

In Defence of Reasonable Doubt

GEORGI GARDINER
Rutgers University

In criminal trials the state must establish, to a particular standard of proof, the defendant's guilt. The most widely-used and important standard of proof for criminal conviction is the "beyond a reasonable doubt" standard. But what legitimates this standard, rather than an alternative? One view holds the standard of proof should be determined or justified—at least in large part—by its consequences.

In this spirit, Laudan uses crime statistics to estimate risks the average citizen runs of being violently victimised and falsely convicted. He argues that since the former risk is higher, and the aggregate harms are worse, the standard of proof should be substantially lowered. He presents a formula for calculating the preferred standard.

In this paper I outline various ways Laudan's uses of crime statistics are flawed, and explain how he substantially overestimates risks of victimhood and underestimates costs of false convictions. I also explain why his formula is mistaken, and illuminate consequences Laudan neglects. I conclude that, even if consequences determine the appropriate standard of proof, Laudan's arguments fail to show the standard is too high. I conclude by suggesting that the inadequacies of Laudan's reasoning might be good news for consequence-based justifications of the standard of proof.

1. The Standard of Proof

The state sometimes performs actions against the wishes or interests of individual citizens. They might stop and search an individual, charge him with a crime, incarcerate him, impose a restraining order or curfew, remove child custody rights, or order him to pay costs or fees, such as child support. Standards of proof are levels of confidence required to perform these actions legitimately. Before the state may perform these actions they must determine, to that confidence level, certain facts to be the case.¹ Employing public, clearly understood standards of proof serves a range of functions, including bolstering citizens' trust in the system and protecting individuals from undue loss of liberty.

There are several standards of proof in use across the world. They vary in demandingness according to the seriousness of the state's action—weaker standards are employed for less severe actions—and the consequences risked by not acting. Formal standards include the "reasonable suspicion" standard in the US, which is required for a brief stop and search.² To satisfy this standard the police officer must have a specific, articulable, and individualised suspicion that crime is afoot before they conduct a brief search of an individual. The "preponderance of evidence" standard—usually understood as requiring that the claim is more likely than not—is operative in many Western democracies for most civil and family court cases.³ The "clear and convincing evidence" standard is even more demanding. It is used in equity cases such as right-to-die hearings, wills, libel, child custody, paternity disputes, and commitment to mental institutions.⁴

The "proof beyond a reasonable doubt" (BARD) standard is more demanding still. It is perhaps one of the most familiar aspects of criminal law. In order to find a defendant guilty of criminal charges, the person's guilt must be established beyond a reasonable doubt.

The standard has been adopted throughout the English-speaking world. It is widely employed across Europe, and in many other democratic countries, such as Israel, India, and in the Rome Statute of the International Criminal Court.⁵ It is a keystone of criminal justice systems in liberal democracies.

The BARD standard is far higher than the “reasonable suspicion”, “preponderance of evidence”, or “clear and convincing evidence” standards. This demandingness reflects the widely-held view that false convictions are a significantly worse error than false acquittals.⁶ William Blackstone famously stated, “It is better that ten guilty persons escape than that one innocent suffer”, and many others have expressed similar sentiments.⁷ All else being equal, a higher standard of proof leads to relatively few false convictions and significantly more false acquittals. (Note that, as I explain in section eight, a standard of proof does not equate to a certain ratio of guilty defendants being acquitted to innocent defendants being convicted.⁸)

To satisfy the BARD standard, a juror need not be absolutely certain, and need not be free from all possible doubts; the standard only requires the absence of reasonable doubts. Unreasonable, imaginary, frivolous, hypothetical, or fanciful doubts are not precluded by the BARD standard.⁹ The standard is typically sketched in negative terms, such as having no plausible reason to believe otherwise. More positively, it is often understood to be around 90 to 95% confidence in the guilt of the defendant.¹⁰

But the BARD standard is controversial. There are debates, largely within legal scholarship, about whether the standard can be reduced to a numerical value and, if so, what number and what the number quantifies.¹¹ There are questions concerning how jurors, scholars, and legal professionals do and should interpret the standard,¹² and whether—and in what ways—the standard should be explained during juror instruction.¹³ An important question asks whether BARD is the standard that criminal justice systems ought to employ and, if so, what justifies this standard.¹⁴ A related question asks whether the justice system should employ differing standards depending on the seriousness of the crime and associated punishment.¹⁵ These questions are all helpfully informed by a better understanding of epistemology: the questions themselves concern epistemic phenomena—such as evidence, credences, and chances—and progress is easily hindered by epistemological errors, such as statistical fallacies, misunderstandings of probabilities, and ignorance about the nature of judgement.

2. The Costs of the Standard

In this paper I focus on whether BARD is the appropriate standard for criminal convictions. The criminal justice system, at least in part, is meant to produce better outcomes. Thus if one standard of proof generates less harm than another, then to that extent it is preferable. One view holds that consequences are all that matter for determining and justifying the standard of proof. A more moderate view holds that consequences are not dispositive—other things, such as rights—matter, but that consequences must be taken into account.¹⁶

Larry Laudan has recently developed a consequence-based approach to determining and legitimating the appropriate standard.¹⁷ He argues that an accurate accounting of the aggregate harms reveals that the BARD standard is far too demanding, and should be replaced by a substantially lower standard. Laudan’s attack on BARD has two parts. He first examines BARD from a consequentialist perspective and argues it does significantly more harm than good, and so should be jettisoned. Secondly, he questions whether citizens have a

right to be evaluated using BARD in criminal trials even though on balance it results in aggregate harm. He concludes there is no such right, and so from a consequences-based, and a rights-based, perspective BARD should be abandoned. In this paper I focus on the first strand of Laudan's argument. I argue that even from a purely consequence-based perspective, Laudan's attack on BARD is unconvincing. I conclude by suggesting that, even when restricted to consequence-based considerations, there is something distinctively valuable about employing the BARD standard.

Note that whilst I focus on consequence-based arguments concerning BARD, facts about the relevant consequences have ramifications for the debate in general. Firstly, Laudan frames the rights-based project as "the moral argument for retaining BARD, despite the fact that it does more harm than good" and characterises those who offer a right-based defence of BARD as those who "insist that defendants have an inalienable right to be judged by BARD [...] *whatever the negative consequences of these policies might be*".¹⁸ Laudan's estimations of the harms of BARD thus generate an obstacle for rights-based defences to overcome. Undermining Laudan's estimates affects how we understand the dialectic by removing the *prima facie* case against the BARD standard. Secondly, no reasonable views about the standard of proof hold that consequences are wholly irrelevant. Even rights-based views exhibit sensitivity to outcomes. Thus a clear picture of the outcomes affects the contours of the terrain, even for non-consequentialist views.¹⁹

Laudan's consequence-based criticism of the BARD standard proceeds by noting that the state has two interrelated duties. One is to protect citizens from harm committed by other citizens, the other is to protect citizens from false conviction.²⁰ They are interrelated because many policies and systems that advance one aim thereby hinder the other. If the system lifts restrictions on admissible evidence, for instance, or allots more weight to police testimony, it is easier to convict guilty defendants, and so citizens are protected from harms the defendants would inflict if acquitted. But such changes are likely to increase the incidence of false convictions. Laudan notes that the standard of proof employed in criminal trials is one of these double-edged considerations: a lower standard likely results in a higher conviction rate and increased incarceration of criminals. This, in turn, decreases the incidence of citizens being criminally harmed.²¹ But the lower standard also generates more erroneous convictions.

The current standard of proof, as noted above, is relatively demanding. In order to convict, the jury must judge unanimously, beyond reasonable doubt, that the accused committed the crime. Laudan argues this standard is far too high because many defendants who are falsely acquitted of violent crimes commit further violent crimes during the period they would have been incarcerated if convicted. And so, argues Laudan, the high standard of proof carries substantial costs—it enables a greater number of violent crimes. For this reason, Laudan concludes, a lower standard is preferable.

