

**EXPLANATION-BASED APPROACHES TO REASONING ABOUT EVIDENCE AND PROOF IN  
CRIMINAL TRIALS**

**Anne Ruth Mackor**

**Hylke Jellema**

**Peter J. van Koppen**

---

---

PENULTIMATE DRAFT 30TH OF MAY 2020

PLEASE CITE FINAL VERSION:

Mackor, A. R., Jellema, H., & van Koppen, P. L. (2021). Explanation-Based Approaches to Reasoning about Evidence and Proof in Criminal Trials. In *Law and Mind: A Survey of Law and the Cognitive Sciences* (pp. 431-470). Cambridge University Press.

---

---

### **1. Introduction**

In this chapter we discuss explanation-based theories on reasoning about evidence and proof in criminal trials, which view such reasoning in terms of evaluating competing explanations. The theories are all grounded in the cognitive psychology of decision making (notably Bennett & Feldman 2014; Cohen, 1977; Gigerenzer, 2007; Gigerenzer & Engel, 2006; Kahneman, Slovic & Tversky, 1982; Pennington & Hastie, 1993). They aim to describe how people actually reason when they have to make sense of a great deal of evidence in, for example, criminal trials. According to these theories in such contexts people typically construct one or more causal explanations for the evidence.

The explanation-based approaches that we discuss in this chapter extend these empirical findings by developing theories that are, in varying degrees, normative. That is, they are concerned with the question how people should reason about legal evidence if they want to do so in a rational manner. These theories investigate to what extent people actually reason in a rational manner but also to what extent they should change their manners. In particular, we discuss three theories, namely Allen and Pardo's relative plausibility theory, Amaya's theory of inference to the most coherent explanation and Van Koppen's theory of anchored narratives or, as it has also been called, the scenario theory. These three theories have in common (i) that causal explanation, more specifically inference to the best explanation (IBE), is at their heart, (ii) that coherence is a core criterion to assess the quality of explanations, and (iii) that as normative theories, they try to stay close to how people actually reason, both in everyday life and in criminal trials. We have chosen to discuss these three theories in some depth instead of giving an overview of all explanation-based theories or trying to give a general characterisation of explanation-based approaches.

In the literature about evidence and proof in criminal law, explanation-based approaches are generally contrasted with probabilistic, more specifically Bayesian, approaches (Fenton, Neil, & Berger, 2016; Schum, 1994; Tillers & Green, 1988). In sections 3 we discuss Bayesian critiques of explanation-based approaches and in section 4 we discuss explanation-based critiques of Bayesian approaches.<sup>1</sup> Here we briefly mention the most salient differences. The approaches differ in particular with respect to the three core characteristics just mentioned.

One important difference between normative explanation-based and probabilistic approaches is that the latter do not aim to or focus on staying as closely as possible to how human beings actually reason. Next to that, although the notion of probability plays a role in explanation-based approaches, it is not the core notion used to assess the quality of the explanations. Instead, coherence plays a central role. The last few years a debate has been going on about the question whether explanatory coherence can be explicated in probabilistic terms (Fitelson, 2003; Olsson, 2005 and 2019; Schippers, 2014; Shogenji, 1999; Siebel, 2011). Finally, causality and explanation mark the perhaps most important differences between explanation-based and probabilistic approaches. There is no room to give an extensive analysis of causality and explanation. However, in a nutshell, to state that A is a cause of B is not merely to claim that B happened after A or that there is a correlation between the occurrence of A's and B's or that the occurrence of A makes the occurrence of B more probable, but that A is efficacious in bringing about B. Accordingly, on an explanation-based approach, first and foremost there is and should be a causal explanation of evidence and only then should factfinders assess the probability that certain events happened. An explanation is an answer to a why-question and a causal explanation is answer to a why-question in which the cause is said to explain why the evidence has occurred. Thus, a causal explanation does not merely predict or retrodict the probability that the evidence occurs, but it aims to answer the question why the evidence has occurred. Moreover, it allows for counterfactual reasoning (for answering questions when and why evidence would have occurred) by offering information about causes and mechanisms. It is said that in this manner, causal explanations aim to offer more than probabilities, namely knowledge or understanding.

In section 5 we discuss the question whether a productive partnership between explanation-based and probabilistic approaches is possible. First, however, we describe the three explanation-based theories in some detail.

---

1. We do not compare the explanation-based approach to a third, argumentative, approach in which logical inference is a core notion (Anderson, Schum, & Twining, 2005; Bex, Van Koppen, Prakken, & Verheij, 2010; Bex, Prakken, Reed, & Walton, 2003; Walton, Reed, & Macagno, 2008). Unlike explanation-based theories and some Bayesian approaches, the argumentative approach is generally seen as useful only to reason about details of a case but not to reason about and to model a complete criminal case (Prakken, 2019). For this reason, the argumentative approach is not dealt with in this chapter.

## 2. Three explanation-based approaches

### 2.1. The story model of Pennington and Hastie<sup>2</sup>

The starting point of the scenario theory of Van Koppen (section 2.4) and, to a lesser extent, of the relative plausibility theory of Allen and Pardo (section 2.2) is psychological research into the way people actually reason when they have to deal with a great amount of evidence. Both Van Koppen and Allen and Pardo refer specifically to the story model of Pennington and Hastie (1993). Therefore, this section starts with an exposition of the model.

