such circumstances is to suppose that at least half of those who are acquitted at trial actually committed the crime(s) they are charged with but the evidence allowed room for a rational doubt about defendant's guilt.[43]

According to Laudan, this means that if the SoP had been lower than it actually is, the rate of false negatives at trial would have also been cut in proportion. However, this conclusion results from surprisingly equating the *degree of confidence* of the (presumably

[40] See Laudan, 2016, pp. x, 5, 48 , and notes 104-109. Among the statistical data Laudan refers to, the most important to our goals here are: Bureau of Justice Statistics (BJS), *Felony Defendants in Large Urban Counties, 2009* (https://www.bjs.gov/content/pub/pdf/fdluc09.pdf); BJS (2009), *Criminal Victimization, 2008* (https://www.bjs.gov/content/pub/pdf/cv08.pdf).
[41] Laudan, 2016, p. 59.
[42] Laudan, 2016, p. 57; emphasis added. Along the same lines, but long beforehand, see Packer, 1964, pp. 12-13.
[43] Laudan, 2016, p. 58.

prudentially rational) prosecutor (e.g. 70%) with the degree of genuinely epistemic appearance of guilt (which would then be 70%). In other words, Laudan takes for granted, somewhat surprisingly, that the degree of prosecutors' confidence covariates with the evidence they actually had at hand.

Fortunately, the argument does not rest on speculation born from a mix of conventional wisdom and a conjecture as to what degree of confidence would be prudentially rational for a prosecutor to achieve before bringing a case to trial. Indeed, the "hunch" is underpinned by empirical data, or at least Laudan yearns for this. He says that there are "two powerful reasons for thinking that this simplistic assumption understates the frequency of guilt among those acquitted at trial."[44] These reasons are analysed below.

In a first step, Laudan resorts to empirical data from Scotland, where BARD is in force, but there are three possible criminal verdicts rather than just two. After a trial, a defendant in Scotland could be declared "guilty," "not guilty," or "not proven." This last verdict plays a crucial role in Laudan's goal, given that it could be dictated at any time

> (i) the jury is persuaded that the defendant is factually guilty (that is, p(guilt)≥0.5) but (ii) the jury is not convinced of that guilt beyond a reasonable doubt.[45]

As stated in a study of criminal prosecutions carried out by the Scottish government between 2005 and 2006, 71% of defendants tried for homicide and acquitted during that term received a verdict of "guilt not proven".[46] That means, Laudan says, that "about 7-in-10 acquittals for murder in Scotland involved defendants *regarded by the jurors as having probably committed the crime.*"[47] Let us assume, as Laudan does, the implausible expectation that 71% of not-proven cases is a relevant sample, and extensible to criminal cases (for violent crimes) decided in the U.S. during 2008. Or to go even further, let us assume that the sample is relevant with regard to all the previous and later history of any criminal legal system governed by BARD. Notwithstanding these assumptions, before continuing, let me be clear about the reasons why they are implausible. First of all, in Laudan's writings there is neither an argument nor empirical proof to any extent

[44] Laudan, 2016, p. 58.
[45] Laudan, 2016, p. 58.
[46] Scottish Government Statistical Bulletin: Statistical Bulletin Criminal Justice Series: CrJ 2006/3: Criminal Proceedings In Scottish Courts, 2004/5 (https://www.gov.scot/Publications/2006/04/25104019/11).
[47] Laudan, 2016, p. 58; emphasis added.

in favour of the thesis that a "not proven" verdict entails a persuasion of jurors ≥0.5 regarding the factual guiltiness of the defendant. In fact, because of the subjective nature of the enterprise, one could also think that "not proven" is being understood (or perhaps misunderstood) by jurors as being "anything between 0.01 and 0.89" (supposing BARD is equivalent to 0.9). Under this interpretation, "not guilty" would be a situation in which there is evidence in favour of the proposition that the crime has not been committed, or in support of the proposition that it was not the defendant who did it. In other words, a not-guilty verdict would be tantamount to a positive declaration of innocence.[48] Second, using Scottish statistics to interpret U.S. ones is unacceptable without further justification, and Laudan fails to provide any argument in this respect. Further justification should show the absence or irrelevance of any idiosyncratic, cultural, economic, social or other relevant differences between the U.S. and Scotland regarding criminality, the criminal legal system and the way each system is understood from the "internal point of view". Both problems impinge on Laudan's argument. However, I will say no more about them here, mainly because I think they are not the most significant problems when it comes to defending his claim.[49]

In a second step, Laudan again refers to the "monumental study" by K-Z, but now not as an indicator of the rate of *invalid* verdicts, but as empirical evidence of the rate of false acquittals at trial. Part of this study consisted in asking judges about every trial (in total 1,191 cases from the 1960s) which resulted in an acquittal. Following Laudan's reconstruction, the researcher's question was,

> ... whether, in the opinion of the judge, the case was 'close' (meaning the apparent guilt of the acquitted defendant verged on proof beyond a reasonable doubt) or whether it was a 'clear' acquittal (meaning that defendant's apparent guilt was well below the BARD standard).[50]

When interviewed, the judges responded that in their opinion, only 5% of the cases resulted in 'clear' acquittals, while 52% were 'clear for conviction'. The ambiguity of the word 'conviction' is suggestive here: even if Laudan is using it in reference to a guilty verdict, given the nature of the study the word should be understood as a degree of confidence, for what jurors are talking about is their own degree of confidence in the

[48] Nonetheless, fortunately for Laudan's argument, a recent study conducted by the Scottish Government gives *some* plausibility to his conjecture, even if concluding generally in that there are important inconsistencies in a juror's understanding of what the "not proven" verdict means. See Ormston, Chalmers, Leverick, Munro, & Murray, 2019, §§ 5 and 6.
[49] For criticism regarding the way Laudan uses statistics see Gardiner, 2017, §§ 3 and 4.
[50] Laudan, 2016, p. 59.

culpability of (presumably erroneously) acquitted defendants. Let's take, again implausibly, these interviews as a relevant sample extensible to our framework: a total of 45,000 U.S. criminal cases (for violent crimes) decided in 2008. What makes this assumption implausible partially coincides with the issue of inappropriately using Scottish statistics to interpret the U.S. situation mentioned above. Here, too, further justification is needed to show that statistics on the U.S. criminal situation in the 1960s should be read at face value in order to draw inferences which fit with particular crimes and judgments in the U.S. in 2008.