3. Laudan's Statistics: A Grim Picture

Laudan's cost estimates are largely informed by crime statistics in the US and UK. He employs Bureau of Justice Statistics research on lifetime likelihoods of victimisation to conclude the lifetime chance of "being a victim of a serious, violent crime (excluding homicide) is about 83%". Laudan estimates the lifetime chance of being wrongfully convicted of serious crime is "at most about 0.1%".²² To motivate that the current standard

of proof is too high, Laudan contrasts these probabilities and concludes the average citizen is 830 times more likely to be violently victimised than falsely convicted of violent crime.²³

Laudan uses data on recidivism rates, and Alfred Blumstein's studies of criminal behaviour, to estimate the rate that the falsely acquitted commit crime during the period they would have been incarcerated.²⁴ Laudan notes,

If we settle on the lower end of Blumstein's estimate [...] we can infer that the average person who is falsely acquitted of a violent crime will continue committing a couple of them every year for about nine years, before abandoning a life of crime.²⁵

and claims,

[Blumstein's] team concluded that the best estimate of the number of felonies committed annually by the typical criminal ranges between ten and fifteen, of which between two and four are violent crimes.²⁶

Laudan argues that since the average length of incarceration for violent crime is 3.6 years then, given Blumstein's estimates, each false acquittal enables on average more than 36 crimes, including seven violent ones.²⁷ These crimes, Laudan contends, are the average committed by those falsely acquitted of violent crime during the period the defendant would have been incarcerated and so unable to commit crime.²⁸ Laudan emphasises that this statistic "undergirds" his appraisal of BARD.²⁹ In a given year, holds Laudan, the average citizen is 6 to 10 times more likely to be criminally victimised by a falsely acquitted defendant during the time he would have been incarcerated, than he is to be falsely convicted of violent crime.³⁰ Laudan concludes,

The object lesson seems to be clear: we are utilizing a standard of proof that is far higher than it should be, because it leads to far more harm in aggregate than a lesser standard of proof would. Indeed, BARD would make sense as a standard of proof for violent crimes just in case the harm done by a false conviction was ten to twenty times greater than the harm done by a false acquittal. There is no way that such an assessment of harms can be plausible, in light of the data rehearsed here on the recidivism of violent felons.³¹

In addition to arguing the standard of proof should be lowered, Laudan also endorses a formula for calculating his preferred standard, and offers a new standard using his estimates of relevant harms. The formula Laudan uses is:³²

$$\text{Standard of Proof} = 1/[1+(\text{COST}_{\text{FA}})/(\text{COST}_{\text{FC}})]$$

Where COST_{FA} is the cost of a false acquittal, and COST_{FC} is the cost of a false conviction. Laudan suggests that given the costs of false acquittals—chiefly, the crimes they enable—and costs of false convictions, the standard of proof should be about 56 to 67%.³³ Laudan suggests, furthermore, the justice system should aim for a ratio of false acquittals to false convictions of about 2:1—a ratio substantially higher than Blackstone's oft-cited 10:1.³⁴

Thus if Laudan's calculations are close to accurate, his argument strongly indicates the current standard causes substantial harm, and consequence-based reasoning recommends dramatically reducing the standard. In the remainder of this essay, I criticise Laudan's consequence-based argument. In sections four and five I assess the statistics Laudan employs, and show the grim picture painted by Laudan's statistics does not reflect reality. My

criticisms point to important ways that applied epistemology illuminates issues concerning the standard of proof. This includes understanding the base rate fallacy, the potential of averages to mislead, the artefacts of sampling methods, diminishing returns, and similar statistical pitfalls. One central subject matter of epistemology concerns how to interpret the world using limited information and how to extrapolate from incomplete data. I demonstrate “epistemology in the field” by applying this to Laudan’s crime rate estimates.³⁵ In sections six and seven I examine costs of conviction and incarceration Laudan neglects. I argue that, when illuminated by a more accurate portrayal of relevant harms and risks, Laudan’s arguments lose their force. The broader lesson is that if we think that consequences matter, consistency requires that we consider all relevant consequences, and not a mere subset. In section eight I assess the formula Laudan advances, and explain its inadequacies. Again, the applied epistemology tools of reasoning with limited information and formal epistemological insights illuminate the debate. In section nine I conclude by examining broader costs that bear on a consequence-based approach to ascertaining a standard of proof, and suggest that—even from a consequence-based perspective—there is something distinctively valuable about BARD. I draw on recent cognitive science, for example, to suggest that people reason better with notions such as “reasonable doubt” than with statistics, and this underpins a distinctive value of the BARD standard. I conclude that Laudan’s arguments for lowering the standard of proof are unsuccessful. I hope to illustrate the important role of applied epistemology in understanding questions about the criminal justice system.

4. A Second Look at the Statistics: Lifetime Risks

One of Laudan’s most stunning statistics is “the lifetime risk an average American runs of being a victim of a serious, violent crime (excluding homicide) is about 83%”. Laudan claims that violent crimes are “homicide, rape, aggravated assault and armed robbery”.³⁶ This frightening statistic vividly proclaims that the average citizen faces significant risk of violence.

But the statistic is not as potent as it first appears. Firstly, note the cited report is from 1987, and uses victimisation rates from 1975 to 1984. The report’s author, Herbert Koppel, assumes victimisation rates will not change.³⁷ But it is widely recognised that crime in America peaked between the 1970s and late 1980s, and has declined significantly since. Between 1993 and 2010, for example, victimisation rates for violent crime and serious violent crime declined 70% and 73% respectively.³⁸ Using recent statistics produces *substantially* lower victimisation risks; countenancing the decline in crime since Koppel’s data was collected reduces the lifetime likelihood to less than 30%.

Note too whilst Laudan claims the 83% figure concerns “serious, violent crime”, the estimate includes all “violent crime”.³⁹ The former is typically worse to experience. Similarly, Laudan claims Koppel’s statistic concerns “rape, aggravated assault and armed robbery”, but Koppel makes explicit the figure concerns “rape, robbery, and assault”.⁴⁰ The figure includes all robbery, in other words, not specifically armed robbery, and all assault, not just aggravated assault. Koppel is explicit, furthermore, that his 83% figure includes threats and attempts.⁴¹ Laudan does not mention this. Restriction to completed crimes halves the figure; including only assaults and robberies that result in physical injury to the victim almost halves it again.⁴² By suggesting that the statistic concerns the completed or more violent version of the crime, Laudan radically over-represents the risks.⁴³

So far we have examined difficulties with Laudan's use of Koppel's statistics, but Koppel's statistics themselves are flawed. Koppel's calculation of lifetime risk presupposes likelihoods are evenly distributed throughout the population. But—as Koppel recognises—this is far from true. Criminal victimisation distribution is strikingly uneven; previous victimisations correlate strongly with further victimisation. Koppel's simplifying assumption leads him to massively overestimate the proportion of people who are victimised at least once, and accordingly underestimate the proportion who are victimised many times in their life. Koppel substantially overestimates, in other words, the lifetime likelihood of victimisation.⁴⁴

Thus, for a range of reasons, Laudan substantially overestimates when he claims “the lifetime risk [...] of being a victim of serious, violent crime (excluding homicide) is about 83%”.⁴⁵ This overestimation is important since it is a central statistic supporting Laudan's case.

Recall Laudan contrasts lifetime risk of violent victimisation with lifetime risk of false conviction for violent crime. He notes the latter risk—which he calculates as “at most 0.1%”—is relatively low.⁴⁶ But Laudan calculates this figure using 1991 incarceration rates, which indicate a lifetime likelihood of imprisonment of 5.1%.⁴⁷ The proportion of citizens incarcerated has risen since, which increases lifetime likelihood of imprisonment and—one would expect—chances of false conviction.⁴⁸ The lifetime likelihood of incarceration calculated in 2001 is about 6.6%, and is now slightly higher.⁴⁹ This is a relatively small difference, but it bears noting that the risks of victimisation decreased, and risk of false conviction may have increased, since Laudan's data was gathered. Laudan's reasoning would have been more accurate, and less compelling, if he used recent statistics.

Note too that risks of violent victimisation and false acquittal have an uneven distribution across social groups. Disabled citizens, inmates, and the poor, for example, have heightened victimisation risks. Race is also an important social factor in estimating lifetime risks. Black citizens are significantly more likely to be falsely convicted of violent crime than white citizens, for example, and a proposed trade-off that does not consider this disparity is likely to further increase this racial discrimination. Given the uneven distribution of risks, a trade-off that seems reasonable to one racial group might well seem unreasonable to another racial group.