The story model is a descriptive psychological theory about cognitive strategies that factfinders use to process trial information in order to take decisions about evidence and proof. Pennington and Hastie claim that factfinders typically use one central strategy, namely active story construction. In doing so, factfinders impose a narrative story organization on the trial information (Bennett & Feldman, 2014; Pennington & Hastie, 1993, p. 194). Their model offers an analysis, both of the structure of these stories and of the dynamics of the way in which people construct and reason about stories.

#### *The structure of stories*

Stories consist of elements, which are called *episodes*. Episodes consist of specific elements, such as an initiating event, a psychological response, sometimes a goal, an action and a consequence. Episodes have a specific structure: the elements are chronologically connected through physical and mental causal relationships. Stories can be thought of as a hierarchy of episodes (Pennington & Hastie, 1993, p. 197). Take for example the following episode that is also a simple story: a husband has an argument with his wife (initiating event), which makes him angry (psychological response). Because he intends to hurt her (goal), he beats his wife (action), which causes her death (consequence).

#### *The dynamics of stories*

Factfinders construct stories by reasoning from *three kinds of knowledge*. They use:

- (a) case specific knowledge, i.e. evidence;
- (b) knowledge about similar events to infer facts and causal relationships;
- (c) knowledge about what makes a story complete: knowledge about the typical elements of stories, episodes and their elements, and about the connections in and between episodes (Pennington & Hastie, 1993, p. 194).

---

<sup>2</sup> This section is an adaptation of section 2.1 of Dahlman and Mackor (2019).

Factfinders use (b) and (c) to 'fill out' a story. Thus, on the story model, a story consists of evidence, inferred facts and causal relations between them. Unlike the normative scenario approach, Pennington and Hastie do not always sharply distinguish between the elements, that is the hypotheses in the story, and the evidence for the story.

Pennington and Hastie (1993, p. 195) in particular mention *three types of reasoning procedures* that factfinders use to establish intermediate and final conclusions:

- deductive reasoning from world knowledge;
- reasoning from analogy to other – experienced and hypothetical – episodes;
- reasoning by evaluating alternate conclusions that contradict the initial conclusion.

Finally, Pennington and Hastie (1993, pp. 198-199) mention three *certainty principles* that factfinders use to assess stories, namely coverage, coherence and uniqueness. These principles help a fact trier to determine how acceptable a story is for him and how confident he is about the truth of the story:

#### 1: Coverage

Coverage deals with the question to what extent the story explains the occurrence of the evidence. The greater the coverage, the more acceptable the story and the more confident the factfinder will be.

#### 2: Uniqueness

A story is unique if it is the only coherent story that can account for the evidence. If there is more than one coherent story, all stories are in principle acceptable, but confidence in each of them will diminish.

#### 3: Coherence

Coherence has three components: consistency, plausibility and completeness.

3a: Consistency is about two questions, namely (1) whether the story is consistent with evidence believed to be true and (2) whether all of its elements are consistent with other parts of the story.

3b: Plausibility deals with the question whether the story fits into the factfinders' background or world knowledge.

3c: Completeness is about the question whether the structure of the story has all its parts, such as episodes, elements of episodes and causal relationships in and between episodes. Missing information (story gaps) and lack of plausible inferences makes a story incomplete and decreases confidence in the story.

Pennington and Hastie state that consistency, plausibility and completeness can be fulfilled to a greater or lesser degree and that the values of the three components combine to yield the overall coherence of the story (Pennington & Hastie, 1993, p. 199). They do not offer a further specification of the components of coherence, nor of the way the individual weight of the components is assessed and the way in which these weights are combined. Moreover, given the fact that a consistent set of statements need not be coherent, their analysis of coherence, which is only in terms of consistency and completeness, may not be fully adequate.

## **2.2. Allen and Pardo's Relative Plausibility Theory**

Over the past decades Allen and Pardo's *relative plausibility theory* has arguably been one of the major contributing factors in a shift in thinking about legal evidence – from probabilistic to explanatory thinking (Allen & Pardo, 2019b, p. 1). We now turn to explaining their theory, drawing mainly on a recent special issue of the *International Journal of Evidence and Proof*, which was dedicated to a renewed defence of their theory and to the responses of fourteen critics (Allen & Pardo, 2019a, 2019b).

Allen and Pardo state that their goal is to explain the different facets of juridical proof of the American legal system, thereby making that system understandable. They propose a theory, which they believe is not just descriptively, but also normatively appropriate in the light of that system (Allen & Pardo, 2019b, p. 2). They focus on proof standards and how to interpret these in terms of explanations (more on that below). However, they seek to explain many other aspects of legal proof as well, including (i) jury decision-making, (ii) how evidence is presented by the parties and used in argumentation, (iii) how evidence is processed by humans, (iv) how rules of evidence create the trial structure, (v) the structure of litigation before and after the trial (Allen & Pardo, 2019b, pp. 3-4). So, while their goal of explaining reasoning and decision-making is related to that of Pennington and Hastie's story model, Allen and Pardo's goal is broader. As they themselves put it, whereas the story model is an empirical description of juror reasoning, relative plausibility theory is an explanation of the 'standards of proof and other features of the proof process' (Allen & Pardo, 2019b, p. 17, footnote 86).

