The reasonable doubt standard as inference to the best explanation

Jellema, Hylke

The reasonable doubt standard as inference to the best explanation

Hylke Jellema[1]

Abstract

Explanationist accounts of rational legal proof view trials as a competition between explanations. Such accounts are often criticized for being underdeveloped. One question in need of further attention is when guilt is proven beyond a reasonable doubt (BARD) in criminal trials. This article defends an inference to the best explanation (IBE)-based approach on which guilt is only established BARD if (1) the best guilt explanation in a case is substantially more plausible than any innocence explanation, and (2) there is no good reason to presume that we have overlooked evidence or alternative explanations that could realistically have exonerated the defendant. This is a comparative account, which I argue is better suited for arriving at accurate verdicts than the non-comparative 'no plausible alternative' account that many explanationists tacitly assume. Furthermore, this account is not susceptible to the most important arguments against IBE in criminal trials or to arguments against other, non-explanationist interpretations of the BARD standard. I use a case study to illustrate how this account provides meaningful guidance for decision makers in criminal trials.

1 Introduction

In criminal trials, fact finders (judges or juries) often face a choice between competing scenarios. For instance, in this article you will read about a real case where a woman was killed in a forest. The killer was most likely either the woman's husband or a madman who jumped out of the bushes. Imagine that you are in the judge's robe and have to choose between these scenarios—'her husband killed the victim' and 'an unknown madman killed the victim'. You ask yourself which of these options best

---

✉ Hylke Jellema
h.jellema@rug.nl

1   University of Groningen, Groningen, The Netherlands

explains the evidence in the case. Now assume that the husband killing her is the best available explanation. Should you then convict him because it is proven that he is the culprit? To put it differently, if fact finders have to choose between several scenarios—some better than others—what determines which they can accept as proven (if any)?

This article focuses on a particular account of rational criminal proof, namely the so-called *explanationist* or *explanation-based* approach. Explanationists view criminal proof as a competition between stories, scenarios or explanations of what happened (I'll be using these terms interchangeably). These approaches have been influential over the past decades, most prominently through Pennington and Hastie's (1991) story model and the work of Pardo and Allen (2008).

However, despite their influence, some criticize explanationist approaches for not being worked out in sufficient detail. One particular question that has received insufficient attention within this framework is when the prosecution has succeeded in proving a scenario that implies the defendant's guilt. The answer to this depends on the criminal proof standard. Common law systems use the *beyond a reasonable doubt* (BARD) standard, which is the standard that I will focus on here.

The BARD standard is notoriously vague. For instance, Larry Laudan calls it "obscure, incoherent, and muddled" (Laudan 2006, pp. 295–296). Though various interpretations of this standard exist, they are all problematic according to Laudan (2006, pp. 35–47). He claims that these interpretations seem to create more confusion and potential for error than they prevent. In particular, he argues that they are too unclear to be meaningfully applied in real cases, too reliant on the subjective hunches of the person who applies it or insufficiently connected to the goal of avoiding false convictions (Laudan 2006, p. 87). My aim in this article is not just to propose an interpretation that fits well within the explanationist framework, but also to argue that this interpretation overcomes Laudan's criticisms of other, non-explanationist interpretations. The interpretation that I propose is as follows:

Guilt is only established BARD if (1) the best guilt explanation in a case is substantially more plausible than any innocence explanation, and (2) there is no good reason to presume that we have overlooked evidence or alternative explanations that could realistically have exonerated the defendant.

My interpretation is an expanded version of Thagard's (2003, 2008) proposal that we could interpret the BARD standard as a form of *inference to the best explanation* (IBE). This is a comparative account of the BARD standard: we determine whether guilt is proven by looking at how well our guilt explanation performs compared to competing innocence explanations.

In contrast with this approach, those explanationists who have briefly addressed the question of how to interpret the BARD standard, tacitly adopt a non-comparative interpretation. This is the *no plausible alternative* (NPA) account. According to NPA, proving guilt BARD requires that we have a sufficiently plausible explanation implying guilt and no sufficiently plausible explanation implying innocence (e.g. Pardo and Allen 2008).

As I will argue in Sect. 3, such a non-comparative interpretation is undesirable on epistemic grounds.[1] In particular, a defendant should only be convicted if the probability that he is guilty is sufficiently high, in order to avoid false convictions. Yet in contrast with the competing probabilistic framework of rational criminal proof, explanationists argue that fact finders often cannot meaningfully attach probabilities to particular scenarios. What they *can* reason meaningfully about, however, is whether an explanation is *plausible*. Plausibility is a qualitative notion that effectively expresses how well an explanation explains the facts in a case. Explanationists assume that the more plausible an explanation is, the more probable it is as well. In other words, plausibility tracks probability. However, as I will argue, in a non-comparative account such as NPA, plausibility is not always a good guide to probability. The reason for this is that the relative performance of explanations influences their probability. So, we need a comparative account that captures the comparative dimension.

Many object to a comparative approach to fact finding in criminal trials. According to such critics, criminal law is about whether the case for guilt is good enough, not whether it is better than the case for innocence. However, as I will argue in Sect. 4, IBE does not imply that we should *only* care about the relative performance of explanations. First, we can look at the absolute performance of explanations to efficiently reach comparative conclusions. Second, and more importantly, IBE does not imply that we can convict a defendant based on an implausible guilt explanation, even when that explanation is better than any innocence explanation. If all our explanations are implausible then we are in a so-called *bad lot* situation. As I will argue, in bad lot situations we have reason to presume that a better investigation could have exonerated the defendant. The comparative case for guilt then lacks *robustness* and we should therefore not convict the defendant. That robustness is a requirement for proof of guilt BARD is a plausible assumption for *any* account of the BARD standard. So, the proposed IBE-based account overcomes the bad lot problem without the need for *ad hoc* modifications.

To illustrate how my account can be applied to criminal practice I will use the Dutch *Simonshaven* case as a running example. I introduce this case in Sect. 2. I also discuss explanationism, NPA, and IBE in more detail in that section. In Sect. 3 I discuss why explanationists need a comparative account of the BARD standard, which NPA does not provide. Finally, in Sect. 4, I argue that my comparative account nonetheless squares with the non-comparative nature of criminal law.

---

[1] In his article within this special issue Roberts (2020) argues against such 'epistemic' interpretations of proof rules. He instead advocates interpreting proof rules, and in particular the presumption of innocence, by looking at how they are actually used within specific legal systems, in a way that is sensitive to the history and institutional details of those systems. I agree to a large extent with what Roberts writes. However, the BARD standard is often seen as one of the most confusing proof rules, which courts make sense of in various, and often highly unclear ways (Laudan 2003; Shealy 2012). Therefore, a jurisprudential approach to interpreting it may not be the best way to go. Instead, an epistemic interpretation, such as the one given here, can clarify just why such a standard exists, where its normative status comes from and how we can apply it to real cases.

2 Preliminaries: the Simonshaven case and explanatory choice

In this section I set the stage for my argument. First, I discuss the 'Simonshaven' case, which I will use as an example throughout my article.[2] After that I introduce my theoretical framework, explanationism and the two explanationist interpretations of the BARD standard, IBE and NPA.

2.1 A murder in a forest

The Simonshaven case took place in the Netherlands in 2009. The case's name comes from a village near the forest where it took place.[3] I will provide more details of the case throughout this article, where needed to illustrate my ideas.[4] The simplified case is as follows:

> A couple has recently stopped living together due to relationship issues. Nonetheless, they sometimes take walks together in a nearby forest. One day, witnesses see them walk into this forest. About an hour later the husband calls his son in a panicked and confused state. The son races to the forest where he finds his dad who says that someone attacked them. The son then calls the police who find the wife's body lying on a nearby path in the forest in a pool of blood. There are severe wounds on her face. The husband is also injured. He has minor facial injuries. When questioned, the husband claims that an attacker suddenly jumped out of the bushes and knocked him out. He claims to remember only vague details about the attacker, such as that he was light skinned and of average height. After he is marked as a suspect by the police, he calls upon his right to remain silent and makes no further statements.
>
> Forensic investigation shows that a gun was fired near the victim's face and that her wounds were caused by being hit with a blunt object, possibly a gun. The police also find shell casings near her body. One of these casings contains DNA-evidence that belongs to neither the victim nor the suspect. However, despite intensive searching, no gun is found near the crime-scene. During the trial the prosecution argues that the suspect could have dumped the gun in a nearby river.
>
> The children of the defendant testify that he was jealous because the victim had recently started dating a new lover. They also testify that their father owned several guns. Apart from the victim's husband, no other suspect is identified.