This skewed distribution suggests an aggregated lifetime risk obscures, rather than illuminates, the true risks. To illustrate this, consider forcible rape. Women face a miniscule risk of false conviction for forcible rape, and a high risk of victimisation, relative to men. Given women's aggregate risks, Laudan's approach recommends a low standard of proof. Men face a higher risk of false conviction, but a lower risk of victimisation. Laudan's approach accordingly recommends a high standard of proof. When risks have a skewed distribution, simply aggregating the risks—as Laudan proposes—may fail to adequately reflect the risks that real people confront. One problem with setting a standard of proof by aggregating the risks is that no real individuals face aggregated risks; risks faced by real people have skewed distributions.⁵⁰

Finally, whilst the comparison between the lifetime risks of being falsely convicted of a violent crime and of being violently victimised might be striking, in the context of adjudicating trade-offs in the trial system the more relevant comparison is with the lifetime risk of being violently victimised by someone who was falsely acquitted during the time that they would have been incarcerated. This is because altering the standard of proof plausibly

bears directly only on that smaller subset of violent victimisations, and influences the rate of other violent victimisations only indirectly, if at all. This more refined risk is a small fraction of the total, and so the relevant comparison is significantly less suggestive of a need to lower the standard of proof. This further undermines Laudan's use of the comparison of lifetime risks.⁵¹

5. The Cost of False Acquittals: Crime Rates and Residual Careers

Another important statistic Laudan deploys—one that “undergirds” his critique of BARD—is that every false acquittal enables an average of more than 36 crimes, including seven violent ones.⁵² These are the crimes, Laudan claims, each falsely acquitted defendant will commit, on average, during the time they would have been incarcerated. The figure is based on Blumstein's studies of criminal behaviour, which Laudan takes to describe the “typical criminal” and so illuminate the average behaviour of the falsely acquitted.⁵³ As I explain below, the difficulties with using Blumstein's research for Laudan's calculations are (at least) threefold. Firstly, Blumstein's study does not describe average criminal behaviour; it describes “career criminals” who are, by definition, more active than average. Secondly, given crime distribution patterns, mean averages are not useful for understanding crime or generating policy. Thirdly, Laudan neglects chances of re-arrest and the distinction between total and residual career durations.

The first problem: Blumstein's research is concerned with distinguishing high-rate from low-rate offenders, and is motivated by the extreme variation in crime rates amongst offenders. This striking distribution—to which Blumstein frequently returns⁵⁴—means that amongst those who commit crime most commit only a handful of crimes each; a tiny minority commit a huge amount each.⁵⁵ Blumstein argues criminal justice policy must countenance the skewed distribution.⁵⁶

The statistics Blumstein uses concern “active violent offenders”, which Laudan claims includes anyone who has committed a violent crime.⁵⁷ But this is *not* Blumstein's use of the term. Blumstein explicitly distinguishes those who have ever committed a crime (“participants”) from “active offenders”.⁵⁸ Blumstein notes the former category is vastly larger and the categories differ so radically they do not illuminate each other. Blumstein glosses “active offender” as anyone who commits at least one crime during an observed period.⁵⁹ This differs from participation statistics, since the shorter the observed time period, the larger the “discovered” average crime rate.

To see why, suppose I decide to research runners' behaviour, so I first identify a sample set of “runners”. If I identify “runners” as anyone who ran at least once during the period “the last five years”, my sample will include many low-rate runners. Accordingly I “discover” a relatively low average running rate amongst runners. Suppose instead I employ the time period “yesterday”. Using this duration my sample has a higher proportion of frequent runners, and I “discover” a higher average running rate. The difference arises because the set of people who ran yesterday contains a higher proportion of frequent runners than the set of people who ran at least once in the last five years. My hypothetical research into runners is not thereby useless. By defining “runners” as those who ran during a certain time period I can determine what runners of various running frequencies tend to wear or weigh, for example. But I cannot discover *average running frequencies* when my method identifies “runners” using running rates. This is a basic feature of statistical sampling.

This has implications for Laudan's use of Blumstein's research. Low-rate offenders are not represented in Blumstein's average. Suppose that to qualify as an "active violent offender" a person must have committed a violent offence within the last month. We would expect the sample to contain a high proportion of high-rate offenders; the entry qualifications filter out those who only rarely commit violent offences. By disqualifying the lower end of the spectrum, the demanding entry requirement for inclusion in the sample inflates the average rate. If inclusion in the sample requires, by contrast, the person to have committed violent crime within the last ten years, the average offence rate will be lower. Suppose Blumstein's "active violent offenders" are those who committed violent crime within the last year. It would then not be surprising that on average active violent offenders commit 2 to 4 offences per year. That is a likely consequence of Blumstein's sampling method.⁶⁰

Given the skewed distribution of crime—most people who commit crime do so rarely—and the fact that the average rate of 2 to 4 crimes per year concerns only *active* violent offenders, we should predict the rate of violent crime committed by people on trial for violent crimes is significantly less.⁶¹ Returning to the running analogy, suppose the vast majority of people run once per year, and some tiny minority run every day. If you classify "runners" as those who ran in the last day, you would "discover" a relatively high average running rate. But when you see someone running in the park, you should not assume she runs at that rate; she is likely to be one of the many people who run infrequently.

Recall that Laudan writes, "Blumstein's [...] team concluded that the best estimate of the number of felonies committed annually by the *typical criminal* ranges between ten and fifteen, of which between two and four are violent crimes".⁶² This is doubly misleading: Firstly, Blumstein's estimate concerns active, not average criminals. Secondly, Blumstein's statistic concerns only violent criminals, not all criminals. It is false that the typical criminal—or even the typical active criminal—commits two to four violent crimes annually.

The second problem: Suppose Laudan were correct that by "active violent offender" Blumstein meant everyone who has ever committed a violent crime, and so the average Laudan cites is the mean average rate of violent crime among participants. The skewed distribution that animates Blumstein's work has consequences for Laudan's use of mean averages. Consider, for illustration, a society comprising 100 people. 99 people each commit one violent crime per year, and one person commits 99 violent crimes per year. This approximates the top-heavy distribution Blumstein describes.⁶³ The mean crime rate is 1.98 violent crimes per year. But consider the "typical criminal": the mode and median averages are one crime per year; the vast majority of criminals (99 out of 100) have very low crime rates. If we pick a criminal at random it is overwhelmingly likely (99%) his crime rate is lower than the mean. Supposing each crime is equally likely to result in trial, from the perspective of the court trying a guilty defendant there is an approximately 50% chance the defendant is a low-rate offender, and an approximately 50% chance he is the high-rate offender. Expecting each acquittal to result in approximately 1.98 crimes per year paints an inaccurate picture. The story is more complex. Approximately half acquittals are of low-rate offenders, and so result in (on average) one crime per year until the person retires from crime, the other (approximately) half of acquittals acquit the high-rate offender and so allow much more crime.⁶⁴ But that person will be tried 99 times as often as the others—as often as the others combined—and so is far more likely to eventually be imprisoned, even with a higher standard of proof. His incarceration means much less crime happens, and the remaining

trials all concern low-rate defendants. Thus the skewed distribution entails that mean averages obscure, rather than illuminate, the true story.⁶⁵

Blumstein's research shows crime distribution is skewed dramatically. Some defendants commit many crimes each, and most defendants commit few. This pattern, which is a universal feature of crime distribution, means targeting high-rate offenders is an effective way to reduce crime. And crime policies should not be based on mean averages, since mean averages obscure crime patterns. Crime policies that affect all offenders should not be based on mean averages of active offenders, since active offenders are the minority of criminals, and (perhaps) appear in the minority of court cases. I suspect that lowering the standard of proof is not an effective way to target high-rate offenders; a more predictable result is that lowering the standard sends disproportionately more low-rate offenders to prison (and, as I suggest in section six, incarceration may make low-rate offenders more active). Instead a relatively high standard of proof might be more effective given the skewed distribution: false convictions are reduced, and the cost of false acquittals for high-rate offenders is low, because high-rate offenders are likely to be prosecuted again in the future.

Thirdly, Laudan writes, "if we settle on the lower end of Blumstein's estimate [...] we can infer that the average person who is falsely acquitted of a violent crime will continue committing a couple of them every year for about nine years, before abandoning a life of crime".⁶⁶ I described above problems with Laudan's annual rate estimates and his making claims about the average falsely acquitted person. There are also problems with Laudan's "nine-year" prediction. Laudan claims the falsely acquitted continue committing crime for an average of nine more years following acquittal. Criminals tend to have an "active career" of a certain length, after which they abandon crime. But Laudan's claim is dubious. Firstly, the falsely acquitted person can be re-arrested, cutting short his criminal activity. Laudan fails to consider the chance of later arrest and conviction. The more active the criminal, the more likely a re-arrest; over time he is more likely be found guilty than a less active acquitted defendant, which reduces the cost of false acquittals. A significant virtue of the justice system, with regard to false acquittals, is that repeat offenders generate another chance to convict, whereas non-repeat offenders do not. This allows the system to target repeat offenders.