The core of the relative plausibility theory is that juridical proof involves determining the comparative plausibility of competing explanations (Allen & Leiter, 2001, pp. 1527-1528). In their words: 'The primary message of relative plausibility is that from beginning to end the legal system pushes the parties to provide competing explanations, and these explanations structure the decision that is subsequently made' (Allen & Pardo, 2019b, p. 4). For instance, parties in dispute try to show weaknesses in the cases of their opponents and try to show that their own version of reality is true.

Factfinders then have to determine which of these competing explanations best explains the available evidence and to decide accordingly (Pardo & Allen, 2008, p. 228).<sup>3</sup> We now turn to three central aspects of the relative plausibility theory: explanations, plausibility and their account of the proof standards.

### *Explanations*

Just like Pennington and Hastie's story model, the relative plausibility theory proceeds from the premise that humans process evidence 'holistically'. According to Allen and Pardo holistic processing means that people structure evidence and evaluate it by tying it to (causal) explanations when making inferences (Allen & Pardo, 2019b, p. 10).<sup>4</sup> In that respect, the two approaches are very similar. Allen and Pardo use the empirical research underlying the story model – which is evidence that people reason holistically – to support their own theory (Allen & Pardo, 2019b, pp. 16-17). However, their notion of an 'explanation' is not exactly the same as that of a 'story'.

Allen and Pardo's definition of an 'explanation' is quite general: is it "an answer to the question 'what happened' in the particular context of the dispute in question" (Allen & Pardo, 2019a, p. 7). In other words, it is an answer to the question what caused the evidence in this particular case. For example, suppose there is a rape case in which DNA evidence was found, where we have a witness statement of the (supposed) victim that the defendant sexually assaulted her and we have another witness who states that he has seen the defendant running away from the crime scene. One explanation of such evidence is that the defendant did indeed assault the victim. A competing explanation is that the evidence was fabricated by the victim. Both are answers to the question: what happened that caused this evidence (the DNA evidence, the report and the witness)?

Allen and Pardo see stories as a specific type of (causal) explanation that have a chronological narrative form (Allen & Pardo, 2019b, p. 34). Such a story is therefore a specific series of chronologically and causally connected events. However, they claim that not all (causal) explanations take such a chronological form. So, while stories are explanations, not all explanations are stories. In contrast with stories, explanations can also be 'disjunctive' and 'general' (Allen & Pardo, 2019b, footnote 86).

---

3. Pardo and Allen (2008) have also called their account 'inference to the best explanation'. However, Laudan (2007) and Nance (Nance, 2001) call inference to the best explanation a misnomer. They point out that in criminal cases, Allen and Pardo's theory implies that the factfinder does not always choose the best theory. That happens, among others, when the best explanation is not good enough or when the disjunction of multiple explanations in favour of innocence are jointly more plausible than the best explanation implying guilt.

4. Allen and Pardo contrast holistic reasoning with the sequential updating presumed by probabilistic models in which, or so they claim, evidence is processed item-by-item. A more in-depth critique of probabilistic approaches is given in section 4. The probabilistic approach to evidence and the probabilistic critique of explanation-based approaches is discussed in section 3.2.

Disjunctive explanations are made up of mutually exclusive events – so ‘this happened or that’, while general explanations can be of the form ‘something (else) happened’ (Pardo, 2013, pp. 598-599).<sup>5</sup>

Furthermore, while they do not say that explicitly, another possible difference is that not all good explanations have to contain the elements that good stories do, such as a motive and an action. Allen and Pardo note that in areas like no-fault divorce, anti-trust litigation and contract litigation, explanations often do not take the form of stories. However, they also point out that in criminal trials and tort cases, explanations are usually stories (Allen & Pardo, 2019b, p. 27).

### *Plausibility*

According to the relative plausibility theory, explanations should be judged and compared in terms of their plausibility. Allen and Pardo contrast the notion of plausibility with that of probability, which is the key concept in probabilistic approaches to legal evidence. While they do not offer a very precise distinction between the two concepts, we can roughly interpret it as follows. According to them probability is a quantitative concept, leading to statements as ‘I believe that the probability that the defendant is guilty is 0.92’ (Allen & Pardo, 2019a, p. 4 footnote 15). Yet, as Pardo and Allen point out, such probabilistic conclusions either require objective numbers or rely on subjective degrees of belief. Yet, objective numbers are usually difficult to obtain or estimate in legal contexts (Allen & Pardo, 2019b, pp. 9-10). Furthermore, probability theory sets hardly any constraints on subjective degrees of belief. The Kolmogorov axioms of probability, which are discussed in section 4, set only very weak limits on what kinds of probabilities are acceptable, such as that the probability of events should be between 0 and 1. However, such constraints do not rule out many kinds of reasoning that we would ordinarily call irrational and that conflict with the legal system’s goal of obtaining accurate outcomes (Allen & Pardo, 2019b, pp. 9-10). For instance, if a factfinder believes that the probability that a defendant is guilty is 100%, on a probabilistic account no amount of evidence can result in another conclusion. However, we would not call a factfinder who convicts all defendants, even when there is overwhelming evidence for their innocence, rational. Yet nothing in probability theory rules out such a dogmatic and irrational stance.