Implausible or not, according to Laudan, this counts as evidence in favour of the conventional thesis, according to which, most defendants acquitted in trial are materially guilty. Moreover, the K-Z data, transposed in time to 2008, leads to the conclusion that 85% of the people acquitted via trial that year could be considered materially guilty, and thus, that those acquittals could be considered *errors*. The argument supporting this conclusion appears thus:

Given that the U.S. statistics from 2008 indicate that about one third of the trials for violent crimes resulted in acquittal, this "would seem to entail [bringing in judicial answers to K-Z study] that only about 15% of the acquittals are 'clearly' acquittals, while some 85% are, in the opinion of the presiding judge, close cases."[51]

This seems to be so, because if 5% of the cases are clear acquittals (according to 1960s K-Z interviews), it is possible to conjecture (Laudan seemingly deems) that only 2,250 violent crime cases decided in 2008 should have been clear cases to acquit (5% of 45,000). Nevertheless, the number of acquittals was 15,000. *Ergo*, Laudan apparently holds, only 15% of acquittals were clearly decided cases. What about the remaining 85%? In virtue of the K-Z study, Laudan asserts that "we can assume" that "more than 12k of defendants are close enough to warrant an assumption that these are probably factually guilty, even if their *apparent guilt* fails to eliminate all reasonable doubts."[52]

The following step in the argument is to combine Scottish data with the K-Z study results in order to get a mid-point and then apply this to the 2008 U.S. statistics:

> ... it seems fair to say that most of those acquitted at trial of a violent crime were nonetheless regarded by the jurors and judges as probably guilty *and thereby* are reasonably assumed to be false negatives. Accordingly, I shall hereafter assume

[51] Laudan, 2016, p. 59.
[52] Laudan, 2016, p. 59; emphasis added.

> that, among those 15k acquittals that emerged in trials for violent crimes in the US in 2008, some 11.2k (75%) were false negatives.[53]

This "fair assumption" governs all of Laudan's quoted work, whereas the "verification" of the high rate of false acquittals the system produces (between 75% and 80% of acquittals after trial), is a banner against BARD's formula *given* its unjustified high exigence *of proof* (among other rules which surely benefit defendants).

Let us analyse Laudan's argument in depth.

The first thing to consider is that his argument takes a triple presupposition as being paramount, i.e., that there *is* a SoP in force, that this SoP is BARD, and that BARD equals a confidence level of 90% or greater.[54] Taking the first step from this triple presupposition, the argument could be reconstructed as follows:

(a) Most fact finders in the U.S. criminal legal system – as in the Scottish system – who must decide whether to convict defendants (for violent crimes) believe that the SoP they need to apply (i.e. BARD) demands a degree of confidence in guilt tantamount to a likelihood of 90% (or above).

[53] Laudan, 2016, p. 59; emphasis added. He says later that "of 80% (12k out of 15k acquittals) among those who were acquitted at trial" (Laudan, 2016, p. 65). The false acquittals rate is to be completed with the rate of wrongly dropped charges. This is not part of the framework of this paper. Nevertheless, our arguments can be easily applied to Laudan's analyses of this phenomena, at least in part. Indeed, he calculates the rate on the basis of interviews which ask prosecutors about their degree of confidence regarding defendants' guilt in cases where charges were dropped. Laudan believes that these interviews lead to the conclusion that 56% of cases dismissed and dropped (for reasons that had nothing to do with evidence) were probably factually guilty, on the basis of the subjective beliefs of officials. He concludes that "we can reasonably suppose that the proportion of guilty among them would be about the same as the proportion of guilty among those who go to trial." This is because, as we have seen, 2/3 of defendants that went to trial in 2008 had been convicted BARD and 75% or 80% of the remaining 1/3 were falsely acquitted. Thus "[t]hat seems to provide a plausible rationale for saying that, among those defendants who had the charges against them dropped for *non-evidentiary* reasons, approximately two-out-of-three (and probably more) are highly likely to be guilty. Hence, we shall assume that about 37% to 38% (that is two-thirds of the 56% of those whom were booted out of the trial system for non-evidentiary reasons) are factually guilty (and, if they had gone to trial, would have been convicted)" (Laudan, 2016, p. 64).

[54] Again, this seems to be supported empirically: a previous study asked judges and jurors in England about the level of probability they consider BARD requires for conviction in a criminal trial. According to (Laudan's presentation of) those interviews, 1/3 of judges put the level between 70% and 90%, and almost all the rest above 90%. Among jurors, 26% declare a willingness to convict on confidence below 70%, and 54% declare that the level of confidence required by BARD was 90% or more. See Laudan, 2006, p. 47, quoting J. Cohen & Christensen, 1970. This assumption cannot be discussed here, mostly because my argument would work even if the assumption were right. But in any case, it is worth mentioning that there are very important reasons to doubt its correctness. See, for example, Lillquist, 2002, p. 112 (analyzing diverse empirical studies whose results are strikingly different), Hastie, 1993 (where studies showing that popular understanding of BARD varies between 51% and 92%), and the recent analysis of Walen, 2015, § 6 (distinguishing between the "customary" and the "practice" understanding of BARD; and arguing (on the base of several empirical studies) that "a substantial number of jurors interpreted the BARD instruction [the practice understanding] to allow conviction on weak evidence, evidence that would not even satisfy the preponderance of the evidence standard for over half of the juries that considered it": Walen, 2015, p. 375).

(b) Thus, the SoP in force demands a 90% probability of guilt to convict.

(c) Combining K-Z's 1960s interviews and the Scottish rate of not-proven verdicts from 2004/5 shows that between the 75% and 80% of trial acquittals, fact finders had a level of confidence higher than 50%, but lower than 90% of probability in the (finally) acquitted defendants being guilty.

(d) *Ergo*, between 75% and 80% of defendants have been acquitted *in spite of the appearance of guilt being higher than 50%* – the percentage that makes them *probably guilty defendants*.

(e) From there it follows (Laudan should be thinking) that 75% of acquittals in favour of defendants whose appearance of guilt is higher than 50%, benefits materially guilty defendants by letting them get away with the crime. This is so, given that that person is *probably* guilty (over 50% likely), even if not *provably* guilty in light of BARD.[55]

(f) The conclusion is that 75% of acquittals at trial are false acquittals and that this is, at least to a certain extent, due to BARD's excessively demanding evidential requirement.