The fact finder (in this case a judge) is now faced with the question which of the following competing explanations to believe (if any):

---

[2] This case is also discussed in the special issue of *Topics on cognitive science*, titled 'Models of Rational Proof in Criminal Law'.

[3] The Hague Appellate Court, 18th February 2015, ECLI:NL:GHDHA:2015:282. In the Netherlands fact finding is done by judges, rather than juries. Furthermore, the Dutch legal system lacks the BARD standard. It has the *conviction raisonnée* standard, which stipulates that conviction is permitted only if the fact finder is (reasonably) convinced of the defendant's guilt. The account that I propose in this article can also be extended to this standard because it can be read as an account of justified belief in the guilt of a defendant.

[4] My aim is not to draw conclusions about the actual case. The above is only a summary that focuses on the most important items of evidence.

The husband explanation: The suspect walked into the forest with the victim. They either got into an argument and the suspect killed the victim in anger, or the suspect had planned to kill her there. The suspect fired a gun in the vicinity of the victim and then beat her to death using his handgun. He then drove to a nearby river to dump his gun after which he drove back and called his son and claimed someone had attacked them.

The madman explanation: The suspect walked into the forest with the victim. During the walk an unknown perpetrator randomly attacked them, first knocking the husband unconscious. The unknown man then fired a gun in the vicinity of the victim. Then he beat her to death using the same gun and fled.

2.2 Explanationism

The judge in the Simonshaven case faced a choice between the competing 'husband' and 'madman' scenarios. The *explanationist* view of legal fact finding emphasizes the importance of such competition for criminal trials. On this view, fact finders reason abductively, by asking how well each of the competing scenarios explains the evidence in the case and then choosing one (or none).

Various authors defend an explanationist account, though not all of them define themselves as such. For my current purposes, I identify explanationists as those authors who understand either the practice (Allen 2010; Pennington and Hastie 1991) or the rationality (Josephson 2000; Pardo and Allen 2008; Bex 2011; van Koppen 2011; Wagenaar et al. 1993) of criminal fact finding as a competition between explanations. Pardo and Allen are a good example of this. They describe the core of their approach as "explanation as a guide to inference" (Pardo and Allen 2008, pp. 229–33). The task of the fact finder, therefore, revolves around reasoning about—and with—explanations. In criminal trials such explanations typically take the form of 'scenarios' or 'stories'. These are essentially time lines—series of causally connected events—that explain the evidence in the case. For instance, the 'madman' and 'husband' scenario both explain how the victim was killed and other evidence such as the husband's wounds. Empirical research suggests that when deciding on a case, fact finders commonly construct such stories and compare them to see which they find most believable. These stories usually contain a central action (such as a killing) and a context that makes this action understandable (Pennington and Hastie 1991, 1992).[5]

Some explanations imply that the defendant committed the alleged crime. If such *guilt explanations* are proven, this can lead to conviction. Other explanations are consistent with the innocence of the defendant (I call these *innocence explanations*). For instance, the 'madman' explanation implies the innocence of the defendant and, if proven, should lead to his acquittal.

Whether an explanation can be accepted as proven depends on the *plausibility* of that explanation and that of its competitors (Pardo and Allen 2008; Allen 2010; Josephson

---

[5] Scenarios should ideally be specific. However, they can also be more general. For instance, in criminal trials the prosecution sometimes try to prove disjunctive explanations, which are composed of multiple, possibly inconsistent possibilities. An example of this is the explanation: 'the defendant either strangled or punched the victim to death'.

2000). A plausible explanation is one that explains the evidence (sufficiently) well. This term is usually further specified as a set of qualitative criteria, with different authors suggesting different criteria. For instance, in the story model of legal proof, how good an explanation is depends on how much of the evidence the explanation accounts for, whether it contains internal contradictions and problematic gaps, whether it has all its parts and how well it fits with our knowledge of the world (Pennington and Hastie 1992, pp. 190–199; Bex 2011, pp. 91–92). Pardo and Allen (2008, p. 230) propose that "all other things being equal [an explanation is] better to the extent that it is consistent, simpler, explains more and different types of facts (consilience), better accords with background beliefs (coherence), is less ad hoc, and so on; and is worse to the extent that it betrays these criteria." Josephson (2000) and Amaya (2009) propose still other criteria.

I do not want to delve too deeply into the debate about which set of criteria best captures plausibility. Instead I will simply offer a working definition, which is an attempt at summarizing the above proposals: *An explanation is more plausible to the extent that it: (a) explains more of the evidence in the case, (b) fits better with the fact finder's background beliefs about how the world usually works, and (c) has parts that fit together coherently.*

So, when we want to determine whether an explanation is plausible, we ask questions such as how well this particular explanation fits with our background beliefs and how much evidence in the case it explains. One thing that is important to note for my argument in the rest of this article is that when we discover, for instance, that a better explanation exists, our formerly best explanation does not become any less plausible. After all, it does not suddenly explain fewer facts or become more internally incoherent.

Plausibility is a qualitative, non-mathematical concept[6] that is distinct from the quantitative notion of *probability*. Nonetheless, explanationists assume a close connection between the two concepts. In particular, an underlying assumption of this framework is that plausibility is a *guide* to the probability of explanations. For instance, as Pardo and Allen (2008, footnote 45) put it "[Explanationism] is, at root, based on the notion that explanatory success tracks likelihood of truth—the better the explanation, the more likely true." In other words, the more plausible an explanation is, there more probable it *ceteris paribus* is as well.

Though explanationists are often somewhat vague about how their theories relate to Bayesian probability theory, we can read their remarks as an expression of the *heuristic conception* of the link between plausibility and probability. On this account explanatory reasoning approximates correct Bayesian reasoning (Okasha 2000; McGrew 2003; Lipton 2003). Yet Bayesian probabilistic reasoning is often not feasible in practical situations, according to proponents of this approach. Instead, reasoning about specific, competing explanations and their plausibility is more suitable for human beings (including judges and juries).

There is a great deal of support for the idea that fact finders who need to make sense of complex sets of evidence do so using explanatory reasoning. This can most clearly be

---

[6] This is similar to how philosophers of science such as Lipton (2003) and Thagard (1978) approach the notion of 'best' in inference to the *best* explanation. In contrast, authors such as Douven (2017) and Schupbach (2016) attempt to define such terms in a mathematically precise way.

seen in the research of Pennington and Hastie (1991). More generally, summarizing existing psychological research, Lombrozo (2016, p. 749) writes that when children and adults generate and evaluate explanations, they recruit explanatory virtues (…) as evaluative constraints on reasoning. Finally, there is ample evidence that ordinary people are quite poor at probabilistic reasoning (Tversky and Kahneman 1983).

So, explanationists agree with probabilistic account of the BARD standard that guilt can only be proven if the probability that the defendant is guilty is sufficiently high (e.g. 95and). As Allen and Pardo (2017, p. 1580) put it: "The explanatory account shares the same ends or goals as probabilistic approaches, which have to do with various policy judgments about the likelihood of disputed facts and allocating the risk of error between the parties." Yet whether the probability of guilt is sufficiently high is often not something that fact finders can meaningfully determine by means of probabilistic reasoning (Allen and Pardo 2007; Laudan 2006, pp. 44–47).

The heuristic view underlying explanationism has not been worked out with much detail. Nonetheless, the probabilistic goal and the notion of plausibility are clear enough that we meaningfully can ask how explanationists should interpret the BARD standard to meet this goal. I now want to compare two candidates, IBE and NPA, and argue for the former.

2.3 Two explanationist interpretations of the BARD standard

The explanationist framework instructs fact finders to look at the competing explanations and to choose between them based on how plausible they are. Choosing an explanation means accepting that explanation as proven. The following scheme expresses this idea (adapted from Laudan 2007):

(1) $e_1, e_2, \ldots, e_n$ are the salient facts to be explained.
(2) Hypotheses $h_1, h_2, \ldots, h_n$ each explain $e_1, e_2, \ldots, e_n$.
(3) Rival explanations for $e_1, e_2, \ldots, e_n$ have been earnestly sought out but the search has produced only $h_1, h_2, \ldots, h_n$.
(4) *[Proof standard]*
(5) Therefore, h_i is probably true.

I have left premise 4—the proof standard—unspecified. Such a proof standard determines on what grounds the fact finder should choose between the available explanations. For criminal cases the proof standard tells fact finders when they can accept a guilt explanation as proven.