Plausibility, on the other hand, is a non-numerical notion and its assessment depends on the evidence and our background beliefs. More specifically, an explanation is more plausible when ‘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 extent [sic] it betrays these criteria’ (Pardo

---

5. For instance, Allen and Pardo (2019b, p. 24) mention that plaintiffs in *res ipsa loquitur* cases may win by proving that “the defendant did something negligently and thereby caused my injuries” without being able to identify a specific cause.

& Allen, 2008, p. 230).<sup>6</sup> However, Allen and Pardo say little else about how these criteria should be interpreted and applied.

While plausibility differs from probability, the goal of assessing the plausibility of competing explanations is to reach probabilistic conclusions. Allen and Pardo write that '[e]xplanatory criteria guide inferences and judgments about likelihood' (Allen & Pardo, 2019b, p. 17) and that 'the better the explanation, the more likely true' (Pardo & Allen, 2008, p. 9, footnote 45). Nevertheless, we should not conclude that 'the best explanation' therefore simply means 'the most likely explanation' or that inferences about plausibility are identical to probabilistic judgments (Allen & Pardo, 2019b, p. 20). Instead, inference to the best explanation is a method of arriving at probabilistic conclusions. That fits with how other authors, such as Lipton (2004) and Poston (2014) have described the role of inference to the best explanation (IBE). We discuss the relation between explanation-based and probabilistic approaches in more detail in sections 4 and 5.<sup>7</sup>

### *Proof standards*

While Allen and Pardo aim to explain various aspects mainly of the American proof process, they focus in particular on proof standards. Proof standards determine when certain key propositions, such as whether the defendant is guilty, count as legally proven. Allen and Pardo interpret such proof standards in terms of a comparison of the plausibility of different explanations. Whether an explanation satisfies the proof standard depends on strength of the possible explanations supporting each side (Allen & Pardo, 2019b, p. 13). For instance, under the 'preponderance of the evidence' standard, which is used in civil cases, the party with the most plausible explanation wins (Allen & Pardo, 2019b, p. 14). Under the more stringent 'clear and convincing evidence' standard, which is also used in civil cases, a party wins if their explanation is clearly more plausible than that of the other party (Allen & Pardo, 2019b, p. 15). However, in this chapter we are primarily concerned with criminal law and therefore we restrict ourselves to a discussion of the beyond a reasonable doubt standard.

On the relative plausibility theory, guilt is proven beyond a reasonable doubt only if there is a sufficiently plausible explanation that implies guilt and there is no plausible explanation which is consistent with innocence (Pardo & Allen, 2008, p. 238). Allen and Pardo contrast their interpretation of that proof standard with a probabilistic interpretation, according to which the standard should be interpreted as a probability threshold. Guilt is then proven if the probability that the defendant committed the criminal act of which he is accused exceeds a certain threshold (e.g. 95%).

---

6. Pardo and Allen (2008) seem to derive these criteria from Thagard (1978). Amaya also builds upon Thagard's theory, but – as we will see below – her theory is more precise and detailed.

7. Allen & Pardo (2019b, p. 12, footnote 74) explicitly refer to Lipton's approach.

In conclusion, Allen and Pardo offer an alternative to probabilistic accounts of legal evidence.<sup>8</sup> In doing so they have put the discussion about explanation-based versus probabilistic approaches on the agenda. However, their account remains underspecified on several points. In particular, they say little about the nature of explanatory reasoning and about what makes explanations plausible. While they state that explanatory reasoning is holistic and distinguish criteria such as coherence, they do not elaborate upon these terms. In the next section we turn to Amaya's theory, which offers a detailed analysis both of the concept of coherence and of the process of coherence maximization.

### **2.3. Amaya's theory of inference to the most coherent explanation<sup>9</sup>**

Amaya's theory of inference to the most coherent explanation (Amaya, 2009, 2013, 2015) is built on Thagard's theory of explanatory coherence (Thagard, 2000). Amaya offers a detailed analysis of coherence and the role it should play in inferences in evidential judgments in law. Thus, Amaya's theory is an analysis of both of the *concept* of coherence and the *process* of coherence maximization. Coherence is claimed to play a role in the generation and the pursuit as well as in the justification of evidential judgments in criminal law.

The purpose of Amaya's theory is not to describe or explain legal practice, in America or elsewhere. Thus, it is much less focused on legal practice than Allen and Pardo's theory. Also, although Thagard developed his theory as competitor to a probabilistic model, Amaya's goal is not, unlike Thagard and Allen and Pardo, to criticize probabilism. Finally, although Amaya's theory focuses on coherence, it does not make use of the story model and the role coherence plays in this model. In particular, she claims that her notion of coherence as constraint satisfaction differs from the notion of narrative coherence as it is used in the descriptive story model (Amaya, 2015, p. 109).