I cannot (despite careful inspection) find the "nine year" number in Blumstein's article, so I cannot fully evaluate Laudan's use of it. I cannot see, for instance, whether "nine years" represents the criminal career of a typical defendant who is falsely acquitted of a violent crime, as Laudan claims, or whether it represents the behaviour of another subset, such as active violent offenders, the once incarcerated, or those currently incarcerated for violent offences.⁶⁷ This is important because active and imprisoned offenders have longer than average careers and, as with many crime statistics, career length trends have a strikingly top-heavy distribution: Most offenders have very short careers, consisting of just a few crimes, and a small—active and long-term—minority continue committing crimes for many years. Most defendants have careers significantly shorter than the mean.⁶⁸ And given the selectivity of the justice system, which targets "problem cases", the incarcerated tend to have longer (and more active) careers. This is because at sentencing repeat offenders are given longer sentences than first-timers, are less likely to receive parole, and are more likely to receive extended sentences owing to prison crime, all of which raises the prevalence of active offenders within prisons. Surveying prisoners over-represents active, long-term offenders, in other words, relative to the total criminal population. Since Laudan does not say where he

finds the “nine year” figure, I cannot adjudicate whether he corrected for these factors. But given his methodological errors concerning crime rates—Laudan employs Blumstein’s figures for “active offenders” as a proxy for crime rates of the falsely acquitted—there is reason to think he committed the same error with career length.

Blumstein claims that most criminal careers are short—perhaps on average about five years for index offences—many ending whilst the offender is still teenage.⁶⁹ Admittedly violent offenders tend to have longer careers than other offenders, but nine years does not accord well with current estimates. Estimating the average length of criminal careers is important because if the expected residual career is shorter than the incarceration, this further reduces the cost of false acquittal. If the average career is five years, for example, and the average sentence is four years, many careers dwindle before the sentence would have ended, and so false acquittal does not enable the full four years’ worth of crimes. This affects Laudan’s calculations. There are two important factors. Firstly, the acquitted have shorter than mean-average careers, because most criminals do (recall the skewed distribution). Secondly, trials are interspersed across a criminal career. The shorter the career, the more likely the criminal is approaching the end of his career when arrested. To the extent the defendant is approaching natural career cessation, the false acquittal allows fewer crimes. In estimating costs of false acquittals, Laudan neglects the distinction between residual career—the length of criminal career remaining at time of trial—and total career length. This omission over-predicts the cost of false acquittals.

Thus Laudan overestimates the average amount of crime enabled by a false acquittal. He overestimates—in a variety of ways—the crime rate of the falsely acquitted and neglects the chance their criminal behaviour ceases. In short, false acquittals are less costly, in terms of crime enabled, than Laudan’s use of Blumstein’s statistics suggest.

6. The Criminogenic Effects of Prison

Laudan estimates the average crime enabled by each false acquittal principally using Blumstein’s research into career criminals, but he also advocates other sources of data, such as recidivism rates upon release from prison.⁷⁰ I noted above that, for a variety of reasons, inmates have a higher average crime rate than the acquitted.⁷¹ But there is a further problem: Plausibly people commit more crime just after leaving prison than at other times. If this is correct, recidivism estimates record a peak, not a typical, crime rate.

Reasons to think recent incarceration increases a person’s criminal activity include that the inmate acquires criminal friends, contacts, and skills, and upon discharge lacks alternative skills, associates and opportunities.⁷² The inmate might associate with gangs for protection in prison, which precipitates criminal behaviour. Inmates are treated as criminals during incarceration, and may internalise the norms and self-image suggested by their treatment.⁷³ Exposure to violence and crime whilst incarcerated desensitises many inmates.⁷⁴ Prison exacerbates mental health issues—especially where the inmate suffers physical and sexual violence or solitary confinement—which correlates with increased criminality.⁷⁵ Experiencing prison rape can lead inmates to commit acts of violence upon release.⁷⁶ The incarcerated can lose employment and employment opportunities. In some states felons cannot work in public employment, are denied professional licences, cannot join the military, and are not permitted to drive. Incarceration can cause gaps in employment records. Work experience and skills can become out-dated. Convicted offenders often forfeit rights to

subsidised housing, welfare, food stamps and other social programmes; incarceration can sever ties to family and friends, and threaten child custody.⁷⁷ All these factors correlate with increased criminal behaviour.⁷⁸

If incarceration causes a spike in criminality, then recidivism rates are a poor guide to the crime rates of the falsely acquitted; recidivism rates over-predict the cost of false acquittal.⁷⁹ But the criminogenic effects of incarceration do not merely influence attempts to estimate crime rates. It also bears on the utility calculus, and thereby—according to Laudan’s reasoning—the preferred standard of proof. If imprisonment *causes* crime, there is additional cost to incarcerating each individual, guilty or not. This consideration affects each individual trial decision by increasing the cost of individual convictions, and affects the overall utility of imprisonment: The higher the prison population, the worse the criminogenic effects of prison.

It is not only the imprisoned who commit crime as a result of incarceration. Families of those incarcerated, especially children, are significantly more likely to commit crime.⁸⁰ The broader community can also suffer increased crime.⁸¹ For some offences, such as selling drugs, incarceration can result, directly and swiftly, in the initiation of a new offender. Not only does the crime continue, the number of criminal participants thereby increases.⁸² And dismantling drug gangs can trigger a peak in violence, as rivals compete for newly available terrain. Some evidence suggests increased incarceration levels cause the general public to commit more crime: once incarceration becomes normalised, prison loses its deterrent effect.⁸³ So more inmates—a predictable result of lowering the standard—might exacerbate society’s crime incidence. These are consequences Laudan ought consider, given his project, but neglects; they increase the costs of conviction and so affect his utility calculus.

7. Prison Crime

Another problem with Laudan’s statistics concerns crime within prison. Prison crime is substantial and bears on the utility calculus. Crime occurring within prisons is rarely represented in official crime statistics, and is not included in the figures Laudan employs.⁸⁴ Laudan’s utility calculations assume the incarcerated do not commit violent crime. But this is mistaken: convicted people can commit violent crime within prison. Much prison crime is the simple transference of crime from outside to within prison. Other prison crime is caused by the incarceration, and would not have otherwise occurred.⁸⁵ Laudan neglects both kinds of prison crime.

Prison crime rates are exacerbated by local social norms in which new inmates use violence to establish their status in the hierarchy, and prison rape culture dictates that an inmate must sometimes choose between succumbing to frequent sexual assaults in a prison “marriage” or fighting to demonstrate strength.⁸⁶ But prison crime is not limited to violence amongst inmates. Incarceration causes guards to commit and suffer crimes they otherwise would not.⁸⁷ These crimes should be represented in Laudan’s calculus. Insofar as prison crime is worsened by overpopulation—by straining the system—lowering the standard is doubly costly: It incarcerates more individuals, which directly generates expected prison violence towards those inmates, and it results in more incarcerations overall, exacerbating the overpopulation problem.

Laudan supposes “it is twice as costly to be falsely convicted of a violent crime than to be the victim of such a crime”, and proposes a revised Blackstonian ratio of 2:1.⁸⁸ But

evidently when adjudicating this Laudan neglected prison rape and other prison crimes; being falsely convicted substantially elevates risks of violent or lethal victimisation. Inmates might well, on average, be victim to more than two violent crimes during incarceration whilst enduring other harms of conviction. Note too victims of prison crime typically do not enjoy societal support structures. Wardens ignore or participate in prison crime, criminal investigations rarely occur, the victim is separated from friends and family, and many members of society think the inmate “deserves” the crime, or that it constitutes part of the punishment. By ignoring prison crime, Laudan underestimates the true costs of conviction.

Laudan writes “BARD would make sense as a standard of proof for violent crimes just in case the harm done by a false conviction was ten to twenty times greater than the harm done by a false acquittal”, and dismisses this possibility outright.⁸⁹ But in light of the prevalence of prison violence (and prison’s other criminogenic effects), the ratio does not seem implausible: convictions are bad for the defendant and those around him, and may cause substantial harm and crime. Plausibly the harms are greater where the convicted is innocent.⁹⁰ Meanwhile false acquittals, as argued above, are less costly than Laudan suggests, since they enable less crime than he estimates. Thus, by Laudan’s own lights, the statistics might well vindicate the current standard.

The lifetime risk of being falsely convicted of violent crime is, according to Laudan, “at most 0.1%”.⁹¹ But in light of the discussion of widespread victimisation within prison, 0.1% is actually quite high. A lifetime chance of 0.1% means that false conviction for violent crime can be expected to happen to roughly one in every thousand people. (Far more, presumably, are falsely convicted for less serious crimes. These inmates also risk victimisation whilst incarcerated.) Given this only counts false conviction for serious crimes, and given how bad prison can be, this is a high proportion, and provides some reason to think we should be more cautious about false convictions, not less. Recall too from section four that this risk should be contrasted with a tiny fraction of the 83% that Laudan originally claimed. A critical appraisal of Laudan’s statistics suggests his statistics-based case for lowering the standard is unsuccessful.