It is highly surprising that after cautioning us on the uncontrollable subjectivity of measuring degrees of confidence in terms of probabilities, Laudan supports an entire argument for a factual hypothesis almost exclusively pointing to judges' or jurors' speculations or self-attributed degrees of confidence. For example, opinions on the specific degree of confidence BARD demands; speculation regarding whether degrees of confidence prosecutors needed in order to take a case to trial were (prudentially) reasonable; judges or jurors self-attributing degrees of confidence regarding the probability of people acquitted at trial being guilty; judges' or jurors' self-attributing degrees of confidence by taking into account the *closeness* (the degree of confidence they thought was demanded) to BARD in cases where they decided to acquit at trial; self-attributing degrees of confidence expressed by not-proven verdicts in Scotland; and so on.

For present purposes, the most important point is (d), which states that BARD is responsible for 75% (or 80%) of false acquittals due to its excessive demand for *evidence*: "... even when the evidence is not skewed or unrepresentative of the crime, there is still

[55] See L. J. Cohen, 1977.

plenty of scope for a verdict that is valid but not true. Indeed, the standard of proof guarantees as much."[56]

The problem is no longer the *validity* of the verdict, i.e. the danger is no longer that the triers of fact can make a mistake in understanding or applying the SoP. On the contrary, it is the objective evidential height of the standard which guarantees false negatives when rightly applied. After explaining invalid verdicts in 2006, Laudan stated:

> It is essential to understand what is meant by "false acquittals" and "false convictions" in the study in question. What Kalven and Zeisel found was not that two of every ten acquittals were someone who was guilty but that two of every ten acquittals, in the opinion of the judges, justified a conviction, that is, satisfied the criminal SoP. Since that standard is quite high, we have every reason to expect that false acquittals, as we are using that term – namely, the acquittal of truly guilty defendants – happen far more often than the 20 percent figure would imply.[57]

Thus, even if a verdict is valid, and considering the enormous number of false acquittals that the argument of precedent 'shows', we can now say that:

> ... the overwhelming reason for most false negatives is not a flawed case made by the prosecutor nor a lack of strong inculpatory evidence but the fact that, to convict someone of a crime, our system generally requires that the jury must be *unanimously persuaded of his guilt to a degree of near certainty* (proof beyond a reasonable doubt) [...] Because the evidence profiles of many truly guilty defendants fail to meet the exacting threshold of proof beyond a reasonable doubt, such persons will receive a verdict of not-guilty (better expressed as 'guilt not proven') even when most or even all of the jurors strongly believe them to be the culprit.[58]

> Viewed from this perspective, it is vividly clear that the current standard of proof is too high. As I have *already shown*, the promiscuous acquittal of, or dropping of charges against, many very probably guilty defendants (because their apparent guilt is perceived as below the BARD standard) is generating much more harm than the occasional conviction of innocent defendants does.[59]

The alleged verification of this flaw and of the number of errors it produces, i.e. the *failure thesis*, constitutes an empirical basis for Laudan to encourage a reduction in the

[56] Laudan, 2006, p. 13.
[57] Laudan, 2006, p. 70, note 14.
[58] Laudan, 2016, pp. 13-14; emphasis added.
[59] Laudan, 2016, p. 98.

epistemic exigence of the SoP, and an elimination or reduction of procedural privileges granted to defendants.[60] Indeed, given that BARD leads to an unbalanced number of false acquittals, a lower threshold of proof would put us on the right track for reducing errors globally, especially, false negatives. Again, according to Laudan, BARD has caused and continues to cause many more false acquittals than Blackstone's *ratio* would be willing to tolerate (i.e. BARD does not fit its underlying justification), and many more than he (according to its *consequentialist thesis*) takes to be morally decent.[61]

However, as I will try to show in the next section, the *failure thesis* is either clearly inconsistent with the first conviction to BARD developed in §2, i.e. with the *confusion thesis*, or, in any case, independently unacceptable. Note that even though Laudan circumscribes his *failure thesis* to only a section of the U.S. criminal system, the *kind* of argument he develops, and that I have reconstructed in this section, has a much greater reach in that it could be reproduced for any other legal system, or section of a legal system, by simply appealing to similar premises. Therefore, by showing that the unacceptability of the claim inheres the *kind* of argument the *failure thesis* is an example of, I will undermine not just its Laudan-US-violent-crimes version, but any other attempt to ground a claim of "too many false acquittals" adducing these sort of premises.

[60] "The object lesson seems to be clear: we are now utilizing a standard of proof much more demanding than it should be, since it leads to far greater risk of harm in aggregate than many lesser standards would. BARD (understood as a surrogate for about 90% confidence) would make sense as a standard of proof for violent crime just in case the harm done by a false conviction was about ten times greater than the harm done by a false acquittal. There is no way that such an assessment of harms can be plausible...": Laudan, 2016, p. 97.

[61] The problem of the *consequentialist thesis* opens an independent line of thought in Laudan's worries. He displays these worries as a direct attack on every deontologist intent on fixing a SoP, and mainly going against Blackstone's *ratio*, as I have pointed out at the beginning of this paper. Conversely, assuming a metaethical utilitarian point of view and developing a complex calculus of (his own) undesirability of false negatives and false positives, Laudan tries to justify a brand-new fair *ratio*. For him, also, convicting an innocent person is worse than acquitting a guilty defendant, but the fair *ratio* would be of 1 false positive for every (less than) 2 false negatives. See Laudan, 2011; Laudan, 2016, ch. 4. I think this laudanian line of argument should also be carefully scrutinized, but this is not the moment to do it. Along these lines, Gardiner, 2017, analyses this a very seriously, and by a different path arrives at the same conclusion I am defending: "Laudan's consequence-based argument for reducing the standard is unsuccessful. Given doubts about statistics employed, the paucity of kinds of consequences considered, and the mistaken formula used to derive the new standard, we should conclude Laudan's argument fails. A consequence-informed approach to crime reduction is important, but Laudan's evaluation of the consequences does not establish the standard of proof is too high. Perhaps this is good news for advocates of consequence-based justifications of the standard of proof: if consequence-based calculations recommend standards of proof as low as 56 to 67%, this casts doubt on the consequence-based project. The standard is implausibly low and conflicts with our understanding of criminal justice. And if, as Laudan claims, an appraisal of the consequences suggests that Blackstone's 10:1 ratio of false acquittals to false convictions should be jettisoned in favour of something closer to 2:1, this suggests that consequences cannot be a good guide to criminal justice. Fortunately for advocates of the consequence-based approach, it appears that Laudan's estimations of the consequences are not close to accurate" (Gardiner, 2017, p. 236). Regarding justification of BARD see also Picinali, 2018.