A key goal of proof standards is achieving a just *error distribution*. The judicial system will inevitably produce errors from time to time. The higher the standard for conviction is the more errors will be false acquittals and the fewer will be false convictions. In criminal law, false convictions are generally considered to be much worse than false acquittals (Epps 2014). This is captured in Blackstone's well-known maxim that "it is better that ten guilty persons escape than that one innocent suffer" (Blackstone 1962). This 10-to-1 ratio should not be taken too literally; various other ratios have been postulated. What matters is the underlying intuition that a good criminal proof standard imposes a high barrier before conviction is justified.

Explanationists have paid little attention to how the BARD standard should be interpreted. In the literature we find two (brief) proposals: Thagard (2003, 2008) has proposed a version of the BARD standard based on IBE. However, Pardo and Allen (2008) claim that *all* proof standards within common law systems (such as the preponderance of evidence standard), *except* BARD, are IBE. For BARD they propose the NPA standard (Pardo and Allen 2008; Allen and Pardo 2019).

2.3.1 Inference to the best explanation

The concept of IBE originated in the philosophy of science and epistemology (Peirce 1965; Harman 1965; Thagard 1978; Lipton 2003). On IBE we infer the probable truth of an explanation on the grounds that it would, if true, provide a better explanation of ones evidence than any other available alternative explanation. It therefore leads to the following criterion for accepting an explanation as proven:

(4) h_i is the best explanation in the set $\{h_1, h_2, \ldots, h_n\}$.

IBE has also been applied to criminal law. Thagard has proposed—but not argued—that the BARD standard could be IBE (Thagard 2003, p. 366; 2008, p. 141). However, his version of IBE requires more than simply choosing the best available explanation. Rather, he proposes that the BARD standard is only met when the best guilt explanation is "substantially more plausible" than any innocence explanation. An explanation being only slightly better than any of the alternatives fails to yield sufficient certainty for conviction. If, for instance, both the prosecution and defense offer explanations that are roughly equally good, their explanations will lack uniqueness, and great uncertainty will result (Pennington and Hastie 1992, p. 191). This would conflict with the goal of a just error distribution mentioned earlier. We want a high degree of certainty before conviction. That is why the guilt explanation must be 'substantially more plausible'. So, what IBE asks the fact-finder to do is to make a judgment about the relative plausibility of the various explanations (I call this a *relative judgment*). On this view, proof of guilt BARD requires the best guilt explanation to be significantly more plausible relative to the best innocence explanation.

2.3.2 No plausible alternative

NPA originates with Pardo and Allen, who explicitly contrast it with IBE:

> In criminal cases, rather than inferring the best explanation from the potential ones, fact finders [should infer] the defendants innocence whenever there is a sufficiently plausible explanation of the evidence consistent with innocence (and ought to convict when there is no plausible explanation consistent with innocence) (Pardo and Allen 2008, pp. 238–239).

Josephson ([2000](#), p. 1642)[7], Allen ([2010](#)) and Allen and Leiter ([2001](#), pp. 1527–1528)[8] give equivalent formulations of the BARD standard.

On NPA a guilt explanation is proven BARD when:

(4) h_i is the only sufficiently plausible explanation in the set $\{h_1, h_2, \ldots, h_n\}$.[9]

How should we interpret this criterion and how does it differ from IBE? As I mentioned above, IBE depends on a relative judgment of the plausibility of explanations. Although Allen and Pardo do not state explicitly how NPA should be interpreted, I follow the same interpretation that others, such as Sullivan ([2018](#)) and Kolflaath ([2019](#)) give, namely that NPA is instead about judging explanations based on their *absolute plausibility*. This means judging explanations on an absolute scale, independently of one another. As Kolflaath ([2019](#), pp. 5–6) puts it:

> [NPA means] that fact-finders should decide first whether there is a plausible explanation consistent with guilt, and then—if they take this to be the case—decide whether there is also a plausible explanation consistent with innocence. Thus, [NPA seems] to presuppose that the first decision can be made independently of the second. And although the need for the second decision depends upon the first decision, whether or not there is a plausible explanation consistent with innocence would not, as far as I can tell, depend upon whether there is a plausible explanation consistent with guilt.

In other words, fact finders have to check whether explanations individually meet some standard of plausibility, rather than how well they perform relative to one another. When there is a good reason to believe that an explanation does not perform sufficiently well on its own—i.e. is insufficiently plausible in an absolute sense of the word—it is eliminated. I will call reasoning about how well an explanation performs on its own *absolute judgment*.[10]

So, while NPA is comparative in the sense that it instructs fact finders to consider and compare the possible alternative explanations on each side, it is *not* comparative in the sense that it takes into account the relative plausibility of explanations (Sullivan [2018](#), p. 5).[11]

---

[7] "Guilt has been established beyond a reasonable doubt when there is no plausible alternative explanation for the data that does not imply the guilt of the defendant."

[8] "[I]n criminal cases the prosecution must provide a plausible account of guilt and show that there is no plausible account of innocence."

[9] It is similar to the *no plausible alternative theory of knowledge*, which states that to know proposition P, we need to rule out all relevant alternatives to P (Lawlor [2013](#)).

[10] These two types of judgments—relative and absolute—are equivalent to two modes of reasoning distinguished in cognitive psychology. 'Selective processing' where something is judged in isolation, whereas 'comparative processing' means judging something in the context of the available alternatives (e.g. Sanbonmatsu et al. [1998](#), [2011](#); Trope and Liberman [1996](#)). On NPA, the judgments about whether there is a 'sufficiently' plausible' guilt and whether there is a sufficiently plausible innocence explanation can be made in isolation, which therefore corresponds to selective processing.

[11] Though Allen and Pardo's call their overall account the *relative plausibility theory* (Allen and Pardo [2019](#)), their definition of the BARD standard is not about relative plausibility. As Allen and Pardo ([2019](#), p. 31) themselves remark: "[The BARD standard is] "comparative" in a different sense, however. It is comparative in the sense that determining whether the threshold is met will involve considering and comparing the possible alternative explanations on each side."

Let me give a toy example to clarify. One aspect that makes an explanation plausible is how much of the evidence it explains. Suppose that an explanation explains 7 items of evidence. This is a conclusion about how well the explanation performs on its own—i.e. how plausible it is in an absolute sense. In contrast, a conclusion about relative plausibility could, for instance, be 'explanation A explains 2 more pieces of evidence than explanation B.' Within this toy example, if we employ IBE we would look at how many more pieces of evidence explanation A explains than explanation B when we decide which explanation to accept. When we employ NPA, we first look at whether explanation A explains a 'sufficient' amount of evidence, say 5, and then we look whether explanation B explains a 'sufficient amount of evidence'.

One important thing to note is that an absolute judgment can be turned in to a relative judgment. For instance, based on the fact that explanation A explains 7 pieces of evidence and explanation B 5, we can reach the relative conclusion that A explains two more pieces of evidence than B. As I will argue in Sect. 4.1, this is one of the reasons why absolute plausibility still plays a key role within IBE—because it is a useful heuristic for reaching relative conclusions. However, a relative judgment about plausibility cannot be turned into an absolute one. For instance, in this toy example we cannot infer that A explains at least 5 pieces of evidence from the fact that it explains 2 pieces of evidence more than B (unless we know how much B explains). This in turn means that because NPA is about how explanations perform in an absolute sense, judgments about relative plausibility serve no function within the framework. In other words, a fact finder never has any reason for asking how much better one explanation performs relative to the other if she is after conclusions on an absolute scale.

However, this makes NPA problematic. Both relative and absolute judgment are important for accurate fact finding. In the next two sections I deal with each form of reasoning in turn. I begin with relative judgment and argue that NPA cannot account for this without losing what makes it distinctive (Sect. 3). After that I show how IBE can account for the value of absolute judgments, while still being about relative conclusions at heart (Sect. 4).

3 The importance of relative plausibility

On IBE proof of guilt BARD is about relative judgments of plausibility, whereas on NPA it is about absolute judgment. As I will argue in this section, relative judgment is (at least sometimes) necessary for accurate conclusions. In particular, sometimes the *plausibility* of one explanation going up can mean the *probability* of a competing explanation going down. The goal of the BARD standard is ultimately probabilistic; fact finders should only accept a guilt explanation as proven if it is sufficiently probable. So, in order for plausibility to adequately track probability we should also take into account the relative plausibility of explanations. NPA fails to account for this (or, if it does, collapses into IBE). However, before looking at this argument, let me first say more about how explanations compete both in terms of probability and plausibility.