8. The Formula

As noted in section three, the formula Laudan advocates for calculating a preferred standard of proof is,

$$\text{Standard of Proof} = 1/[1+(\text{COST}_{\text{FA}})/(\text{COST}_{\text{FC}})]$$

This formula cannot be correct; it is too simple. Firstly, it fails to consider relative prevalence of guilt and innocence in the defendant population, a variable that affects which standard maximises expected utility. To illustrate, compare two cases. In Scenario A the ratio of innocent to guilty defendants is 100:1. Perhaps in this society few people commit crimes, yet prosecutors are wildly over-zealous. In Scenario B the ratio is 1:100. Perhaps in this society prosecutors try only cases with persuasive incriminating evidence, and so the overwhelming majority of defendants are guilty. There are many innocent defendants in the first scenario, and many guilty in the second. For any given standard, the courts oversee more false convictions, and less false acquittals, in Scenario A; they oversee more false acquittals and less false convictions in Scenario B.⁹² Comparing scenarios A and B illustrates that Laudan

cannot use utility calculus to calculate his preferred standard without considering this variable. By neglecting relative prevalence of guilt and innocence in the defendant population, Laudan commits the base rate fallacy.

Secondly, Laudan fails to consider the distribution of credences amongst jurors, which also bears on Laudan's preferred standard. To illustrate, consider a standard of proof of 70%, and compare two cases: In Scenario C, jury confidence bunches around 40% confidence in the verdict "guilty", with few confidence levels higher or lower. Perhaps jurors are diffident or prosecuting lawyers incompetent. This means many defendants are acquitted—falsely or not—and few convicted. In Scenario D, jury confidence levels tend to bunch around 80%, with few confidence levels higher or lower. Perhaps prosecuting lawyers are excellent, and defence lawyers terrible. In Scenario D many defendants are convicted and few acquitted. This shows Laudan's calculus must include a variable for, or assumption about, credence distribution in jurors' judgements.

There is a third error: A utility calculus to determine a preferred standard must take into account the relative likelihood of each mistake: false acquittal and false conviction. If it is almost impossible, for example, to find an innocent defendant guilty, then for any given standard of proof, there would be more false acquittals and fewer false convictions; if it is relatively difficult to find a guilty person innocent, the opposite obtains. If police officers were to routinely lie and tamper with evidence to secure convictions, including false ones, for example, this would make false convictions more common relative to false acquittals. If jurors distrust the system and so are reluctant to convict, then false acquittals would be more common than false convictions. The formula Laudan advances does not account for these variables, and so is inadequate.

Laudan notes a typical citizen faces a risk of being falsely convicted of violent crime and of being the victim of a violent crime committed by someone who was falsely acquitted (during the time they would have been incarcerated). He claims, "Faced with these two risks, presumably he would like to see the standard of proof and rules of trial set at that point that minimises the aggregate dangers that he runs from these two risks."⁹³ This cannot be correct: if no one were ever charged with violent crimes, the aggregate risk would be zero: no one is falsely convicted or falsely acquitted. But this is not desirable. The relevant risk is not "victimisation by someone who was falsely acquitted (during the time they would have been incarcerated)"; it is "victimisation simpliciter". Even if—as seems doubtful—consequences are all that matter to the standard, Laudan's utility calculus is inadequate.

9. Other Relevant Consequences

The standard of proof Laudan suggests is about 56% to 67%.⁹⁴ A consequence-based endorsement of standards within this range provides good evidence that there are consequences Laudan neglects. A standard as low as 67% (let alone 56%) is implausible. A standard of 67% means someone could convict with about the same confidence I have in the proposition "this fair dice will not land one or two". A standard of 56% approximates a coin toss. This is not a good basis for a criminal justice system.

Laudan neglects all consequences independent from whether any individual trial ends in acquittal or conviction. He thus neglects the consequences of the standard itself. This is reminiscent of someone's adopting a Byzantine decision procedure because the outcome of each individual subsequent decision is improved, without realising they now spend their

entire lives making complicated decisions. When adjudicating among decision procedures, one should not simply aggregate the result of the individual decisions. The procedure itself has consequences that must be considered, but Laudan does not do this.

A significantly lower standard is likely to increase the number of arrests and charges, as police recognise they need less evidence to secure conviction. This will be costly. Citizens would distrust a system that could incarcerate on such meagre evidence, and might well feel bitter and aggrieved. The sense of injustice could lead to riots and chaos. A drastic change in the standard will generate a destabilising effect, and may skyrocket the prison population, inflating costs and precipitating violence. The American prison population is already enormous, and an increase to this population—a likely result of Laudan’s proposal—will exacerbate overpopulation and expense problems.

Criminal justice systems worldwide exhibit severe racial disparities, which a lower standard will exacerbate. One way a lower standard would increase injustice is simply by sending more people—and so disproportionately more black and Hispanic people—to prison. But worse, there is evidence that black defendants on trial are more likely to be falsely convicted than white defendants.⁹⁵ False convictions are already distributed disproportionately, and lowering the standard will worsen this. Lower standards are also likely to disproportionately affect those who fit “criminal” stereotypes because if standards are lower, juror prejudices are more potent. These foreseeable further increases in racial disparities are costs of Laudan’s proposal.

Another consequence of radically lowering the standard within one country—such as America—concerns nations influenced by that country. If America jettisons BARD, as Laudan proposes, other countries—ones with less adequate justice systems or human rights protections—might follow suit, generating severe harm. The BARD standard is currently spreading across the globe,⁹⁶ but America’s radically lowering the standard of proof may well interrupt this spread, with devastating effects.

Laudan neglects the diminishing returns of incarceration: Laudan uses consequence-based reasoning to advance incarceration-based crime reduction policies, so ought countenance the widely-recognised fact that increasing incarceration reduces the marginal expected utility of each new incarceration.⁹⁷ Given the diminishing returns of incarceration, an uptick in the incarcerated population—the aim of Laudan’s proposal—reduces the expected value of each conviction.

Another cost of altering the standard is there might be something special about BARD in reasoning. Laudan holds the standard of proof is a “malleable tool” that can be “fine-tuned” because it can be set at any percentage, which allows enormous variation.⁹⁸ But recent research in cognitive science suggests people are not adept at reasoning using probabilities expressed as percentages.⁹⁹ If correct, jurors are ill-equipped to reliably adjudicate whether, for example, evidence indicates a 63% or 73% chance of guilt. People are better, research suggests, at adjudicating standards that are robustly normative and qualitative. This seems to be how normal reasoning functions. Judging whether a claim is “reasonable” is more familiar and practicable than judging whether something is 60% likely. If correct, there is a cost to jettisoning the “reasonable doubt” standard, as Laudan advocates. It means we must adopt a standard, $n\%$, that people are less adept at applying.¹⁰⁰

There are other difficulties with Laudan’s utility-calculus approach to determining his preferred standard. Crime statistics change over time. Indeed criminal justice policies aim for precisely this result. But when relevant crime rates change, this raises the question of

whether the standard of proof should change accordingly. If crime rates decline, for example, should the standard rise? If Laudan thinks the standard should not shift with crime rates, he is faced with the consequence that the standard does not maximise expected utility. If he thinks the standard should evolve with the crime statistics, however, there is the cost of operating a justice system whose standard of proof fluctuates. This is likely to cause confusion and complication. And if instead Laudan argues the standard should change with crime statistics, but only occasionally—when the discrepancies are large—a version of both problems arise: most defendants are tried according to a standard that does not maximise utility and, whilst the standard does not change often, when it changes it alters by large amounts; a wildly varying standard will contribute to distrust and instability in the system.

A further outcome of Laudan's proposal is that different nations, and perhaps even different provinces within one nation, should employ different standards. The standard depends on their current crime rate. This itself seems like a cost. There is plausibly a value in each criminal justice system employing the same standard.

Laudan focuses largely on the standard of proof for violent crime. This raises the question of whether the standard for non-violent crime should be lowered to the same amount.¹⁰¹ If it should, then the standard of proof for non-violent crime would probably not maximise expected utility. But if the standard should remain high for non-violent offences, there are implausible consequences. A person might be tried for murder and petty theft, for example. On the same quality of evidence, he ought be convicted of murder and acquitted of petty theft. It seems implausible that a court should require less evidence to find a defendant guilty of the more serious crime; the punishment is much more significant. As noted in section one, typically higher standards of proof are required for more severe state actions.