4 Imperceptible crimes[62]

Laudan's *failure thesis* rejection requires two steps. First, I will argue that arriving at the conclusion that BARD is the main cause of the false negatives produced by the system *because* of its high demand for evidence, essentially depends on the previous ascertainment of the precise *epistemic exigence* of the formula. Exactly the same thing should be said for the claim that some verdicts are invalid when BARD is applied wrongly. Nonetheless, Laudan's *confusion thesis* supposes that this ascertainment is hopeless. This is the core of the inconsistency between Laudan's theses, as previously stated. As I have already pointed out, this apparent inconsistency is relevant as it is a touchstone from where it is possible to see the requirements the *failure thesis* should fulfil in order to be suitable. In the second step I will argue that Laudan's thesis, which states that BARD produces too many false acquittals is in part a case of the wrong-sort-of-reasons fallacy, and is groundless.

4.1 An inexistent threshold cannot be too high: an inconsistency?

One way to determine if a SoP in force is epistemically too demanding would be to pre-identify the exact threshold of evidence the standard requires. However, according to Laudan's *confusion thesis*, this is not possible when the SoP in force is BARD. On one hand, it is an extremely ambiguous formula. On the other, and more relevant to our purpose, the usual understandings of the formula – as Laudan evinces – turned it into a travesty system of proof, a fake epistemic-criterion. This is what would happen, especially with attempts to articulate the standard in probabilistic terms. According to Laudan, these attempts are an illusion, not only due to the impossibility of quantifying degrees of confidence in contexts such as the legal process, but because the result of doing so is "fundamentally subjective"; "[a]nd nothing about it is more peculiar than the fact that it offers no free-standing standard *of proof*."[63]

What this has taught us is that any degree of confidence a judge or juror could express regarding the guilt of a defendant is compatible with a large amount of evidence, with a medium amount of evidence, and even with no evidence at all. In other words, no

[62] I'm evoking the original title of Gillermo Martinez's novel *The Oxford Murders*, MacAdam/Cage, 2005.
[63] Laudan, 2006, p. 79; emphasis added.

covariation between degree of confidence and degree of epistemic justification is guaranteed.[64]

> After all, the credulous among us are not likely to harbour doubts, even when they should and the skeptics – especially if they are philosophers – are likely to have lingering doubts no matter how powerful the proof. By using a standard as amorphous as BARD, the state is essentially inviting triers of facts to trust their gut instincts [...] In a society whose citizens are variously sure that the Bible is divinely inspired and inerrant, that their children are uniformly smarter than their teachers think, and that we will all be reincarnated in another body after death, the existence of a high level of subjective confidence about x among twelve ordinary citizens does not inspire much confidence that x is either true or that x has been proved in any even mildly robust sense of the term.[65]

Therefore, the opinion of judges, jurors, or citizens in general about the degree of confidence the hypothesis of guilt ought to achieve to convict according to BARD says absolutely nothing about the epistemic exigence it entails. As Laudan puts it in referring to BARD: "The standard of proof (...) in criminal trials in the United States is, I have just claimed, a mess. It is not only ill defined, but it smacks of the arbitrary."[66]

It is worth remembering that the whole argument in favour of the *failure thesis* had taken its starting point from the presupposition that what BARD requires is just a high *degree of confidence* in the guilt of the defendant to convict him or her (between 90 and 95%). This means that the standard Laudan presupposes as the one in force in his view, is one that improperly reduces the threshold of evidence to a subjective degree of confidence, excluding any reference to epistemic reasons. But it is precisely because of this assumption that no conclusion about epistemic exigence of BARD could be drawn from citizens', jurors' or juries' opinions regarding the level of confidence the formula requires.

Therefore, even if the rate of false acquittals at trial was what Laudan purports, to ascribe those errors to the *demand of evidence* of BARD just because some people say that the formula requires a high degree of confidence is gratuitous. The rate of false acquittals should be *proved* anyway, as they could be due to a number of other causes. Without a "free-standing standard *of proof*" we have no parameter to say whether the evidence

[64] See Corea-Levy, 2012, § IV. I am not denying that a degree of confidence in truth of a certain proposition *could* be epistemically rational. What I am saying, in Laudan's vein, is that the mere confidence says *per se* nothing about its own epistemic grounds.
[65] Laudan, 2011, p. 215.
[66] Laudan, 2006, p. 64.

demand is satisfactory or not. And as Laudan revealed, resorting to BARD we have no such free-standing SoP but one "... grievously inadequate, deliberately unclear, wholly subjective, and open to about as many interpretations as there are judges, to whom it falls to explain this notion to hapless jurors."[67]

Given that this is the notion guiding (or failing to guide) every decision made in U.S. courts during 2008 and earlier, as well as in Scotland and many other countries,[68] Laudan's conclusions regarding the demands made by the threshold *of proof* and the errors caused by *that* threshold seem to become idle. No conclusion could be drawn regarding the adequacy of the evidentiary threshold of BARD, nor regarding the validity of its applications to certain cases facing disagreement between jurors or judges.

However, there is a way around this obstacle. It is possible to avoid inconsistency in defending the *confusion thesis* and the *failure thesis* at the same time. After all, one could prove that even if BARD leads to the *land of confusion*, it happens to be that in a particular legal system, in a certain community, the triers of fact only convict in cases where there is a huge amount of evidence, and in many other cases, acquitting the defendant when the evidence of guilt was quite good. In these circumstances, one could try to argue, as Laudan does, that a conviction would have been epistemically rational and morally convenient. This could even lead to claiming that in spite of its abstract indeterminacy, the way in which BARD is applied in the relevant community happens to be excessively exigent regarding evidence. However, in order to escape inconsistency, one should properly ground this claim, resorting to the right sort of reasons; simply invoking the high degree of confidence people declare BARD requires is not sufficient. This is precisely here where Laudan's error becomes serious.

4.2 Appearance of guilt that smacks of the arbitrary

This brings us to the main point of this paper. To avoid the inconsistency described above, Laudan was able to reformulate his *failure thesis* in a way that would make it worthy of consideration, to wit: as an empirical thesis independent from the degree of certainty BARD requires. According to this reformulation, the claim would be that, notwithstanding the *confusion thesis*, in the actual practice of adjudication, jurors generally applied the formula in a way that leads them to acquit in too many cases despite the presence of significant amounts of evidence of guilt. To a certain degree, we

[67] Laudan, 2006, p. 30.
[68] See Gardiner, 2017; Mulrine, 1997; Picinali, 2009; Walen, 2015.

can say this is what Laudan *tries* to demonstrate, but fails to, given the unappropriated grounds of his argument.