3.1 Explanatory competition

The *relative plausibility* of explanations is important to accurately assess their probability when these explanations compete. What do we mean by competing explanations?[12] One sense in which, for instance, the 'husband' and the 'madman' explanation compete is that they attempt to explain the same facts (e.g. the victim's dead body). However, not all explanations that explain the same facts compete. Events can have multiple causes, which do not have to conflict in any way. As I use the term here, two explanations—A and B—compete when the following condition is met: if evidence raises the probability of explanation A, the probability of explanation B is thereby lowered. In other words, the confirmation of explanation A implies the disconfirmation of explanation B. This is also called *explaining away*: when an event has multiple possible causes, and the confirmation of one cause reduces the need to invoke alternative causes (Kjaerulff and Madsen 2008). The probabilistics of explaining away can best be shown in a Bayesian network. For instance, Hepler et al. (2007) give the following idiom[13]:
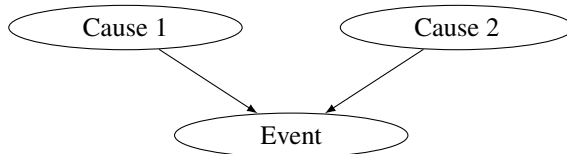

Without going into too much detail about how Bayesian networks work, this idiom can be understood as follows: The top nodes are potential causes of the event in the bottom node. For instance, in Hepler et al.'s (2007) example, the event is 'the defendant confesses to the crime', and the potential causes are 'the defendant is guilty' and 'the defendant was coerced by an interrogating official.' However, the causes could also be 'madman' and 'husband' and the event 'dead body'. What matters is that there is a path between the two causes through the event that they explain. This expresses the explaining away effect that I just mentioned, where the probability of one explanation influences that of its competitors.[14]

In such a network, evidence can change the probability of an explanation in two ways. First, some evidence directly confirms or disconfirms an explanation. Second, evidence can (dis)confirm an explanation by (dis)confirming one of its competitors. For instance, take the piece of evidence from the Simonshaven case that the victim had a new lover, which implies that the defendant had a motive for killing her. We can include this evidence in the Bayesian network as follows:

---

[12] See Schupbach and Glass (2017) for an extensive discussion on this topic.

[13] Modeling mutually exclusive hypotheses requires a more sophisticated approach (Fenton et al. 2013). However, the general idea remains the same.

[14] Assuming we have some evidence that the event occurred. For instance, that we know that there is indeed a dead body.

Motive

Husband   Madman

Dead body
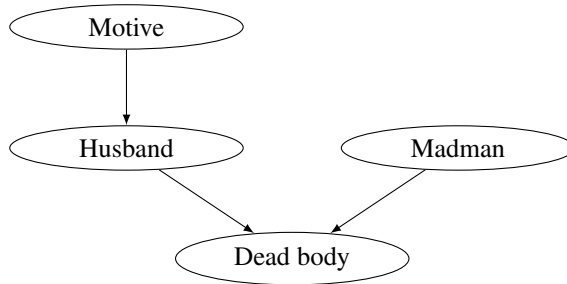
There is a direct probabilistic link between the 'husband' explanation and the newly introduced evidence. In this case it means that the piece of evidence increases the probability of the explanation that the husband was the killer. This piece of evidence directly confirms the explanation. Yet because the two explanations compete, this piece of evidence also indirectly disconfirms the 'madman' explanation.

What does all this have to do with the plausibility of explanations? Criteria of plausibility—such as how many pieces of evidence an explanation explains—are intended to be a guide to the probability of these explanations. However, the two ways in which evidence can *probabilistically* (dis)confirm an explanation are reflected differently in terms of their *plausibility*. Evidence that directly confirms an explanation makes this explanation more plausible in an absolute sense. For example, when we introduce the evidence that the defendant had a motive for the killing, the 'husband' explanation now explains one more piece of evidence. It becomes more plausible as a result, regardless of what this evidence does to its competitors. Conversely, evidence that is indirectly relevant influences the *relative* plausibility of explanations. It disconfirms one explanation probabilistically, by directly supporting a competing explanation. For example, the motive evidence makes the 'madman' explanation less probable, by making the 'husband' explanation more plausible.

Suppose that we look at the competing explanations independently from one another. For instance, we first look at the madman' explanation, and try to determine whether it is plausible. The question whether the husband had a motive does not enter into this equation. To see why, imagine that we knew with certainty that the husband did not kill his wife. Now suppose that we learned that the defendant had a motive for murder. Since we know for sure that he did not kill his wife, this does not give us any reason to disbelieve the explanation that a madman killed her. Conversely, if we did *not* know with certainty that the 'husband' explanation is false, this evidence *would* be relevant. After all, it makes the husband explanation more plausible. So, the relevance of evidence about motive to the 'madman' explanation can only be determined by looking at both explanations in relation to one another—i.e. to their relative plausibility.

3.2 NPA, IBE and relative plausibility

The above is relevant for the discussion between IBE and NPA because a great difference in relative *plausibility* can mean that the two principles lead to different recommendations about whether the fact finder should convict or acquit due to the

resulting *probabilities* of these explanations. In particular, in some cases we have a highly plausible guilt explanation that explains a great deal of evidence, making it much more plausible than any innocence explanation. IBE then dictates conviction. However, if there is also an innocence explanation that is plausible when looked at on its own, then NPA dictates acquittal.

Yet in such cases, the evidence for the guilt explanation might indirectly disconfirm the innocence explanations to such a great degree that they become highly improbable. For instance, if we believe that the defendant in the Simonshaven case had motive, means and opportunity, then this supports the explanation that he killed his wife. It thereby disconfirms the competing 'madman' explanation. The more plausible we make the 'husband' explanation, the less probable the 'madman' explanation becomes. Add enough evidence supporting the 'husband' explanation and it (indirectly) disconfirms the 'madman' explanation to such a great extent that no reasonable fact finder could believe it. For example, one piece of evidence that I have not mentioned is that multiple witnesses testified that the defendant had threatened the victim several times in the past. First, the son of the defendant said that the defendant had repeatedly told the victim "If anything happens I'm taking you with me". Second, the brother of the victim reported that the defendant told him "if you leave me, I'll kill you." We can imagine even more of such evidence for the 'husband explanation' being found, which would increase the plausibility of the 'husband' explanation, thereby lowering the probability of the 'madman' explanation.

So, whether the defendant can be convicted—at least sometimes—depends on the relative plausibility of explanations. Fact finders should therefore be attentive to the relative plausibility of explanations, especially when the difference in plausibility is large. However, NPA does not take relative plausibility into account. Instead it directs fact finders to look at how well explanations perform on their own. It therefore ignores the probabilistic influence of how explanations perform relative to one another. Because we only want to accept an explanation as proven if it is sufficiently probable we should *ceteris paribus* prefer IBE's conclusion over that of NPA, where the two differ. After all, IBE does pay attention to the relative plausibility of explanations.

One move that proponents of NPA can make in response to this, is to suggest that their account should be interpreted differently; that fact finders should look at both absolute and relative plausibility. This is, for instance, what Josephson (2000, p. 1626) proposes. He puts forward the NPA definition of BARD[15] but also says that what makes an explanation plausible depends both on how decisively the leading hypothesis surpasses the alternatives and how well it stands by itself, independently of the alternatives. Presumably, on such an approach, fact finders should convict when the best guilt explanation is substantially more plausible than the best innocence explanation, even if the best innocence explanation is plausible when looked at on its own.

I have no issue with such an interpretation, because NPA would boil down to IBE.[16] NPA would then instruct fact finders to primarily look at the absolute plausibility of explanations, but also that relative plausibility can trump considerations of absolute

---

[15] "Guilt has been established beyond a reasonable doubt when there is no plausible alternative explanation for the data that does not imply the guilt of the defendant" (Josephson 2000, p. 1642).

[16] Indeed, despite offering the NPA definition of the BARD standard, Josephson refers to his account as 'inference to the best explanation'.

plausibility to avoid (probabilistically) unwanted outcomes. However, this is precisely what I suggest fact finders should do on IBE. As I will argue in Sect. 4, on IBE it is also important to check how well an explanation stands by itself. In fact, as I will argue, it is often useful to primarily look at how well explanations perform in an absolute sense. So, the two approaches would then give fact finders identical instructions.

Yet this amended form of NPA would be mostly uninformative about how fact finders should reason or when guilt can be proven in particular cases. It would amount to little more than a reiteration of the explanationist framework—which tells fact finders to look at competing explanations of the evidence.