The two most important ingredients of Amaya's theory are coherence and inference to the best explanation. On Amaya's view, the best explanation of a (criminal) fact is the one that does best on a test of coherence (Amaya, 2009, p. 137). A third and fourth component of Amaya's theory are a responsibilist epistemology (Amaya, 2013, p. 24) and the demand that standards of justification should be contextualized (Amaya, 2013, p. 27). On a responsibilist epistemology justification is not analysed exclusively in terms of evidential support, but also in terms of what a factfinder has done or failed to do, more in particular in terms of how *thorough* or *robust* the investigations have been (Amaya, 2009, p. 154). The fourth component, the demand of contextualization, implies that the theory of inference to the most coherent explanation should *fit the specific context* in which it is applied. In this section, the focus is on criminal law, more specifically on its rules and principles.

---

8. We discuss their critique of probabilistic interpretation of the proof standard in section 4.

9. This section is an adaptation of Dahlman and Mackor (2019) section 2.2.

## Coherence

Amaya relies on Thagard's analysis of coherence, in particular on his analysis of explanatory coherence. Thagard (2000, pp. 15 ff) defines coherence as the satisfaction of a set of positive and negative constraints – coherence and incoherence relations – among a set of elements. The set of elements E is divided into two disjoint subsets A (accepted) and R (rejected).<sup>10</sup>

The main elements in the assessment of evidentiary judgments in a criminal case are hypotheses (H) and evidence (E). The main, but not the only, type of coherence involved in these judgments is explanatory coherence.<sup>11</sup>

Thagard (2000, p. 43) distinguishes seven principles of explanatory coherence. Amaya states that when applying Thagard's theory to a particular problem, it must be further specified and contextualized. In her application of Thagard's theory in the context of criminal law, she adds two coherence principles, namely to E4 and to E7 (see below) and she stresses that next to explanatory coherence, deliberative coherence should play a role too. She does not, however, offer a further specification of the elements (H and E) in the set. More specifically, Amaya does not distinguish between kinds of hypotheses and between types of evidence. In particular she does not incorporate the distinction between episodes and the different kinds of elements that the story model distinguishes.

We list Thagard's seven principles of explanatory coherence and also the two sub-principles Amaya adds to them:

### E1: symmetry

Explanatory coherence is a symmetrical relation, unlike explanatory relations and relations of conditional probability.<sup>12</sup>

### E2: explanation

- a) a hypothesis (H) coheres with what it explains, evidence (E) or another hypothesis (H);
- b) hypotheses that together explain another proposition (H or E) cohere;

---

10. A problem with the idea of constraint-satisfaction is that the constraints are not equally important and must therefore be weighted. Thagard's theory does not answer the question according to which criteria weights should be assigned.

11. Thagard (2000) distinguishes explanatory, analogical, deductive, visual, conceptual and deliberative coherence.

12. For instance: whereas two propositions H and E cohere with each other equally, explanation is asymmetrical in that H causally explains E, but E does not causally explain H. E is part of an evidential explanation of H, i.e. E is a reason to believe H. The asymmetry of conditional probability is even more obvious in that the probability that H is true given E, i.e.  $p(H|E)$ , is not the same as the probability that E is true given H, i.e.  $p(E|H)$ . Those who do not distinguish between these two probabilities commit the so-called prosecutor's fallacy.

c) the more hypotheses are needed to explain something, the lower the degree of coherence.<sup>13</sup>

E3: analogy

Similar hypotheses that explain similar pieces of evidence cohere.<sup>14</sup>

E4: data priority

Propositions that describe the results of observations and evidence have a degree of acceptability on their own.

For the context of criminal law, Amaya (2013, p. 13) adds the principle that factual hypotheses that are compatible with innocence have a degree of acceptability on their own. She adds this principle to satisfy the demand of the presumption of innocence.

E5: contradiction

Contradictory propositions are incoherent with each other.

E6: competition

Hypotheses that both explain a proposition but that are not explanatorily connected are incoherent (i.e., they are incoherent even if they do not contradict each other).

E7: acceptance

The acceptability of a proposition in a system depends on its coherence with other propositions. For the context of criminal law Amaya (2013, p. 13) adds the principle that the guilt hypothesis may be accepted only if it is justified to a degree sufficient to satisfy the reasonable doubt standard.

### *Inference to the most coherent explanation*

Amaya (2009, 2013, 2015) not only offers an analysis of the concept of coherence but also of the process of coherence maximization. She argues that that process is an explanatory inference, that fits the model of inference to the best explanation.<sup>15</sup> Amaya uses Lycan's (1988, 2002) definition of IBE:

---

13. That principle expresses the epistemic virtue of simplicity. Note that simplicity is an ambiguous term. It can refer to the number of hypotheses but it can also refer to the number of hypotheses in relation to the number of pieces of evidence they explain. Thagard (1978) also discusses other epistemic virtues, such as consilience and analogy. Allen and Pardo (2019b) also refer to epistemic virtues, among which simplicity and consilience, as criteria to assess the quality of explanations.

14. Thus, analogical coherence plays a role in the assessment of explanatory coherence. Pennington and Hastie (1993, p. 195) too claim that people reason by analogy to experienced or hypothetical episodes and that in doing so analogies play an important role in 'filling out' the story on parts where evidence is missing.