The problem is no longer that Laudan assumes BARD is too demanding, given the opinions of jurors and juries regarding the degree of confidence BARD requires. The problem is now quite different and concerns Laudan's strategy to prove that large amounts of evidence of guilt were present in the majority of the cases acquitted. As found in Laudan's contentious interpretation of the Scottish study, the judges manifested their *degree of confidence* in guilt tantamount or superior to 50% by means of guilt-not-proven verdicts. As shown by the K-Z interviews, judges and jurors manifested their *degree of confidence* that only 5% of all 1,192 cases of the acquittals sample were clear acquittals. From this, Laudan infers that in all the remaining acquittals, the degree of confidence should have been tantamount or superior to 50%.[69]

This raises at least two problems. The first regards the miss-invocation of reasons of the wrong sort; the second concerns the absence of reasons of the right sort.

4.2.1 A CASE OF THE WRONG-SORT-OF-REASONS ARGUMENT

The first problem concerns the reasons Laudan invokes to infer the rate of false acquittals he claims the system generates. Indeed, he adopts an entirely subjective SoP formulated in terms of probabilities, and by doing so, he falls into the very same flaw he had been criticizing: the assumption that the jurors' alleged degree of confidence verged towards the one BARD supposedly requires (i.e. bordering on 90%) in cases they had acquitted, and that this was equivalent to the epistemic appearance of guilt. In other words, he falls into the *confusion error*.

This is, in fact, the result of a double confusion. First, Laudan confuses what Richard Moran would call "theoretical" and "deliberative" answers regarding states of mind.[70] Laudan takes for granted that the information he uses, i.e. jurors' or judges' self-attributions of the degree of confidence about defendant's guilt *from the theoretical point of view*, is tantamount to the degree of confidence which would have been right for them to achieve *from the deliberative point of view*. However, the self-attributions of (past) degrees of confidence regarding certain propositions does not demonstrate the

[69] While working on this paper, Findley's 2018 article came into my hands. Some of his criticisms of Laudan are along the same lines as mine. See especially Findley, 2018, pp. 1278-1280 for an in-depth analysis of K-Z's study.
[70] See Moran, 2001, §§ 2.5 and 2.6.

reasonableness of the doxastic attitude at stake.[71] Second, and even worse, Laudan takes for granted that both answers reflect the objective degree of epistemic justification of the proposition at stake: to wit, the every defendant's guilt. Laudan assumes, with no apparent reason at all, that subjective probability expressed by means of degrees of confidence (close to BARD, and therefore close to 90%) *corresponds* in a perfectly symmetrical way to the degree of epistemic justification. From this, he concludes that there actually was an *epistemic* probability higher than 50% to support a conviction verdict in every case in the sample.[72]

Nevertheless, degrees of confidence about a proposition *p* which are self-attributed are not *epistemic reasons* to accept that proposition as true. They are the wrong sort of reasons to justify accepting *p*. We should recall here, paradoxically, Laudan's lesson about assigning probability degrees to beliefs. He asserts this is an arbitrary task, if it is even possible at all; and something that "neither he nor most jurors could do with any degree of reliability." As a result, if assigning probabilistic degrees of confidence can be irrational when a probabilistic interpretation of BARD is at stake, the same has to be said for assigning probabilistic degrees of confidence which are allegedly *verged on* BARD: this also remains open to the allegedly "devastating criticism that it confuses mere strength of belief (which may be wildly irrational) with warranted belief." Rephrasing Laudan's previously quoted words regardless of how they arrived at their close-to-BARD confidence, we have a close-to-BARD-proof. This also reverses the situation, inasmuch as closeness-to-BARD fails to address the robustness of the evidence. As long as they remain fixed purely on the strength of the juror's belief,

[71] Let alone the problem of the accuracy of the opinions of jurors or juries regarding their own task and of self-attributions of degrees of confidence. They could not only have been lying when answering questions in interviews, but could also have been mistaken about their own performances. See D. A. Moore & Healy, 2008. Laudan overlooks any problem of this kind. Thanks to Hernán Bouvier for pressing me to clarify this point.

[72] Even in his 2006 book, after scrutinizing the opinion of "one distinguished member of the fraternity of defense attorneys" according to which almost all criminal defendants are, in fact, guilty and all criminal defense lawyers, prosecutors and judges share this understanding, Laudan remarked that "[s]till, it would be preferable to have something better than anecdotal evidence to go on. Where direct evidence is concerned, we obviously have very little since true guilt and true innocence are largely inscrutable in the vast majority of cases. But there are some plausible inferences we can draw from that data we have at hand" (Laudan, 2006, p. 108). After taking the midpoint of 35 percent of acquittal rate at criminal trials across the U.S., he continued: "The first and most obvious (but not necessarily the most telling) inference is that, in the opinion of the jurors (and judges), most defendants are guilty. Indeed, their apparent guilt is so striking that a clear majority have been found to be guilty BARD, and there were probably numerous others whose apparent guilt was well above 0.5 and who, therefore, may well have been guilty" (Laudan, 2006, p. 108). Now, even if the first part of the inference is indeed obvious (i.e. that *in the opinion of the jurors and judges* most defendants going to trial were guilty), the second is neither obvious nor plausible. For there is no good reason to conclude that the opinion of jurors and judges reflects an epistemic appearance of guilt, to a striking measure, or to any other.

closeness-to-BARD will remain open to the devastating criticism that it confuses mere strength of belief (which may be wildly irrational) with warranted belief.[73] If BARD smacks of the arbitrary, so too does closeness-to-BARD. This brings us to the second type of problem.

4.2.2 THE COMPLETE ABSENCE OF THE RIGHT SORT OF REASONS

To defend the *failure thesis* without falling into the "wrong-sort-of-reasons" fallacy, the claim that in most of the acquitted cases the defendant was actually guilt should be justified. To justify this claim, a necessary (yet insufficient) condition must exist: to wit, bringing the evidence underpinning the defendant's guilt to the fore. In light of this condition, to invoke triers of fact's high degree of confidence in a defendant's guilt in acquitted cases, as Laudan does, would be relevant if, and only if, one can prove that the relevant degree of confidence of the trier of fact is tantamount to the degree of epistemic justification. Yet, it seems that the only way of proving this point is by means of the body of evidence grounding a hypothesis of guilt. Then, the only acceptable way to estimate the degree of epistemic justification grounding hypotheses of guilt in acquitted cases would be, ultimately, contrasting the body of evidence evaluated by the trier of fact in every acquitted case. In other words, the only appropriate reason to estimate this is the evidence jurors had when they decided to acquit.