In contrast, on IBE, the goal of the fact finder is always to reach relative conclusions: defendants can only be convicted if the case for guilt is much stronger than the case for innocence. Absolute judgment can serve such a relative conclusion in at least two ways. First, it helps to reach such relative conclusions efficiently (Sect. 4.1). Second, it tells us something about the possibility of unconsidered explanations and undiscovered evidence that could have overturned the (relative) case for guilt (Sect. 4.2). So, my account of IBE explains *why* fact finders should look at both relative and absolute plausibility and how these two aspects interrelate. Furthermore, it provides meaningful guidance to fact finders on how to reason about both relative and absolute plausibility as I will show below.

3.3 Relative judgment in practice

By only looking at explanations on their own, fact finders risk making inaccurate decisions. In fact, psychological data suggests that judging explanations in isolation may more often lead to biased conclusions than looking at how explanations perform relative to one another (Sanbonmatsu et al. 1998). For instance, people often fail to ask whether evidence that supports one explanation also supports alternative explanations. Furthermore, they may selectively challenge, discount or ignore certain evidence. Such effects often cause individuals to overestimate the probability of potential causes of events when reasoning in absolute terms (Sanbonmatsu et al. 1993). So, relative judgment appears to be an important method of avoiding biased conclusions.

Apart from its epistemic benefits, relative judgment can also be *strategically* preferable to absolute judgment. Parties may want to make their case by focusing their arguments on the relative plausibility of explanations when it is easier to do so than to argue for an absolute conclusion. I am thinking in particular about cases where the defense gives an unfalsifiable *ad hoc* explanation. For instance, a guilty defendant may give an untrue, vague explanation that only explains the evidence that has been presented to him. Giving direct arguments against this explanation can then be difficult. Yet the prosecution can still make a case by giving reasons *for* the guilt explanation. For example, the madman explanation is low on details. It does not contain many specifics about what the attacker looked like or how the attack occurred. This makes it difficult to argue that the explanation is implausible based on the evidence. Furthermore, given that it *could* have happened, a fact finder might be hard pressed to provide reasons why it conflicts so much with our background beliefs that it can be rejected. In other words, it might be hard to find convincing reasons against this explanation that do not

invoke how it performs compared to the alternative explanation. If the fact finder then wants to convict, a strategy based on relative judgment may then be more prudent than one based on absolute judgment.

In other words, a defendant does not have to be acquitted only because it is difficult to offer reasons why his story is implausible. He can also be convicted if the case for his guilt is substantially more plausible than the case for his innocence (as long as we are sufficiently certain that we have considered all relevant evidence and discovered all relevant evidence, see next section).[17]

Yet is it possible for a fact finder to reason meaningfully about whether one explanation is 'substantially more plausible' than another, as IBE dictates? Larry Laudan objects to this. He believes that jurors will not be able to easily reach consensus about whether the prosecution's case is substantially more plausible than the defendant's. According to him, IBE therefore seems "an invitation to confusion and disparate interpretations of the [BARD] standard" (Laudan 2007, p. 300). I agree with Laudan that people can have different intuitions about whether an explanation is substantially more plausible than another. However, this is not a worry unique to IBE. The same worry applies to NPA (when are explanations 'sufficiently good'?) and to probabilism (when is the case for guilt 'sufficiently probable'?). What matters is that fact finders can meaningfully reason about the relative plausibility of explanations. Let me use the example of the Simonshaven case to show that this type of reasoning is not nearly as difficult as Laudan makes it out to be. Again, I am not interested in drawing conclusions about how fact finders should have reasoned in the actual case. Instead, my point is that rational fact finders *could* reason in relative terms, as follows:

"Guilt is proven in this case, because the 'husband' explanation is substantially more plausible than the 'madman' explanation. To begin, partners killing their wives is a much more common occurrence than random attackers. After all, while random attackers are rare, when a woman is killed, it is often her (ex)partner who did it.[18] Furthermore, the couple was walking together in the forest at the time of the killing, so the husband had the opportunity. Also, he had a motive to kill her—he was jealous because of her new lover—whereas the attacker would have to have attacked without an obvious motive. Finally, it is strange that he called his son, but not the emergency services, when his wife was just killed. Catching the killer would require alerting the authorities as soon as possible. These two facts do not seem to fit well together. Yet the 'husband' explanation does not presume such irregular behavior. The explanation that the husband did it therefore fits much better with the evidence and our general

---

[17] One possible worry is that such a view conflicts with the presumption of innocence because the prosecution has not given a direct argument against the defendant's case. However, while there are different ways in which we can interpret the presumption of innocence, none of them clash with this account. For instance, some interpret it dictating that the onus is on the prosecution to prove the defendant's guilt. Others claim that the presumption creates certain procedural rights for the defendant such as the right to remain silent. Finally, others equate the presumption with the proof standard, meaning that there should be a high burden before we can prove guilt (Roberts 2020). However, my account does set a high burden for proving guilt, because the best innocence explanation should be much more plausible than any innocence explanation. Furthermore, my account does not rule out procedural rights or presumptions in favour of the defendant as it relates purely to the proof standard.

[18] See Dobash et al. (2015).

understanding of the world and is more internally coherent, than the explanation that there was a random unknown attacker."

Note that the fact finder is not establishing that the explanations are (in)sufficiently good when looked at on their own. When looked at on its own, the random attacker explanation might not be so unbelievable. After all, it is something that could have happened and that would explain the evidence in the case. On the fact finders reasoning above, it is the relative plausibility that makes the difference. The reasons the fact finder gives here—such as husbands killing their wives being more common than madmen killing people—establish that one explanation is more plausible than the other. Whether fact finders can easily reach consensus about the validity of such a relative argument is an empirical matter. However, in contrast with Laudan, I see no reason why relative judgment is inherently problematic or even more problematic than other kinds of reasoning in the context of criminal trials.

4 How the comparative approach accounts for absolute plausibility

So far, I have argued that our interpretation of the BARD standard should be attentive to the relative plausibility of explanations. In other words, the case for guilt being substantially better than the case for innocence is sometimes a good reason to consider the guilt of the defendant proven. Yet this does not mean that this is *all* we should care about. For instance, the main reason why (Pardo and Allen 2008) reject an IBE-based interpretatation of the BARD standard is because this does not fit with the structure of criminal proof. After all, we do not want to convict based on an implausible (and therefore, all other things being equal, improbable) case for guilt, even when this is much better than the case for innocence. Can the comparative interpretation proposed here account for this, without *ad hoc* modifications?

I argue that it can. In particular, on IBE we should also look at how plausible both the innocence and guilt explanations are in an absolute sense. First, we can look at the absolute performance of explanations to efficiently reach comparative conclusions (4.1). Second, IBE does not imply that we can convict a defendant based on an implausible guilt explanation (4.2). If all our explanations are implausible then we have reason to presume that a better investigation could have exonerated the defendant. The comparative case for guilt then lacks *robustness* and we should therefore not convict the defendant. That robustness is a requirement for proof of guilt BARD is an indepedently plausible assumption for *any* account of the BARD standard.

4.1 The absolute plausibility of innocence explanations

Proof of guilt BARD often involves eliminating sources of doubt that come in the form of innocence explanations. This idea is found in interpretations of the standard throughout English case law and other jurisdictions that employ the standard, going back at least as far as English evidence treatises of the nineteenth century (Ho 2008, pp. 156–61). For instance, an influential formulation of the BARD standard by Wills includes the formulation that "Every other reasonable supposition by which the facts

may be explained consistently with the hypothesis of innocence must (…) be rigorously examined and successively eliminated" (Wills et al. 1905, p. 312). Similarly, Josephson writes that how confident we should be in the truth of an explanation largely comes "from ruling out or discrediting alternative explanations" (Josephson 2000, pp. 1641–2).

Yet it may seem at first sight as if IBE cannot account for the importance of such successive elimination—i.e. eliminating explanations when looking at those explanations on their own—given that IBE is about how explanations perform relative to each other. However, this view is mistaken. For instance, Bird (2007) describes a special version of IBE that he calls 'Inference to the only explanation'. This is a form of IBE from the philosophy of science that is similar to NPA. Bird describes situations where scientists eliminate all but one of the available explanations because they are not good enough. He argues that this often occurs in science by careful experimental design. This is still IBE, because the remaining explanation is then also the best explanation. After all, it is the only good explanation left. So, elimination through absolute judgment can be a way of inferring to the best explanation.