15. Note that just as a distinction is made between the descriptive story model and the normative scenario approach, a distinction is made between descriptive and normative theories about IBE. Lipton (2004), for instance, offers an analysis that is primarily descriptive, giving an account of how people actually reason. He also offers a tentative normative approach, giving an account of how people should reason if they want to be rational.

F1 ... Fn are facts in need of explanation

Hypothesis H explains F1 ... Fn

No available competing hypothesis explains Fi as well as H does

-----

Therefore, probably H is true

As was explicated above, Amaya's criterion of best explanation is best on a coherence test. She states that an inference to the most coherent explanation consists of the following explanatory inference steps (Amaya, 2013, p. 16):

1. the specification of a base of coherence, i.e. the set of factual hypotheses and evidence over which the coherence calculus proceeds;
2. the construction of a contrast set that contains a number of alternative theories from which the most coherent is to be selected;
3. refining and revising the alternative theories by means of coherence-making mechanisms, in particular addition, subtraction and, the combination of the two, reinterpretation. That can result in a revision of the contrast set (2), but it can also lead to revision of the base set (1);
4. the evaluation of the coherence of the alternative theories by means of the principles of explanatory coherence E1-E7;
5. the selection as justified of the most coherent theory, provided that its degree of justification satisfies the applicable legal standard of proof, in the context of criminal law the beyond a reasonable doubt standard.

#### *Coherence versus intuition*

Amaya explicitly considers the possibility that the theory that best satisfies the criteria of coherence nevertheless seems intuitively unjustified (Amaya, 2013, p. 18). She points out that errors may have been made in step 1, the selection of the elements of the base set of hypotheses and evidence. In particular, it is possible that relevant evidence has been ignored. In step 2, the constructing of the contrast set, factfinders can also fail and as a consequence of that they will at best end up with an inference to the best of a bad lot (Van Fraassen, 1989). That critique is discussed in section 3.2.

Problems can also arise in the inference to the most coherent explanation (step 4). Amaya points out that people suffer from coherence bias, since they may inflate some alternatives and deflate others in order to maximize coherence. By ignoring or misrepresenting evidence or alternative hypotheses, they distort the set of evidence and hypotheses that threaten their beliefs.

How can factfinders prevent these errors? Amaya explicitly mentions the duty to actively search for alternative hypotheses (Amaya, 2013, p. 26) and she also mentions the duty to gather additional evidence about propositions that are less certain and she instructs factfinders to believe all and only propositions that are supported by available evidence (Amaya, 2013, pp. 25-26). These duties are part of her responsibilist epistemology which entails the demand to perform robust investigations.

In conclusion, Amaya offers a detailed analysis, both of the concept of coherence and of the process of coherence maximization. Amaya does not, however, make use of the story model. We now turn to an approach that explicitly takes the descriptive story model as its starting point for a normative theory.

#### **2.4. Van Koppen's scenario theory<sup>16</sup>**

The theory of anchored narratives or the scenario approach as we call it here, has its roots in the story model of Pennington and Hastie. Wagenaar, Van Koppen & Crombag (1993) developed the story model into both a descriptive and a normative theory. In this chapter we only discuss the normative version of the scenario approach. The normative theory only deals with criminal trials and explicates how people should reason if they want to make rational decisions about evidence and proof. Since 1993 the approach has been further developed, especially by Van Koppen (2011, 2013) and more recently by Van Koppen and Mackor (2019).

Like the theories of Allen and Pardo and Amaya, the scenario approach conceives of reasoning about evidence in criminal trials in terms of explanations, more specifically causal explanations. An important difference between the scenario approach and the theories of Amaya and of Allen and Pardo is the central role that stories or scenarios play. Amaya does not make use of the story model and we have seen that Allen and Pardo refer to Pennington and Hastie's theory, but claim that parties offer explanations which can but need not be stories (Allen & Pardo, 2019b, p 3 note 7 and p 13 note 86).

Another difference between the relative plausibility theory and the scenario theory is that the former is primarily an explanatory account of the American system of legal proof whereas the latter is primarily a normative theory. A possible reason for that difference may be that the American legal system has already developed a complex set of rules of evidence. One of Allen and Pardo's aims is to show that these rules are rational. Such a set of rules is lacking in The Netherlands and in many other countries. Accordingly, in The Netherlands and elsewhere there was and still is a need for a normative theory that offers legal practitioners a set of criteria that can guide them through reasoning with evidence in criminal law.

---

16. This section is based on Van Koppen and Mackor (2019, section 1).

A final difference between the relative plausibility theory and the scenario approach is that the former emphasizes that probabilism is incorrect and incompatible with trials (Allen & Pardo, 2019b, abstract), whereas the latter does not make such a bold claim. On the contrary, in its more recent version (Van Koppen & Mackor, 2019), the value of triangulation – i.e. of comparing the analyses of different approaches to reasoning about and with evidence – is emphasized and it is argued that explanation-based and probabilistic approaches are not only compatible but also complementary.<sup>17</sup>

The scenario theory is underpinned by insights from three different disciplines: psychology, epistemology and philosophy of science. From psychology it takes the notion of stories or scenarios (Pennington & Hastie, 1993). From psychology and epistemology, it takes the idea that in everyday life people both use and should use an inference to the best explanation (Harman, 1965; Lipton, 2004) to create, evaluate and select scenarios. From epistemology the scenario theory takes a coherentist view of knowledge (Amaya, 2009, 2013, 2015; Thagard, 2000) as opposed to foundationalist theories. Finally, in the scenario theory insights from philosophy of science are applied, in particular Popper's falsificationist theory (Popper, 1963), drawing on the analogy between the assessment of scientific theories and the assessment of scenarios.