We can now see why Laudan's *failure thesis* turns out to be unacceptable: the only proper grounds to claim that a certain number of acquitted defendants were probably epistemically guilty (yet not probably enough to satisfy the SoP in the way jurors or judges presumptively understood it) would be the real evidence against each one of those defendants; the only reasons *making guiltiness provable*. Yet, Laudan fails to provide any evidence to encourage us to believe that not even one of the cases of acquittal he takes under consideration had good epistemic reasons to believe in guilt.

This should be enough to reject not only his *failure thesis*, but also his speculation (perhaps more a hunch) of the high rate of false acquittals at trial, and his political program of decreasing evidence exigency to convict people. He fails to fulfil a necessary condition of the argument, namely, to properly justify that there was evidence in favour of a guilt hypothesis in cases of acquittal.

[73] As Laudan said when speaking about moral philosophers, who in his view endanger the discussion when they "suppose themselves competent to hold forth on questions involving the rigor or robustness of methods of proofs" - now is Laudan himself who "stand[s] hoist on their own petard": Laudan, 2011, p. 218.

But there is an additional reason for this rejection. Let us suppose that BARD is in force, and that it actually requires 90% confidence in a defendant's guilt in order to secure a conviction. Let us also imagine a community that is infallibly rational from the epistemic point of view; one in which the degree of confidence in factual propositions covariates with the degree of epistemic justification (which is, by hypothesis, also quantifiable). Let us even suppose that for every case decided in that community, one can contrast all the evidence fact finders had in order to make a decision, and every epistemic inference they drew. Now, let us assume that of the total amount of acquittals after trial in criminal cases in that community there is a very high rate (let us say, 85%) favouring defendants whose guilt was proven between, say 60% and 89%.

Even under these circumstances, to claim that the rate of *false acquittals* in this epistemically ideal community is 85% would be unacceptable, to say the least. Since this claim would be the outcome of simply and dubiously applying a less demanding SoP than the one in force. For this claim to be minimally acceptable as *proving a factual error*, new evidence should be brought to the fore (maybe even evidence excluded from the procedure due to some exclusionary rule).

By contrast, one could not legitimately claim that because the SoP in force demands x% of epistemic justification in order to secure a conviction (x being a number higher than 51), and every not-proven decision where proof of guilt was between 50% and x_{-1}% constitutes a *false negative*. This would be mere trickery. Without additional evidence, one could at most argue that to acquit someone whose guilt had been proven between 50% and x_{-1}% was morally or politically undesirable. However, we must be aware that this kind of disagreement is a moral or political one, not a factual, nor an epistemic one. This is because in order to justify the claim that the defendant's guilt was 'sufficiently proven' in most of the acquitted cases requires more than simply providing the evidence to prove the defendant's guilt (the necessary affirmation). In addition, the claim entails an evaluative judgment on the desirability or convenience of a conviction. Then, even if Laudan's argument on the rate of acquittals with *probable* guilt (i.e. above 50%) were sound (and we have seen it is not), his conclusion that this is an indicator of acquittals *of guilty people* should be taken as no more than a mere evaluative judgment where a certain degree of probability of guilt, which less than that demanded by the SoP supposedly in force, *should be* enough to convict.[74]

[74] This again leads us to Laudan's *consequentialist* calculous. See n. 60.

5 Conclusion

The two worries central to Larry Laudan's contributions to legal-evidential reasoning are the radical arbitrariness of the SoP in force in the U.S. to decide criminal cases, i.e. BARD; and the unreasonable number of false negatives given the exacting requirements of the standard.

As I have tried to demonstrate, there is something wrong in the way he sets these worries out: to insist the standard is subjective and arbitrary, to label it a travesty and a fake system of proof seems inconsistent with saying that the standard is both excessively high and that it is responsible for a certain number of errors. Furthermore, his argument in favour of the claim that the system produces too many false acquittals is unacceptable, and not just because of this inconsistency. The main reasons are as follows: on the one hand, Laudan himself confuses (presumptive) strength of belief of jurors or judges with warranted belief; and on the other hand, he fails to provide any appropriate reason to show that the defendants acquitted were materially guilty.

If Professor Laudan tries to persuade his audience that the epistemic demand of the SoP needs to be diminished and they react in an epistemically rational way, then the demand should be rejected. This should be the case at least until he adduces (i) evidence proving the relevant epistemic probability of guilt of (a relevant number of) acquitted defendants and (ii) evidence to the effect that those acquittals benefitting people who are probably guilty are because a SoP with an excessively high threshold of evidence is applied. Trying to support the demand presuming that the SoP requires 90% or 95% confidence is, as we have seen, futile. Indeed, the mere fact that the whole of Laudan's argument on the rate of false acquittals hinges on interpreting BARD in terms of probabilities is dubious. After all, it is Laudan himself who clearly states that translating a SoP into probabilities leads to subjectivity. But in his 2006 book he clearly made the point:

> My reasons for doing so [for making use of the language of probability] were chiefly heuristic; we can learn a great deal about how a SoP behaves, and especially about its role in the distribution of errors, by drawing on well-established results driving from the analysis of probability and statistics. But like Ludwig Wittgenstein's famous ladder that we use for climbing a wall and

> then – having reached the top – discard, the technical discourse of probabilities is now best laid aside as a tool for formulating a standard of proof.[75]

The problem is that inferring that the criminal system produces a definite number of false acquittals (or any errors whatsoever) and that BARD is (partly) the cause of that by starting from a probabilistic interpretation of that SoP, goes far beyond a *chiefly heuristic use* of the probabilistic model of discourse. There is nothing wrong with using a ladder to reach the top of the wall and then discarding it. The problem is that Laudan's eagerness to disclose the rate of false acquittals at trial seems to have made him forget that after reaching the top he had to *land on the wall* before discarding the ladder.

[75] Laudan, 2006, p. 77.

Bibliography

Alston, W. P. (1989 [1985]). Concepts of Epistemic Justification. In *Epistemic justification. Essays in the theory of knowledge* (pp. 81-115). Ithaca: Cornell University Press.