Such successive elimination is not only a way of getting to the best explanation, it is often more efficient than relative judgment. Psychologists have found extensive evidence suggesting that people often do not use "comparative processing"—i.e. judging something by comparing it to its rivals—"because it is cognitively demanding" (Sanbonmatsu et al. 2011, p. 369). Such a relative comparison requires you to consider features of both the object you are comparing and the alternatives you are comparing it to and the extra step of weighing how these features perform relative to one another. Looking at how explanations do on their own is often simpler; it only requires checking whether each explanation individually meets the desired plausibility threshold. Finding a single piece of evidence that contradicts the explanation may, for instance, suffice for elimination. This often makes it desirable to look at explanations on their own, rather than to how explanations perform relative to one another.

Of course absolute judgment does not take into account the explaining away effects that I discussed in the previous section. However, this is only problematic when the difference in plausibility between the explanations is large, but the innocence explanation is still (somewhat) plausible. Such cases involve a highly plausible guilt explanation that outperforms any barely plausible innocence explanations. In those situations, absolute judgment alone is insufficient and we need to take relative plausibility into account as well.

Yet there is a converse problem for relative judgment in cases where the best guilt explanation is not plausible but is still substantially more plausible than any innocence explanation. In such cases it would be a relative judgment that leads us in the wrong direction, because it would require us to accept a bad explanation. This is the *bad lot problem* that I turn to now.

4.2 The bad lot problem

The bad lot problem is possibly the best known argument against IBE in scientific contexts (Van Fraassen 1989). It has also been raised against IBE in legal contexts.

For instance, Amaya calls this "the most serious problem that a model of IBE for law has to face" (Amaya 2009, footnote 13). The bad lot problem arises when a bad explanation is much better than any others only because those explanations are even worse. Of course, if there is no plausible guilt explanation, the defendant should not be convicted. However, because IBE looks at how explanations perform relative to one another it seems that it would then dictate conviction.

In the philosophy of science, the bad lot problem is often presented as a general worry for *any* truth claim based on IBE. After all, even if our explanation is much more plausible than any alternative, an even better explanation may exist (at least in theory) even if there is no reason to suspect so at the moment. The bad lot problem is worrisome for scientific realists in part because history teaches us that many of our best scientific theories have turned out to be false, even though there was no reason to suspect their falsity at the time. This argument is known as *pessimistic induction* (Laudan 1981).

Ribeiro (2018, p. 32) mentions this sweeping version of the bad lot problem as a potential difficulty for IBE in criminal trials. However, this is not the version of the bad lot problem that we are worried about in criminal law. Suppose that we have a case where the guilt explanation is good and much better than any innocence explanation and there is no other reason to doubt its truth. It would be ludicrous to suggest that the fact finder should acquit in such a case, based on the theoretical possibility that a better explanation might be found. Things would be different if in most or even many of such cases, the verdict later turned out to be erroneous. However, this is obviously not the case.

Instead, the bad lot problem for IBE in criminal law concerns those cases where we *do* have reasons to doubt the truth of the best explanation. For instance, because we only have poor explanations. Let me use aspect an of the Simonshaven case that I have not mentioned so far as an example of what such a bad lot looks like. At some point in the case, the defense argued that the defendant could not have killed his wife because he had arthritis in his hands. They argued that this made him incapable of inflicting the kind of violence that was brought upon the victim. If the defense could prove this claim, it would be destructive to the case the prosecution was trying to make. Yet suppose, counterfactually, that the explanation that the husband was the killer is the *only* available explanation, for instance because the defense never offered the random attacker explanation and the suspect called on his right to remain silent from the beginning. In that case we would only have the explanation that the husband killed his wife. This explanation would be implausible because of the arthritis. It would also be the best explanation, because it would be the only explanation.

One solution to this kind of bad lot problem is to amend the definition of IBE by adding the demand that the best explanation should also be sufficiently good on its own. Some philosophers of science have attempted to save IBE from bad lots this way (Lipton 2003, p. 63; Musgrave 1988, pp. 238–239). However, this would be an *ad hoc* workaround—meant only to deal with the bad lot problem—that does not fit well with the essence of IBE, which is about relative conclusions. Moreover, IBE does not require an amendment at all to deal with the bad lot problem. Rather, the solution rests on a prerequisite condition it shares with NPA (and any other adequate account of proof BARD). This condition is that the case for guilt should be 'robust' (Logue 1997, p. 21), 'stable' (Ho 2008, p. 278]), or 'less likely to be shaken by potential additions

to its information base' (Stein 2005, p. 88). That proof BARD requires a robust case for guilt is a common idea in the philosophy of criminal law.

When is the case for guilt robust on IBE? Within IBE, the defendant can only be convicted if the best guilt explanation is substantially more plausible than the best innocence explanation. However, suppose that we have a reason to suspect that our investigation failed to uncover crucial evidence that might have supported the innocence of the defendant. For instance, a key defense witness might not have been questioned. We then know that if investigators would have questioned the witness, the relative conclusion—that the best guilt explanation is substantially more plausible than the best innocence explanation—might no longer hold. More generally, the case for guilt lacks robustness when we have a good reason to presume that a timely, thorough investigation could have led to discoveries that would have overturned this case. Conversely, if we have no (good) reason to believe that a better investigation would have led to a different outcome, the case for guilt is robust.

There are at least two kinds of discoveries that can overturn the case for guilt. First, those who write about robustness typically link this term to the evidence in the case. We may sometimes have reason to believe that the set of evidence in a case is incomplete, misleading or biased. If we have reason to believe that a better investigation could have uncovered further evidence that would have exonerated the defendant, it is irrational to believe in the defendant's guilt (Ballantyne 2015). For instance, the defense in the Simonshaven case argued that the case for guilt lacked robustness, because of potential missing evidence. This related to a curious feature that I have not discussed so far. During the investigation a different criminal case came up. This case was about a serial killer whose *modus operandi* loosely fitted with how the madman supposedly attacked the defendant and his wife. This serial killer had been active in the broad geographical area of the crime. The police investigated this possible connection further, but failed to find any evidence in its favor. For example, the suspected killer did not fit the description the defendant gave of the supposed attacker. Furthermore, his DNA did not match with any DNA found on the crime scene. Finally, there were no phone records that showed that his phone was near the crime scene at the time of the murder. However, the defense argued that the police's investigation came too late. They claimed that relevant evidence may have been destroyed or lost. Their argument was therefore that the case for guilt lacked robustness because a more timely investigation could have produced relevant evidence. The court of appeal ultimately rejected this claim. However, this *is* an example of an argument against the robustness of the case for guilt. Second, an investigation is about more than collecting evidence. It is also a creative effort, where investigators formulate possible explanations of the evidence. This leads to a different way in which the case for guilt can lack robustness, namely when investigators overlook plausible innocence explanations. When unconsidered innocence explanations are discovered that are better (or at least not much worse) than the best guilt explanation, this guilt explanation is no longer substantially more plausible than any innocence explanation. The requirement that the case for guilt should also be robust in this sense is expressed in the following condition from the definitions of both NPA and IBE (see Sect. 2.2): "Rival explanations for $e_1, e_2, \ldots, e_n$ have been earnestly sought out but the search has produced only $h_1, h_2, \ldots, h_n$." Amaya (2009, pp. 154–155) makes a similar point in response to the bad lot problem: "we need to

have some reason to believe that the set of hypotheses from which we have inferred to the best is 'good enough'." Likewise, Josephson (2000, p. 293) claims that for IBE to work well, it is important to consider the exhaustiveness of our set of explanations. Whether the case for guilt lacks robustness depends on a number of factors. For instance, Ho writes the following: "The [robustness] of an inference depends not just on the evidence available to us, but also on the evidence that we know we do not have; it depends on how well tested the inference is, the rigour of the challenges it has managed to survive" (Ho 2008, p. 278). I propose that the case for guilt *prima facie* also lacks robustness when we face a bad lot situation. In such situations, all available explanations are implausible. This gives us a reason to believe that either some better, unconsidered explanation has to exist or that our evidence is insufficient or misleading (or both). Take the arthritis example again. If the husband had arthritis, this makes it unlikely that he was able to inflict the degree of violence that we know killed the victim. Yet we know that *someone* killed the victim. So, if we have no plausible explanations about how the murder occurred, then it becomes more likely that some other, better explanation for the killing exists, that we have not considered. Alternatively, our evidence that the husband has arthritis might be false.