Different social groups commit crimes at different rates. Women, for example, commit fewer violent crimes than men. Suppose a purely consequence-based approach recommends different standards of proof for different groups, perhaps by recommending a lower standard for male defendants. Given this possibility, Laudan must choose: either he acknowledges some rights-based reason to try men and women by the same standards, despite the harm-inducing consequences, or he agrees that, inline with the utility calculus, men should be tried by lower standards. The former recognises the role of rights, and thereby weakens his argument against rights-based considerations in law, by revealing that harm-reducing consequences are not the only relevant considerations. By choosing this horn, Laudan acknowledges he must deploy right-based considerations at some points to protect his view from implausibility. But if he chooses the second horn, and advocates gender discrimination in standards of proof, the result is unpalatable and renders his position untenable to all but the most diehard consequentialists about policy.¹⁰²

Thus 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.¹⁰³

- ¹ The presumption of innocence in criminal law, which is widespread amongst industrialised countries, means the burden of proof falls on the state: the state must establish that the defendant performed the offence. The defendant need not, by contrast, prove that he did not. There are some exceptions, in which the burden shifts to the defendant. In England and Wales, for example, if the defence relies on some "exception, exemption, proviso, excuse or qualification", then the burden of proof for establishing the exception rests on the defendant. The defendant must establish the exception to the "balance of probabilities" standard (Magistrates' Courts Act 1980, s.101).
- ² *Terry v. Ohio*, 392 U.S. 1 (1968).
- ³ This standard is also known as the "balance of probabilities" standard. See *Winship*, 397 U.S. at 371–72; House of Lords in *Re B (A Child)* (2008) UKHL 35.
- ⁴ Leonard B. Sand and Danielle L. Rose, "Proof Beyond All Possible Doubt: Is There a Need for a Higher Burden of Proof when the Sentence may be Death?" *Chicago-Kent Law Review* 78 (2003): 1359–1375, at pp. 1362–3.
- ⁵ Alec Walen, "Proof Beyond a Reasonable Doubt: A Balanced Retributive Account" *Louisiana Law Review* (forthcoming); Thomas Mulrine, "Reasonable Doubt: How in the World Is It Defined?" *American University Journal of International Law and Policy* 12, 1 (1997): 195–225.
- ⁶ Following Laudan, I use "false conviction" to denote when someone is convicted or pleads guilty even though they did not commit the crime, and "false acquittal" for whenever the person committed the crime, but the charges are dismissed by a judge, dropped by a prosecutor, or the person is found not guilty. See Larry Laudan, "The Rules of Trial, Political Morality and the Costs of Error: Or, Is Proof Beyond a Reasonable Doubt Doing More Harm Than Good?" in Leslie Green and Brian Leiter (eds.), *Oxford Studies in Philosophy of Law* (Oxford: Oxford University Press, 2011), pp. 195–227 at p. 202.
- ⁷ William Blackstone, *Commentaries on the Laws of England* (Clarendon, 1765–69). For an overview of similar claims see Alexander Volokh "n Guilty Men" *University of Pennsylvania Law Review* 146 (1997): 173–216.
- ⁸ The standard of proof is not the only factor that influences false conviction and acquittal rates. Other factors include relative prevalence of guilt and innocence in the defendant population and distribution of credences among jurors. As I explain in section eight, these variables underlie why the Blackstone ratio does not simply translate to a standard of proof of 90%, as is often supposed. Thanks to an anonymous referee for pressing me on this point.
- ⁹ Some judicial systems understand "reasonable doubt" with reference to the "reasonable person" standard. This standard is itself fraught with philosophical and legal questions, which would benefit from the insights of applied epistemologists. See also Richard Lippke, "Punishing the Guilty, Not Punishing the Innocent" *Journal of Moral Philosophy* 7, 4 (2010): 462–88, at p. 475; Alex Stein, *Foundations of Evidence Law* (Oxford: Oxford University Press, 2005).
- ¹⁰ Lippke op. cit.; Laudan op. cit.; Laurence Tribe, "Trial By Mathematics: Precision and Ritual in the Legal Process" *Harvard Law Review* 84, 6 (1971): 1329–93. I do not argue for this claim here, but I strongly suspect that quantification of the standard is both unhelpful and untenable.
- ¹¹ Charles R. Nesson, "Reasonable Doubt and Permissive Inferences: The Value of Complexity" *Harvard Law Review* 92, 6 (1979): 1187–1225; Jon Newman, "Quantifying the Standard of Proof Beyond a Reasonable Doubt: A Comment on Three Comments" *Law, Probability, and Risk* 5 (2006): 267–269; Walen op. cit.; Sand and Rose op. cit.; *McCullough v. State*, 657 P.2d 1157, 1159 (Nev. 1983); Laudan 2011, op. cit.; Larry Laudan, *Truth, Error, and Criminal Law: An Essay in Legal Epistemology* (Cambridge: Cambridge University Press, 2006); Larry Laudan, "Is it Finally Time to Put 'Proof Beyond a Reasonable Doubt' Out to Pasture?" in Andrei Marmor (ed.), *Routledge Companion To Philosophy Of Law* (New York: Routledge, 2012), pp. 317–332; Tribe op. cit.; Lawrence Solan, "Refocusing the Burden of Proof in Criminal Cases" *Texas Law Review* 78 (1999): 105–47; James Franklin, "Case Comment—United States v. Copeland, 369 F. Supp. 2d 275 (E.D.N.Y. 2005): Quantification of the 'Proof Beyond Reasonable Doubt' Standard" *Law, Probability, and Risk* 5 (2006): 159–165; Peter Tillers and Jonathan Gottfried, "Case comment—United

-
- States v. Copeland, 369 F. Supp. 2d275 (E.D.N.Y. 2005): A Collateral Attack on the Legal Maxim That Proof Beyond A Reasonable Doubt Is Unquantifiable?” *Law, Probability, and Risk* 5 (2006): 135–157; Jack Weinstein and Ian Dewsbery “Comment on the Meaning of ‘Proof Beyond a Reasonable Doubt’” *Law, Probability and Risk* 5 (2006): 167–173.
- ¹² Solan op. cit.; Erik Lillquist, “Recasting Reasonable Doubt: Decision Theory and the Virtues of Variability” *U.C. Davis Law Review* 36 (2002): 85–179; C. M. A. McCauliff, “Burdens of Proof: Degrees of Belief, Quanta of Evidence, or Constitutional Guarantees?” *Vanderbilt Law Review* 35 (1982): 1293–1335. McCullough v. State; Rita James Simon and Linda Mahan, “Quantifying Burdens of Proof: A View from the Bench, the Jury, and the Classroom” *Law and Society Review* 5 (1971): 319–330; Irwin A. Horowitz and Laird C. Kirkpatrick, “A Concept in Search of a Definition: The Effects of Reasonable Doubt Instructions on Certainty of Guilt Standards and Jury Verdicts” *Law and Human Behavior* 20, 6, (1996): 655–670.
- ¹³ Lillquist op. cit.; Weinstein and Dewsberry op. cit.; Barbara J. Shapiro, “‘To a Moral Certainty’: Theories of Knowledge and Anglo-American Juries 1600–1850”, *Hastings Law Review* 38 (1986); Jon Newman, op. cit.
- ¹⁴ Laudan 2006, op. cit.; Laudan 2011, op. cit.; Laudan 2012, op. cit.; Walen op. cit.; Sand and Rose op. cit.; Trive op. cit.; Lippke op. cit.; Stein op. cit.; Antony Duff, Lindsay Farmer, Sandra Marshall and Victor Tadros, (eds) *The Trial on Trial* Vol. 3. (Oxford: Hart Publishing, 2007); Rinat Kitai, “Protecting the Guilty” *Buffalo Criminal Law Review* 6, 2 (2003): 1163–1187; Katherine Goldwasser, “Vindicating the Right to Trial by Jury and the Requirement of Proof Beyond a Reasonable Doubt: A Critique of the Conventional Wisdom About Excluding Defense Evidence” *Georgetown Law Journal* 86, 3 (1998); Ronald Dworkin, *A Matter of Principle* (Harvard: Harvard University Press, 1985); Jerome Frank and Barbara Frank, *Not Guilty* (New York: Doubleday, 1957).
- ¹⁵ Walen op. cit.; Lillquist op. cit.; Sand and Rose op. cit.; Laudan op. cit.; *Re B (A Child)* (2008) UKHL 35.
- ¹⁶ Many theorists discuss the role of consequences in determining or justifying a standard of proof. A small sample includes Sand and Rose op. cit.; Laudan 2006, op. cit.; Laudan 2011, op. cit.; Laudan 2012, op. cit.; Tribe op. cit.; Lillquist op. cit.; John Kaplan, “Decision Theory and the Factfinding Process” *Stanford Law Review* 20, 6 (1968): 1065–1092.
- ¹⁷ See, for example, Laudan 2006, op. cit.; Laudan 2011, op. cit.; Laudan 2012, op. cit.
- ¹⁸ Laudan 2011, op. cit., p. 209, emphasis in original.
- ¹⁹ Note too that whilst Laudan’s argument—and thus my response—focus on crime in the US and UK, the debate has global ramifications. If Laudan is correct about the US, this bears on the suitability of BARD for other nations. I return to the global ramifications of this debate in section nine.
- ²⁰ *Ibid.*, pp. 195–6.
- ²¹ As I emphasise in section seven, it only reduces the expected harm of non-incarcerated people. As I explain Laudan neglects crimes occurring within prison.
- ²² *Ibid.*, p. 199.
- ²³ *Loc. cit.*
- ²⁴ *Ibid.*, pp. 200–2.
- ²⁵ *Ibid.*, p. 202.
- ²⁶ *Ibid.*, p. 201.
- ²⁷ *Ibid.*, p. 202. On page 204 this figure of 36 crimes, of which seven are violent, increases to 40 crimes in addition to seven violent ones. As I explain, neither estimate is supported by Blumstein’s research.
- ²⁸ As I emphasise in sections six and seven, incarcerated individuals can and do commit crime. Violence within prisons is often not recorded in official crime statistics, and so becomes “hidden crime”.
- ²⁹ *Ibid.*, p. 202.
- ³⁰ *Loc. cit.*
- ³¹ *Ibid.*, p. 207. Laudan derives the “ten to twenty” figure through the statistical errors that I describe in section eight
- ³² Laudan 2011, op. cit. p. 207. The formula comes from Kaplan op. cit.
- ³³ Laudan 2011, op. cit. p. 207.
- ³⁴ Laudan 2012, op. cit.
- ³⁵ Thanks to Elisabeth Camp for helpful discussion here.
- ³⁶ Laudan 2011, op. cit. p. 199.
- ³⁷ Herbert Koppel, “Lifetime Likelihood of Victimization” (Bureau of Justice Statistics, 1987), p. 3.