Bayón Mohino, J. C. (2009). Epistemología, moral y prueba de los hechos: hacia un enfoque no benthamiano. *Analisi e diritto, 2008*, 15-34

Blackburn, S. (1992). Through Thick and Thin. *Proceedings of the Aristotelian Society, Supplementary Volumes, 66*, 284-299. doi: 10.2307/4106979.

Burton, R. A. (2008). *On Being Certain: Believing You Are Right Even When You're Not*. New York: St. Martin's Press.

Clermont, K. M. (2015). Trial by Traditional Probability, Relative Plausibility, or Belief Function? *Case Western Reserve Law Review, 66*(2), 353-391

Cohen, J., & Christensen, I. (1970). *Information and choice* Edinburgh: Oliver & Boyd.

Cohen, J. L. (1989). *An Introduction to the Philosophy of Induction and Probability*. Oxford: Clarendon Press.

Cohen, L. J. (1977). *The Probable and the Provable*. Oxford: Clarendon Press.

Corea-Levy, Y. (2012). Making Sense Of Reasonable Doubt: Understanding Certainty, Doubt, And Rule-Based Bias Filtering. *American University Criminal Law, 8*(1), 48-62

Dahlman, C. (2017). Determining the Base rate for Guilt. *Law, Probability and Risk, 17*(1), 15-28. doi: 10.1093/lpr/mgx009.

Dancy, J. (1995). In Defense of Thick Concepts. *Midwest Studies In Philosophy, 20*(1), 263-279. doi: 10.1111/j.1475-4975.1995.tb00316.x.

Darwall, S. (2006). *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge - Massachusetts - London: Harvard University Press.

Darwall, S. (2013). *Morality, Authority, & Law: Essays in Second-Personal Ethics* (1 ed. Vol. I): Oxford University Press.

de Finetti, B. (1993 [1973]). Bayesianesimo: il suo ruolo unificante per i fondamenti e le applicazioni della statistica. In P. Monari & D. Cocchi (Eds.), *Probabilità e induzione* (pp. 205-228). Bologna: Cooperativa Libraria Universitaria Editrice Bologna.

Dei Vecchi, D. (2014). Acerca de la fuerza de los enunciados probatorios: El salto constitutivo. *Doxa: Cuadernos de filosofía del derecho, 34*, 237-261

DeKay, M. L. (1996). The Difference between Blackstone-Like Error Ratios and Probabilistic Standards of Proof. *Law & Social Inquiry, 21*(1), 95-132. doi: 10.1111/j.1747-4469.1996.tb00013.x.

Dhami, M. K. (2008). On measuring quantitative interpretations of reasonable doubt. *Journal of Experimental Psychology: Applied, 14*(4), 353-363. doi: doi: 10.1037/a0013344.

Eklund, M. (2011). What are Thick Concepts? *Canadian Journal of Philosophy, 41*(1), 25-49

Ferrer Beltrán, J. (2018). Prolegómenos para una teoría sobre los estándares de prueba. El test case de la responsabilidad del Estado por prisión preventiva errónea. In D. M. Papayannis & E. Pereira Fredes (Eds.), *Filosofía del derecho privado (en prensa)* (pp. 401-430). Madrid-Barcelona-Sao Paolo-Buenos Aires: Marcial Pons.

Findley, K. A. (2018). Reducing Error in the Criminal Justice System. *Seton Hall Law Review, 48*(4), 1265-1318

Forst, B. (2004). *Errors of Justice. Nature, Sources, and Remedies*. Cambridge-New York: Cambridge University Press.

Gardiner, G. (2017). In Defence of Reasonable Doubt. *Journal of Applied Philosophy, 34*(2), 221-241. doi: 10.1111/japp.12173.

Gibbard, A. (1992). Thick Concepts and Warrant for Feelings. *Proceedings of the Aristotelian Society, Supplementary Volumes, 66*, 267-283. doi: 10.2307/4106979.

Gibbard, A., & Simon, B. (1992). Morality and Thick Concepts. *Proceedings of the Aristotelian Society, Supplementary Volumes, 66*, 267-299. doi: 10.2307/4106979.

González Lagier, D. (2018). *Prueba y argumentación ¿Es posible formular un estándar de prueba preciso y objetivo? Algunas dudas desde un enfoque argumentativo de la prueba*. Paper presented at the Congreso Mundial de Razonamiento Probatorio, Girona.

Haack, S. (2009 [1993]). *Evidence and Inquiry. A Pragmatist Reconstruction of Epistemology* (2nd, expanded ed ed.). Amherst, N.Y.: Prometheus Books.

Haack, S. (2010). Belief in Naturalism: An Epistemologist's Philosophy of Mind. *Logos and Episteme, 1*(1), 67-83

Haack, S. (2014a). Epistemology and the Law of Evidence. In S. Haack (Ed.), *Evidence Matters: Science, Proof, and Truth in the Law* (pp. 1-26). New York, NY: Cambridge University Press.

Haack, S. (2014b). Legal Probabilism: An Epistemological Dissent. In S. Haack (Ed.), *Evidence Matters: Science, Proof, and Truth in the Law* (pp. 47-77). New York, NY: Cambridge University Press.

Harman, G. H. (1965). The inference to the Best Explanation. *The Philosophical Review, 74*, 88-95

Hastie, R. (1993). Algebraic Models of Juror Decision Processes. In *Inside the Juror: The Psychology of Juror Decision Making* (pp. 84-115). Cambridge-New York: Cambridge University Press.

Kalven, H., & Zeisel, H. (1966). *The American jury*. Boston: Little.

Kaplow, L. (2011). On the Optimal Burden of Proof. *Journal of Political Economy, 119*(6), 1104-1140. doi: 10.1086/664614.

Kaplow, L. (2012). Burden of Proof. *The Yale Law Journal, 121*(4), 738-859

Kirchin, S. (2013). *Thick Concepts*. Oxford: Oxford University Press.

Laudan, L. (2003). Is reasonable doubt reasonable? *Legal Theory, 9*(4), 295-331

Laudan, L. (2006). *Truth, Error, and Criminal Law: An Essay in Legal Epistemilogy* (1st ed.). Cambridge: Cambridge University Press.

Laudan, L. (2007). Strange Bedfellows. Inference to the Best Explanation and the Criminal Standard of Proof. *International Journal of Evidence & Proof, 11*(4), 292

Laudan, L. (2011). The Rules of Trial, Political Morality, and the Costs of Error: Or, Is proof Beyond a Reasonable Doubt Doing More Harm than Good? In L. Green & B. Leiter (Eds.), *Oxford studies in philosophy of law* (Vol. 1, pp. 195-227). Oxford: Oxford University Press.