To summarize, while IBE is about the relative assessment of explanations, making sure that the guilt explanation is plausible on its own is still of key importance. Insufficiently plausible guilt explanations should usually lead to acquittal, even when they are substantially better than any innocence explanation. This is similar to the solution to the bad lot problem mentioned above, that the best explanation also has to be good enough on its own. However, my account is subtly different. Whether the case for guilt is plausible enough for conviction is part of the larger question, whether the investigation (probably) uncovered all relevant explanations and evidence (including evidence about the reliability of our evidence). This means combining information about the quality of the available explanations with other evidence about the robustness of the case for guilt—in particular evidence about quality of the investigation. Compared to the solution that the best guilt explanation should also be 'good enough on its own', this leads to a more elaborate account of when the case for guilt is good enough to warrant conviction.

One particularly important implication is that an *a priori* implausible guilt explanation can sometimes become acceptable when a thorough investigation fails to turn up any alternative explanations or evidence. Such a search for further evidence and attempts to formulate alternative explanations should increase our confidence that the case for guilt is robust. This is important because fact finders in criminal trials often face explanations that are somewhat implausible. After all, criminal cases, especially hard cases, are complex matters. Summaries of the evidence and investigation alone can span hundreds of pages. It would be surprising if all the evidence pointed to one side. Furthermore, criminal cases deal almost by definition with extraordinary circumstances. Explanations are therefore usually, to a certain extent, in conflict with our background knowledge about how the world typically works. Yet suppose that investigators have spent countless hours investigating possible alternative explanations and looking for additional evidence, but this search turned up nothing. Surely this should

have a positive impact on how probable it is that the (*a priori*) implausible explanation is true.[19]

For instance, take the bad lot case I sketched above with the only explanation being the 'husband' explanation, which is implausible if we know that the husband had arthritis. Let us take an extra leap of imagination. Assume the couple was on a small, deserted island rather than in a forest in the Netherlands. If we know for a fact that (1) someone killed the victim and that (2) the investigation has shown that no one else could have been on the island, then the 'husband' explanation suddenly becomes highly probable despite the arthritis.

One additional advantage to linking bad lots to robustness, is that it points to ways in which we can respond to bad lot situations. Such a situation is evidence that a better investigation could potentially have made a difference. This suggests that further investigation might be worthwhile.

A more thorough investigation means looking for further evidence and coming up with new or better explanations. For instance, the claim that the defendant had arthritis led to further investigation which produced a witness who testified that he saw the defendant hauling heavy car parts weeks before the murder. This refuted the defense's claim that the defendant's hands were too weak to exert any serious force.

However, suppose that the investigators had found no further evidence. They could then attempt to introduce a better explanation. This often involves altering the current explanation. Turning to a different example, in the Simonshaven case the prosecution did not find the murder weapon. In response to this stated that the defendant probably threw it in a nearby river after he had killed his wife. In doing so they amended the explanation to make it plausibly fit with the (absence of) evidence about the gun. Similarly, in the arthritis example, the prosecution could argue that the defendant is lying or that having arthritis is compatible with exerting the required force to kill someone. Of course, the fact finder then still has to check whether the amended explanation is better than any innocence explanation and good enough on its own (to warrant the claim that the case for guilt is robust).

The more time and energy investigators have spent, the more robust the case for guilt typically becomes. However, I want to emphasize that even if the investigation was as good realistically possible, the case for guilt can still lack robustness. While a good investigation *is* a reason to assume robustness, such a reason has to be weighed against reasons that point to a lack of robustness of the case for guilt. If all available explanations are (very) implausible then this is a strong reason to believe that an unconsidered explanation exists or that our set of evidence is misleading or incomplete. The case for guilt does then not become robust even if no better investigation was realistically possible. In such a case we have a reason to presume that we have missed important evidence or explanations even if these could only have been discovered by an investigation that would have required resources, opportunities, creativity or time that are unavailable or even impossible in the real world. In contrast, when the best explanation is only slightly implausible, but the investigation was very thorough, we may do not have a reason to presume that such evidence or explanations exist.

---

[19] That the probability of an existing explanation can be raised when a thorough investigation fails to turn up alternative explanations or disconfirming evidence has also been formalized within the Bayesian framework (Dawid et al. 2015; Oaksford and Hahn 2004).

In conclusion, there is no conflict between IBE and looking at how well both innocence and guilt explanations do on their own. However, such absolute judgment is always in service of a conclusion about how the explanations perform relative to one another.

5 Conclusion

Explanationists view criminal trials in terms of competing explanations. These explanations can be more or less plausible. But when is an explanation that implies the defendant's guilt proven beyond a reasonable doubt? In this article I proposed an inference to the best explanation (IBE)-based approach on which guilt is only established BARD if (1) the best guilt explanation in a case is substantially more plausible than any innocence explanation, and (2) there is no good reason to presume that we have overlooked evidence or alternative explanations that could realistically have exonerated the defendant.

This is a comparative account on which we determine whether guilt has been proven by looking at how well our guilt explanation performs compared to any innocence explanations. As I have argued, accurate fact finding—at least sometimes—requires looking at how explanations perform relative to one another. This is because confirming one explanation can lead to the disconfirmation of another. The non-comparative 'no plausible alternative' view of the BARD standard, which many explanationists tactictly assume, overlooks the importance of such relative performance.

Of course we should not only care about the relative performance of explanations. If the case for guilt is bad, we should not convict the defendant. Yet as I have argued, IBE does not imply that we should. When we only have implausible explanations— i.e. when we are in a *bad lot* situation—this suggests that we have missed something and that the relative case for guilt would have been overturned if we had come up with these alternative explanations or if we had found further evidence. Our relative conclusion—that the best guilt explanation is substantially better than any innocence explanation—then lacks robustness. That we should not convict if the case for guilt lacks robustness is an indepedently plausible assumption for any account of the BARD standard. So, the proposed IBE-based account overcomes the bad lot problem without the need for *ad hoc* modifciations.

The proposed account connects various strands of criminal fact finding into a single framework. For instance, Josephson (2000, p. 1626) claims that the strength of any abductive justification is determined by:

1. how decisively the leading hypothesis surpasses the alternatives;
2. how well the hypothesis stands by itself, independently of the alternatives;
3. how thorough the search was for alternative explanations.

Josephson does not discuss why fact finders should care about these specific criteria or how they relate to one another. In contrast, my account offers a coherent account that answers these questions. On this account, how decisively our best explanation surpasses its alternatives is the central question for criminal fact finders. How well the guilt explanation stands by itself relates to the question whether the case for guilt is robust. Finally, the question of whether the search for alternatives was thorough also partially determines whether the case for guilt is robust.

As mentioned in the introduction, Laudan (2006) argues that the most prominent interpretations of the BARD standard are all problematic. He claims that such interpretations are either too unclear to be applied, too reliant on the subjective hunches of the person who applies it or insufficiently connected to the goal of avoiding false convictions. My interpretation of the BARD standard arguably overcomes such problems. First, as I have shown using the Simonshaven case, IBE can meaningfully be applied to real criminal cases. Second, the interpretation frames the standard in terms of the plausibility of explanations and connects this to the goal of getting at probabilistic conclusions. This means that what counts as rational on this interpretation is getting to conclusions that are probably true, not just the hunches of the fact finder. Third, the connection to probability theory also connects this interpretation to the goal of avoiding false convictions.

In conclusion, when fact finders face a choice between competing scenarios, they should infer to the best explanation.

References

Allen, R. J. (2010). No plausible alternative to a plausible story of guilt as the rule of decision in criminal cases. In L. Laudan, J. Cruz (Eds.) *Proof and standards of proof in the law* (pp. 10–27). Northwestern University School of Law.

Allen, R. J., & Leiter, B. (2001). Naturalized epistemology and the law of evidence. *Virginia Law Review* 1491–1550.

Allen, R. J., & Pardo, M. S. (2007). The problematic value of mathematical models of evidence. *The Journal of Legal Studies*, *36*(1), 107–140.

Allen, R. J., & Pardo, M. S. (2017). Explanations and the Preponderance Standard: Still Kicking Rocks with Dr. Johnson. *Seton Hall Law Review*, *48*(1), 1579–1587.

Allen, R. J., & Pardo, M. S. (2019). Relative plausibility and its critics. *The International Journal of Evidence & Proof*.

Amaya, A. (2009). Inference to the best legal explanation. In H. Prakken, H. Kaptein, & B. Verheij (Eds.), *Legal evidence and proof: Statistics, stories, logic*. Aldershot: AshgatePublishing.