-
- ³⁸ Jennifer Truman, “Criminal Victimization, 2010” (Bureau of Justice Statistics, 2011), p. 3; Jennifer Truman and Michael Planty, “Criminal Victimization, 2012” (Bureau of Justice Statistics, 2013); James Alan Fox and Marianne Zawitz, “Homicide Trends in the United States” (Bureau of Justice Statistics, 2002).
- ³⁹ Laudan 2011, op. cit., p. 199; Koppel op. cit., p. 2.
- ⁴⁰ Laudan loc. cit.; Koppel op. cit., p. 4.
- ⁴¹ For inclusion in Koppel’s statistic there need not be weapon, physical force, or success, *pace* Laudan.
- ⁴² Koppel op. cit. pp. 2–4.
- ⁴³ Laudan notes the figure of 83% excludes homicides. Adding homicides to the target statistic increases the likelihood by less than 0.3%. (Alexia Cooper and Erica Smith, “Homicide Trends in the United States, 1980–2008” (Bureau of Justice Statistics, 2011)). Homicide is a terrible, but rare, outcome. It is especially rare to be murdered as a first victimisation, furthermore, but these are the only homicides that increase the relevant risk.
- ⁴⁴ Koppel op. cit., p. 4. Lynch criticises this aspect of Koppel’s work. Lynch dismisses the 83% figure as “suspicious” and “implausible”, and condemns Koppel’s methods and data as inappropriate, arguing Koppel substantially overestimates lifetime risk of victimization. James Lynch, “A Review: An Evaluation of Lifetime Likelihood of Victimization” *Public Opinion Quarterly* 53, 2 (1989): 262–264.
- ⁴⁵ Laudan 2011, op. cit., p. 199.
- ⁴⁶ Loc. cit.
- ⁴⁷ Ibid. p. 224.
- ⁴⁸ Calculating the risk of false conviction in various nations is itself a delicate issue, and one I will not pursue here. Many policy decisions affect false conviction rates. These include kinds of evidence permitted at trial, access to defence lawyers, the role of eyewitness testimony etc. Michelle Alexander argues that long mandatory minimum sentences lead to widespread plea bargains, which increases false conviction rates. She notes that the accused are keen to severe mandatory minimums in the US, and so often plea bargain even if innocent. Corey Ayling, “Corroborating Confessions: An Empirical Analysis of Legal Safeguards Against False Confessions” *Wisconsin Law Review* (1984): 1121–1204; Michelle Alexander, *The New Jim Crow: Mass Incarceration in the Age of Colorblindness* (New York: The New Press, 2010); Jed S. Rakoff, “Why Innocent People Plead Guilty” *The New York Review of Books* 61, 18 (2014).
- ⁴⁹ Thomas Bonczar. “Prevalence of Imprisonment in the U.S. Population, 1974–2001” (Bureau of Justice Statistics, 2003); Ann Carson and Daniela Golinelli, “Prisoners in 2012—Advance Counts” (Bureau of Justice Statistics, 2013); David Brown, “The Limited Benefit of Prison in Controlling Crime” *Current Issues in Criminal Justice* 22, 1 (2010): 137–148.
- ⁵⁰ I am grateful to the Guest Editors for helpful discussion of these issues. D. Michael Risinger also emphasises some of these points. See D. Michael Risinger, “Tragic Consequences of Deadly Dilemmas: A Response to Allen and Laudan” *Seton Hall Law Review* 40, 3 (2010).
- ⁵¹ Thanks to an anonymous reviewer for emphasising this point.
- ⁵² Laudan 2011, op cit. p. 202.
- ⁵³ Ibid., 201–202.
- ⁵⁴ See, for example, Alfred Blumstein, Jacqueline Cohen, Jeffrey Roth, and Christy Visser, *Criminal Careers and “Career Criminals”* (National Academy Press, 1986), pp. 61ff, 94, 128, 205, 298, 306, 313, 320, 336.
- ⁵⁵ See, for example, *ibid.*, pp. ix, 1, 4, 18, 62, 352, 358.
- ⁵⁶ See, for example, *ibid.*, pp. 1, 11, 14, 24, 42, 128, 358.
- ⁵⁷ Laudan 2011, op cit. p. 201, n. 10.
- ⁵⁸ Blumstein op. cit. pp. 1–2. Indeed Blumstein’s thesis is that, given the skewed distribution, crime prevention policies must target the two categories—participants and active offenders—separately. It is crucial to not conflate these groups.
- ⁵⁹ Blumstein op. cit. pp. 1, 17, 31, 55.
- ⁶⁰ As far as I can tell, Blumstein does not render explicit his classification requirements. One must commit at least one violent crime within a timeframe, but (as far as I can see) Blumstein does not state the timeframe. This does not undermine *Blumstein’s* research, however, since Blumstein aims to illuminate career criminals and their behaviour. He wants to discern, for instance, whether active offenders take more drugs than mere participants. So Blumstein employs rates of criminal activity to classify offenders. But this lack undermines Laudan’s use of Blumstein’s work, since Laudan aims to determine average crime rates. For this project, the effect of the qualifying time period, as illustrated by the “runners” study, is vital.