The scenario approach is based on the following core notions: scenario; coherence; background knowledge; IBE, alternative scenarios and discriminating facts; falsification; and prediction. These notions are spelled out below.

### *Scenarios*

The core notion of the scenario approach is 'story', 'narrative' or 'scenario'. Following Pennington and Hastie (1993) the scenario theory entails that scenario is a hypothesis about an action (or event), offering a chronological and causal description of that action. A scenario consists of at least a central action and a scene that makes the central action understandable. A scenario can have more elements that can be categorized under one of the headings: scene, motive, action, actor, consequences. In a complete scenario all these components are described. The scenario approach emphasizes that the scenario of the prosecution should be compared to at least one alternative scenario.

### *Coherence*

The central question in the scenario approach is: why should we believe that the indictment is true? The scenario approach endorses a coherentist view of knowledge to answer that question. Coherentist views should be distinguished from foundationalist, in particular empiricist, views. Scenarios are not

---

17. Van Koppen's (2011) more critical view of probabilistic approaches to evidence and proof is discussed in section 4.

anchored, as foundationalist empiricists see it, in facts about the case under investigation or in our sensory experiences of these facts, but in a coherent web of statements about the world. In that web, some statements, in particular statements about observations that are accepted as true, have more weight than others, but none of the statements are foundational or anchored in 'the world' or in our observations of 'the world'.<sup>18</sup>

According to the scenario approach, for a scenario to be coherent it should minimally fulfil the following three criteria:

- a. it must be internally coherent,
- b. it must be coherent with background knowledge,
- c. it must be coherent with all the evidence about the case.

These criteria are derived from, but differ slightly from, those of Pennington and Hastie (1993).

First (a), a scenario must be internally coherent. For a scenario to be internally coherent, it must be internally consistent (i), complete (ii) and detailed (iii).

- (i) The criterion of internal consistency demands that statements within one scenario do not contradict each other.
- (ii) A scenario can be more or less complete. That means that it can have more or less of the elements mentioned earlier, as for instance scene, motive, action, actor, consequences.
- (iii) Detailedness concerns the level of detail of the content of each of the elements.

The more details there are, the less story gaps the scenario contains. We distinguish between *story gaps* and *evidence gaps*. A story gap is a gap in the story being told. For instance, the simple story 'I was in Amsterdam this morning and am in London now', has a gap, namely how I got from Amsterdam to London. An evidence gap is a gap in the anchoring of a part of the scenario via sub-scenarios in background knowledge. So, if there is no evidence that I was in Amsterdam this morning, the story has an evidence gap on that point. If there is a story gap, there cannot – for lack of a story – be an evidence gap. Conversely, the more details in the story, the more options to predict and confirm facts, but also the larger the risk of falsifications and evidence gaps (Popper, 1935).

Second (b), a scenario must be coherent with our general background knowledge. In fact, world knowledge is paramount in every aspect of the scenario theory. It determines our judgement about

---

18. Compare Thagard's principle E4 about data priority. See section 2.3 above.

whether a piece of evidence is relevant to what needs to be proven. It also determines how strong evidence is considered as proof of something. It also determines how we interpret scenarios and evidence. And it determines what additional probing must be done before a proposition is considered proven. At the same time, our general – and usually shared – background knowledge of the world changes all the time. In the sense background knowledge may be as dynamic as scenarios can be.

The central claim the scenario approach makes is that each and every word and sentence in each scenario carries a lot of general world knowledge we take for granted. It is usually left implicit, but that world knowledge guides us through the scenarios and sub-scenarios. Part of that background knowledge is contained in scripts and the concept of a scenario is related to that of a script (DiMaggio, 1997; Kleider, Pezdek, Goldinger & Kirk, 2008; Taylor & Crocker, 1981). Scripts contain general background knowledge about the world and about object of the script.

For example, if a scenario entails that a husband killed his wife, we can derive that suspect and victim knew each other, had a heterosexual relation, maybe have children, and much more. That information is based on what we call our shared knowledge of the world. It is often taken for granted, but it may become important anytime in a case and can lead to a discussion about evidence.

The distinction between the internal coherence of a scenario and coherence of the scenario with background knowledge can be explicated as follows. A fairy tale, for example, can be internally coherent, but not coherent with background knowledge. The same holds for a story, that is internally consistent, but empirically impossible or extremely implausible, as the scenario in which the defendant has travelled by car at an average speed of 300 km/h.

Third and finally (c.), a scenario must be coherent with specific knowledge about the case under consideration.