Ballantyne, N. (2015). The significance of unpossessed evidence. *The Philosophical Quarterly*, *65*(260), 315–335.

Bex, F. J. (2011). *Arguments, stories and criminal evidence: A formal hybrid theory* (Vol. 92). Dordrecht: Springer.

Bird, A. (2007). Inference to the only explanation. *Philosophy and Phenomenological Research*, *74*(2), 424–432.

Blackstone, W. (1962). *Commentaries on the Laws of England*, vol 2. Beacon, Boston (Original work published 1769).

Callen, C. R. (2017). Spotting a preponderance of the evidence in the wild: Inference to the best explanation and sufficiency of the evidence. *Seton Hall Law Review*, *48*, 1517.

Dawid, R., Hartmann, S., & Sprenger, J. (2015). The no alternatives argument. *The British Journal for the Philosophy of Science*, *66*(1), 213–234.

Di Bello, M .(2013). *Statistics and probability in criminal trials*. PhD thesis, Stanford University.

Dobash, R. E., Dobash, R. P., & Dobash, R. (2015). When men murder women. *Interpersonal Violence*.

Douven, I. (2017). Inference to the best explanation: What is it? and why should we care. In T. Poston & K. McCain (Eds.), *Best explanations: New essays on inference to the best explanation* (pp. 4–22). Oxford: Oxford University Press.

Epps, D. (2014). The consequences of error in criminal justice. *Harvard Law Review*, *128*, 1065.

Fenton, N., Neil, M., & Lagnado, D. A. (2013). A general structure for legal arguments about evidence using Bayesian networks. *Cognitive Science*, *37*(1), 61–102.

Friedman, R. D. (1992). Infinite strands, infinitesimally thin: Storytelling, Bayesianism, Hearsay and other evidence. *Cardozo Law Review*, *14*, 79.

Harman, G. H. (1965). The inference to the best explanation. *The Philosophical Review*, *74*(1), 88–95.

Hepler, A. B., Dawid, A. P., & Leucari, V. (2007). Object-oriented graphical representations of complex patterns of evidence. *Law, Probability & Risk*, *6*(1–4), 275–293.

Ho, H. L. (2008). *A philosophy of evidence law: Justice in the search for truth*. Oxford: Oxford University Press on Demand.

Josephson, J. R. (2000). On the proof dynamics of inference to the best explanation. *Cardozo Law Review*, *22*, 1621–1643.

Kolflaath, E. (2019). Relative plausibility and a prescriptive theory of evidence assessment. *The International Journal of Evidence & Proof*.

Kjaerulff, U. B., & Madsen, A. L. (2008). Bayesian networks and influence diagrams. *Springer Science+Business Media*, *200*, 114.

van Koppen, P. J. (2011). *Overtuigend bewijs: Indammen van rechterlijke dwalingen [Convincing evidence: Reducing the number of miscarriages of justice]*. Amsterdam: Nieuw Amsterdam.

Laudan, L. (1981). A confutation of convergent realism. *Philosophy of Science*, *48*(1), 19–49.

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 epistemology*. Cambridge: Cambridge University Press.

Laudan, L. (2007). Strange bedfellows: Inference to the best explanation and the criminal standard of proof. *The International Journal of Evidence & Proof*, *11*(4), 292–306.

Lawlor, K. (2013). *Assurance: An Austinian view of knowledge and knowledge claims*. Oxford: Oxford University Press.

Lipton, P. (2003). *Inference to the best explanation*. London: Routledge.

Logue, J. (1997). Resiliency, robustness and rationality of probability judgements. *International Studies in the Philosophy of Science*, *11*(1), 21–34.

Lombrozo, T. (2016). Explanatory preferences shape learning and inference. *Trends in Cognitive Sciences*, *20*(10), 748–759.

McGrew, T. (2003). Confirmation, heuristics, and explanatory reasoning. *The British Journal for the Philosophy of Science*, *54*(4), 553–567.

Musgrave, A. (1988). The ultimate argument for scientific realism. *Relativism and realism in science* (pp. 229–252). Dordrecht: Springer.

Nance, D. A. (2001). Naturalized epistemology and the critique of evidence theory. *Virginia Law Review* (pp. 1551–1618).

Niiniluoto, I. (2004). *Truth-seeking by abduction. In Induction and deduction in the sciences*. Dordrecht: Springer.

Oaksford, M., & Hahn, U. (2004). A Bayesian approach to the argument from ignorance. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie exprimentale*, *58*(2), 75–85.

Okasha, S. (2000). Van fraassen's critique of inference to the best explanation. *Studies in History and Philosophy of Science*, *31*, 691–710.

Pardo, M. S., & Allen, R. J. (2008). Juridical proof and the best explanation. *Law and Philosophy*, *27*(3), 223–268.

Peirce, C. S. (1965). *Collected papers of Charles Sanders Peirce* (Vol. 5). Harvard: Harvard University Press.

Pennington, N., & Hastie, R. (1991). A cognitive theory of juror decision making: The story model. *Cardozo Law Review*, *13*, 519.

Pennington, N., & Hastie, R. (1992). Explaining the evidence: Tests of the Story Model for juror decision making. *Journal of Personality and Social Psychology*, *62*(2), 189.

Ribeiro, G. (2018). Can there be a burden of the best explanation? *The International Journal of Evidence & Proof*, *22*(2), 91–123.

Roberts, P. (2020). Presumptuous or pluralistic presumptions of innocence? A methodological diagnosis. *Synthese*.

Sanbonmatsu, D. M., Akimoto, S. A., & Biggs, E. (1993). Overestimating causality: Attributional effects of confirmatory processing. *Journal of Personality and Social Psychology*, *65*(5), 892.

Sanbonmatsu, D. M., Posavac, S. S., Kardes, F. R., & Mantel, S. P. (1998). Selective hypothesis testing. *Psychonomic Bulletin & Review*, *5*(2), 197–220.

Sanbonmatsu, D. M., Vanous, S., Hook, C., Posavac, S. S., & Kardes, F. R. (2011). Whither the alternatives: Determinants and consequences of selective versus comparative judgemental processing. *Thinking & Reasoning*, *17*(4), 367–386. https://doi.org/10.1080/13546783.2011.625659.

Schupbach, J. N. (2016). Inference to the best explanation, cleaned up and made respectable. In T. Poston & K. McCain (Eds.), *Best explanations: New essays on inference to the best explanation*. Oxford: Oxford University Press.

Schupbach, J. N., & Glass, D. H. (2017). Hypothesis competition beyond mutual exclusivity. *Philosophy of Science*, *84*(5), 810–824. https://doi.org/10.1086/693928.

Shealy, M. W. (2012). A reasonable doubt about reasonable doubt. *Oklahoma Law Review*, *65*(225), 225–302.

Stein, A. (2005). *Foundations of evidence law*. Oxford: Oxford University Press.

Sullivan, S. P. (2018). Challenges for comparative fact-finding. *The International Journal of Evidence & Proof*.

Thagard, P. R. (1978). The best explanation: Criteria for theory choice. *The Journal of Philosophy*, *75*(2), 76–92.

Thagard, P. (2003). Why wasn't OJ convicted? Emotional coherence in legal inference. *Cognition and Emotion*, *17*(3), 361–383.

Thagard, P. (2008). *Hot thought: Mechanisms and applications of emotional cognition*. Cambridge, MA: MIT Press.

Trope, Y., & Liberman, A. (1996). Social hypothesis testing: Cognitive and motivational mechanisms. In E. Higgins & A. Kruglanski (Eds.), *Social psychology: Handbook of basic principles*. New York: Guilford Press.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*(4), 293–315.

Urbaniak, R. (2017). Narration in judiciary fact-finding: A probabilistic explication. *Artificial Intelligence and Law*, 1–32.

Van Fraassen, B. C. (1989). *Laws and symmetry*. Oxford: Oxford University Press.

Vlek, C. S. (2016). *When stories and numbers meet in court: Constructing and explaining Bayesian networks for criminal cases with scenarios*. PhD thesis, University of Groningen.

Wagenaar, W. A., Van Koppen, P. J., & Crombag, H. F. (1993). *Anchored narratives: The psychology of criminal evidence*. New York: St Martin's Press.

Wills, W., Beers, G. E., & Corbin, A. L. (1905). *An Essay on the principles of circumstantial evidence: Illustrated by numerous cases*. Boston: Boston Book Company.

Ylikoski, P., & Kuorikoski, J. (2010). Dissecting explanatory power. *Philosophical Studies*, *148*(2), 201–219.