-
- ⁶¹ Slight complications arise if either active offenders or mere participants are disproportionately likely to be arrested, plea bargain, or face trial. This effect, however, is likely to be small.
- ⁶² Laudan 2011, op. cit. p. 201, emphasis mine.
- ⁶³ Blumstein op. cit. pp. 4, 61ff, 94, 128, 205, 298, 306, 313; Todd Clear, “Backfire: When Incarceration Increases Crime” in Karen Fulbright (ed.) *The Unintended Consequences of Incarceration* (New York: Vera Institute, 1996), p. 5.
- ⁶⁴ To simplify the illustration, I set aside crime within prisons and crime caused by incarceration.
- ⁶⁵ Blumstein recognises mean averages do not accurately reflect crime distribution, and that most offenders committed crimes at rates far below the mean. See op. cit. pp. 62, 320, 336.
- ⁶⁶ Laudan 2011, op. cit. p. 202.
- ⁶⁷ Much of Blumstein’s research is on current inmates.
- ⁶⁸ Blumstein op. cit. pp. 5, 92–95.
- ⁶⁹ See, e.g. Blumstein op. cit. pp. 5, 18, 23, 92–94; Clear op. cit. p. 4. Index offences in the US are murder, non-negligent manslaughter, forcible rape, robbery, aggravated assault, burglary, larceny theft, motor vehicle theft, and arson.
- ⁷⁰ Laudan 2011, op. cit. p. 200–1.
- ⁷¹ Blumstein op. cit. p. 352; Clear op. cit. p. 5.
- ⁷² Martin Pritikin, “Is Prison Increasing Crime?” *Wisconsin Law Review* 2 (2008): 1049–1108, at p. 1055; Carl Taylor “Growing Up Behind Bars: Confinement, Youth Development, And Crime” in Karen Fulbright op. cit. pp. 53–4. Thanks to an anonymous reviewer for insights on these issues.
- ⁷³ Pritikin op. cit. p. 1060, Taylor op. cit. pp. 41ff.
- ⁷⁴ Pritikin op. cit. pp. 1056–8.
- ⁷⁵ Taylor op. cit. p. 41 ff; Pritikin op. cit. pp. 1057–60.
- ⁷⁶ James Gilligan, *Violence: Reflections on a National Epidemic* (New York: Vintage Books, 1997).
- ⁷⁷ Alexander op. cit. pp. 93–95; 142ff; Pritikin op. cit. pp. 1060–3; Blumstein op. cit.
- ⁷⁸ Pritikin *ibid.* pp. 1055–6; Karen Fulbright (ed.), *The Unintended Consequences of Incarceration* (New York: Vera Institute, 1996); Alexander op. cit.
- ⁷⁹ For studies of incarceration increasing individual crime rates, see Pritikin *ibid.*; Blumstein op. cit. 26 ff; Fulbright *ibid.*; Brown op. cit. pp. 140 ff; Gilligan op. cit.; and Todd Clear, Dina Rose, Elin Waring and Kristen Scully, “Coercive Mobility and Crime: A Preliminary Examination of Concentrated Incarceration and Social Disorganization” *Justice Quarterly* 20, 1 (2003): 33–64.
- ⁸⁰ Pritikin op. cit. pp. 1065–7, 1081; Clear op. cit. pp. 12ff; Blumstein op. cit. p. 44; Clear et al. op. cit.; John Hagan, “The Next Generation: Children of Prisoners” in Fulbright op. cit.
- ⁸¹ Pritikin op. cit. p. 1067 ff; Clear op. cit. pp. 7–13; Joan Moore, “Bearing the Burden: How Incarceration Policies Weaken Inner-City Communities” in Fulbright op. cit.; Clear et al. op. cit.; Brown op. cit.; Demetra Smith-Nightingale and Harold Watts, “Adding It Up: The Economic Impact of Incarceration on Individuals, Families, and Communities” in Fulbright op. cit.
- ⁸² Ekland-Olson, et al. dub this process “vacancy-chains”, and suggests it largely undermines the value of incarceration for some kinds of offence. See Sheldon Ekland-Olson, William Kelly, Hee-Jong Loo, Jeffrey Olbrich and Michael Eisenberg, *Justice Under Pressure: A Comparison of Recidivism Patterns Among Four Successive Parolee Cohorts* (New York: Springer-Verlag, 1993). See also Clear op. cit. p. 6; Blumstein op. cit. p. 27; Pritikin op. cit. p. 1072; Alexander op. cit. p. 126.
- ⁸³ Brown op. cit. p. 142; Clear et al. op. cit.; Pritikin op. cit. pp. 1064–5, 1072–3; Clear op. cit. pp. 9–12; Alexander op. cit. 236–7. Theorists discuss a tipping point after which higher incarceration rates *increase* aggregate crime. Clear suggests America has reached this point.
- ⁸⁴ The incarcerated represent approximately 1% of the US population, but substantially more than 1% of violent crimes occur within prison. Before the Bureau of Justice Statistics studies on prison crime began, Pritikin op. cit. writes, “The FBI’s Uniform Crime Reports (UCR) index figure [for the year 2000 suggests that] in-prison violent crime accounted for 3.6% of total reported violent crime [...] This estimate of the ratio of in-prison crime to total crime is conservative. It excludes nonviolent crime and it ignores the greater rate of underreporting for in-prison assaults than for violent crime generally”. As Pritikin fears, the real figure might well be much worse. In 2012 1.2 million violent crimes were reported to police outside prisons, whilst 5.8 million were self-reported by inmates to the Bureau of Justice Statistics.
- ⁸⁵ Thanks to an anonymous reviewer for pressing me to clarify this.

-
- ⁸⁶ The Bureau of Justice Statistics recently began to collect data on sex crime within prisons. According to the reports, 4.0% of prisoners and 9.5% juvenile detention prisoners report having been sexually abused within the last year, compared with approximately 0.1% of the non-incarcerated population. Often these abuses were by staff, and most victimisation is repeated. See Allen J. Beck, Marcus Berzofsky, Rachel Caspar, and Christopher Krebs, “Sexual Victimization in Prisons and Jails Reported by Inmates, 2011–2012” (Bureau of Justice Statistics, 2013). See also Gilligan op. cit., Pritikin op. cit., Clear op. cit., and Brown op. cit.
- ⁸⁷ Taylor op. cit. p. 50; Beck et al. op. cit.
- ⁸⁸ Laudan 2011, op. cit. p. 207; Laudan 2012, op. cit.
- ⁸⁹ Laudan 2011, loc. cit. Note that Laudan’s claim about the relationship between the standard of proof and relative harms stems from the errors described in section eight.
- ⁹⁰ There might, furthermore, be particular disvalue in government subjecting a citizen to increased risk of violence when its role ought be protective. Via incarceration the government can *trap* its citizens in a violent situation.
- ⁹¹ Laudan 2011, op. cit. p. 199.
- ⁹² These results are for all standards of proof except 0% and 100%. This illustration supposes a non-extreme distribution of credence levels among jurors. Almost any distribution suffices, except aberrant ones, such as all credences being very low.
- ⁹³ Ibid. p. 205.
- ⁹⁴ Ibid. p. 207.
- ⁹⁵ In America, for example, African American prisoners receive DNA-based exonerations at twice the rate of white prisoners, a ratio that far exceeds that of the relevant convictions. This provides evidence that African Americans are more likely to be falsely convicted than white Americans.
- ⁹⁶ Walen op. cit.
- ⁹⁷ The diminishing returns of incarceration arise because high-rate serious offenders tend to be imprisoned earlier. As society imprisons more people, they incarcerate increasingly lower-rate offenders for less serious offences. See Brown op. cit. p. 142. Recall the 99:1 population described in section five. The high-rate offender is likely to be incarcerated early, and all remaining offenders are low-rate.
- ⁹⁸ Laudan 2011, op. cit. p. 206.
- ⁹⁹ See, for example, Richard Samuels and Stephen Stich, “Rationality and Psychology” in Alfred Mele & Piers Rawling (eds.), *Oxford Handbook of Rationality* (Oxford University Press, 2004): 279–300.
- ¹⁰⁰ It bears noting there are few standards available in the vicinity of reasonable doubt, which is further evidence it is special and distinctive. Laudan attempts to list alternative qualitative standards (Laudan 2011, op. cit. p. 216). It is instructive that almost all his suggestions are *not* alternative standards to BARD; they are alternative regulations about evidence and convictions. The only genuine alternative standard Laudan lists is “beyond all doubt (thereby permitting irrational doubt to trump a conviction)” (loc. cit). This lack indicates there are not many alternative qualitative standards available. Some other potential standards include beyond-all-but-paranoid-doubt, beyond-all-but-conspiracy-laden-doubt, being sure, and certainty. But perhaps these are only two distinct standards, not four.
- ¹⁰¹ A version of this problem arises for any differing kinds of crime, such as whether vehicle theft should have a different standard of proof from house burglary, for example.
- ¹⁰² A further problem for Laudan’s project concerns the role of plea bargains in criminal justice. Plea bargains are extremely common. In the US approximately 97% of federal cases are resolved through plea bargains, and the rate for state felony cases is similar. This has several implications for Laudan’s project. Firstly, it is a weakness of Laudan’s formula that it covers a tiny fraction of cases in the US; for most criminal cases, the standard of proof plays at most an indirect role in influencing the outcome. Secondly, reducing the standard of proof would likely alter the plea bargain rate, which influences the formula Laudan ought endorse. These considerations play no role in Laudan’s calculations. A lower standard of proof is likely to increase the plea bargain rate, because trials become more likely to result in conviction. This is a cost, since during plea bargains in many jurisdictions the defendant has too little power relative to the prosecutor, lacks appropriate access to legal counsel, and is perhaps more likely to falsely plead guilty. For discussion, see Alexander op. cit. pp. 87–9; Rakoff op. cit.
- ¹⁰³ Many thanks to Alex Anthony, Bob Beddor, David Black, Elisabeth Camp, Will Fleisher, Jon Garthoff, E.J. Green, Doug Husak, David Rose, Mary Salvaggio, Ernie Sosa, Una Stojnić, Nick Tourville, Alec

Walen, the Guest Editors, and two anonymous referees for comments on an earlier draft of this paper.
Thanks also to Greg Arutiunov, Beth Henzel, and Alexander Leone for helpful discussion of these issues.