Laudan, L. (2016). *The Law's Flaws. Rethinking Trials and Errors?* Milton Keynes: Lightning Source.

Laudan, L., & Saunders, H. (2009). Re-Thinking the Criminal Standard of Proof: Seeking Consensus About the Utilities of Trial Outcomes. Retrieved from SSRN: https://ssrn.com/abstract=1369996 website: doi:http://dx.doi.org/10.2139/ssrn.1369996

Lee, Y. (2016). Reasonable Doubt and Moral Elements. *Journal of Criminal Law and Criminology, 105*(1)

Lee, Y. (2017). Reasonable Doubt and Disagreement. *Legal Theory, 23*(4), 203-257. doi: 10.1017/S1352325217000180.

Lillquist, E., . UC Davis Law Review, Vol. 36, November 2002. Available at SSRN: https://ssrn.com/abstract=349820 or http://dx.doi.org/10.2139/ssrn.349820. (2002). Recasting Reasonable Doubt: Decision Theory and the Virtues of Variability. *UC Davis Law Review*(36), 85-197

Lipton, P. (1991). *Inference to the best explanation*. London-New York: Routledge.

Montmarquet, J. A. (2007). "Pure" versus "Practical" Epistemic Justification. *Metaphilosophy, 38*(1), 71-87

Moore, A. W. (2006). Maxims and Thick Ethical Concepts. *Ratio, 19*(2), 129-147. doi: 10.1111/j.1467-9329.2006.00315.x.

Moore, D. A., & Healy, P. J. (2008). The Trouble With Overconfidence. *Psychological Review, 115*(2), 502–517

Moran, R. (2001). *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton, N.J.: Princeton University Press.

Mulrine, T. V. (1997). Reasonable Doubt: How in the World Is It Defined? *American University International Law Review, 12*(1), 195-225

Nance, D. A. (2001). Naturalized Epistemology and the Critique of Evidence Theory *Va. L. Rev.*(87), 1551-1618

Nance, D. A. (2016). *The Burdens of Proof*. Cambridge: Cambridge University Press.

Nance, D. A. (2018). Belief Functions and Burdens of Proof. *Law, Probability and Risk, 18*(1), 53-76. doi: 10.1093/lpr/mgy020.

Newman, J. O. (2007). Quantifying the Standard of Proof Beyond a Reasonable Doubt: A Comment on Three Comments. *Law, Probability and Risk, 5*, 267−269. doi: doi:10.1093/lpr/mgm010.

Ormston, R., Chalmers, J., Leverick, F., Munro, V., & Murray, L. (2019). *Scottish Jury Research: Findings from a Large Scale Mock Jury Study*. Edinburgh: Scottish Government.

Packer, H. L. (1964). Two Models of the Criminal Process. *University of Pennsylvania Law Review, 113*(1), 1-68. doi: 10.2307/3310562.

Pardo, M. S., & Allen, R. J. (2008). Juridical Proof and the Best Explanation. *Law and Philosophy, 27*(3), 223 - 268

Payne, A. (2005). A New Account of Thick Concepts. *Journal of Value Inquiry, 39*(1), 89-103

Picinali, F. (2009). Is "Proof Beyond a Reasonable Doubt" a Self-Evident Concept? Considering the U.S. and the Italian Legal Cultures towards the Understanding of the Standard of Persuasion in Criminal Cases. *Global Jurist, 9*(4)

Picinali, F. (2012). Structuring Inferential Reasoning in Criminal Fact Finding: An Analogical Theory. *Law, Probability and Risk, 11*(2-3), 197-223. doi: 10.1093/lpr/mgs006.

Picinali, F. (2015). The Threshold Lies in the Method: Instructing Jurors about Reasoning Beyond Reasonable Doubt. *International Journal of Evidence & Proof, 19*(3), 139-153. doi: 10.1177/1365712715571511.

Picinali, F. (2018). Can the Reasonable Doubt Standard be Justified? A Reconstructed Dialogue. *Canadian Journal of Law & Jurisprudence, 31*(2), 365-402. doi: 10.1017/cjlj.2018.17.

Ramsey, F. P. (1954). Truth and Probability. In R. B. Braithwaite (Ed.), *The foundations of mathematics and other logical essays* (pp. 156-198). London: Routledge & Kegan Paul.

Schauer, F. F. (1991). *Playing by the Rules. A Philosophical Examination of Rule-Based Decision-Making in Law and in Life*. Oxford: University Press.

Scheffler, S. (1987). Morality through thick and thin a critical notice of ethics and the limits of philosophy. *Philosophical Review, 96*(3), 411-434

Stein, A. (2005). *Foundations of Evidence Law*. Oxford - New York: Oxford University Press.

Strawson, P. F. (1962). Freedom and resentment. In G. Watson (Ed.), *Proceedings of the British Academy, Volume 48: 1962* (Vol. 48, pp. 1-25): Oup Oxford.

Tappolet, C. (2004). Through thick and thin: Good and its determinates. *Dialectica, 58*(2), 207–221

Tillers, P., & Green, E. D. (1988). *Probability and Inference in the Law of Evidence: the Uses and Limits of Bayesianism*. Dordrecht [etc.]: Kluwer Academic Publishers.

Tribe, L. H. (1971). Trial by Mathematics: Precision and Ritual in the Legal Process. *Harvard Law Review, 84*(6), 1329-1393

Väyrynen, P. (2011). Thick Concepts and Variability. *Philosophers' Imprint, 11*(1), 1-17

Väyrynen, P. (2012). Thick Concepts: Where's Evaluation? In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics, Volume 7* (pp. 235-270): Oxford University Press.

Väyrynen, P. (2013). Thick Concepts and Underdetermination. In S. Kirchin (Ed.), *Thick Concepts*: Oxford University Press.

von Wright, G. H. (2001 [1951]). *A treatise on induction and probability*. London: Routledge.

Walen, A. (2015). Proof Beyond a Reasonable Doubt: A Balanced Retributive Account. *Louisiana Law Review, 76*(2), 355-446

Whitman, J. Q. (2008). *The Origins of Reasonable Doubt: Theological Roots of the Criminal Trial*. New Haven-London: Yale University Press.

Williams, B. A. O. (2011). *Ethics and the limits of philosophy*. London - New York: Routledge.