The scenario approach is also called theory of anchored narratives because we can rephrase criteria 2 and 3 as the demand that the scenario must be ‘anchored’, not in ‘the world’, but in narratives about the world that are accepted as true. The answer to the question ‘Why should we believe ...?’ is the evidence that will always take the form of another scenario or multiple scenarios. These scenarios are called sub-scenarios. These sub-scenarios, when accepted, form the evidence in the case that, in the end, are anchored in generally accepted knowledge about the world. In practise the, in principle endless, regression into next levels of sub-scenarios ends.

An important question, however, is why we should be justified in accepting scenario’s that are internally coherent, coherent with evidence at hand and with our general background knowledge. Stated differently, how can we argue that the story model is not just a good descriptive theory of how

people reason, but also a good normative model of how people should reason? Another important question is whether and in which respect the normative model differs from the descriptive model.

#### *Inference to the best explanation (IBE), alternative scenarios (AS) and discriminating facts*

According to the scenario approach factfinders are instructed to construct at least two scenarios and to assess and compare them. Like the other two theories discussed in this chapter, that assessment and comparison is interpreted in terms of inference to the best explanation (IBE) which consists in accepting a scenario on the grounds that it provides a better explanation of the evidence than any alternative scenario that has been proposed.

According to the scenario approach, a fact that 'makes the difference' in the comparison of one set of scenarios, a discriminating fact, needs not make a relevant difference in the comparison of a different set of scenarios or of one of the scenarios against another scenario. So, for example, seemingly incriminating evidence that there is blood of the victim on the defendant does not discriminate between the 'guilty' scenario and the alternative scenario if the defendant told the police that he was walking close to the victim when they were attacked by an unknown perpetrator (Van Koppen, 2011, 2013; this example is discussed in Van Koppen & Mackor, 2019). Again, it is pointed out that our general knowledge plays an ineliminable role in assessments of relevance and weight of evidence, in choice of scenarios, in choice of evidence and in assessing the relation between evidence.

#### *Falsification*

In the scenario approach insights from Popper's philosophy of science are applied, emphasizing the importance of falsification. Popper's theory entails that one should not only look for evidence that a scenario can explain, but first and foremost for evidence that can be used to falsify the scenario (Popper, 1963). On that point the normative scenario approach is different from actual practice since people are inclined to confirm, not to falsify (Ask & Granhag, 2005; Lewicka, 1998; Meissner & Kassin, 2004; Oswald & Grosjean, 2004).

The process of falsification is testing whether a scenario is inconsistent with evidence. In the scenario theory, evidence is not 'facts out there' but sub-scenarios that are accepted as true. The only good manner to do so is searching for reasonable alternatives for each scenario and sub-scenario and then trying to find evidence that best discriminates between the scenario and its alternatives. The importance of assessing and comparing alternative scenarios is also stressed by Allen and Pardo (2019b) and by Amaya (2009, 2013, 2015). However, neither Allen and Pardo nor Amaya mention an important manner to test a scenario, namely to derive sufficiently precise and risky predictions from the scenario and investigate whether they can be falsified or confirmed.

### *Creation, accommodation, prediction*

Imagine a scenario that was created only after all the evidence was known. In that situation, the scenario is likely to offer neat explanations for all the facts. That, for instance, often holds for the prosecution's scenario that is usually formed after the police did an extensive investigation. Then the scenario is made to accommodate for all the evidence uncovered by the police. The same often holds for the defendant who, until trial, invoked his right of silence and only then comes up with an explanation to accommodate all the evidence presented by the prosecution. However, sometimes these explanations are in some way *ad hoc*. For example, sometimes implausible auxiliary hypotheses are added which play no other role than to make the scenario fit the facts (see Mackor, 2017, for the analysis of such a Dutch case).

Therefore, it is relevant to assess whether a scenario was solely created on the basis of the evidence or whether the scenario in part *predicted* the evidence. More specifically, we need to distinguish three different relations between a scenario and facts, namely creation, accommodation and prediction (Mackor, 2017). One of the strengths of explanation-based approaches is that a hypothesis, or more specifically a scenario, enables factfinders to predict facts. Testing of predictions is an important way to assess the quality of explanations, both in science and in criminal trials.

The first step in a criminal case is the creation of a scenario. It is created, typically by the police during the investigation, on the basis of one or a few or many known facts. The scenario is created on the basis of certain facts, but it is also created in order to explain those very same facts. After the scenario has been created, police investigations will continue. Three situations can be distinguished. First, if a novel piece of evidence does not cohere with the scenario, the conclusion may be that the scenario should be rejected. However, secondly, one may also conclude that the scenario need not be rejected, but that it must be accommodated. If the core of a scenario, e.g. defendant is the culprit, is falsified by evidence, the scenario will be rejected. If only a minor aspect of a scenario is falsified, the scenario can be accommodated. By means of adding or deleting elements, the scenario can be improved in order to explain facts that were not yet known when creating the scenario (Popper, 1963). Third, the police may discover evidence that was actually predicted on the basis of the scenario. A scenario gains strength especially if a detailed and risky prediction is confirmed (Lipton, 2007). However, it is also possible that no confirmation or falsification is found. In that case the scenario has an evidence gap. A scenario with evidence gaps is less strongly supported than a scenario without, depending on the number and sizes of the evidence gaps (see Mackor 2017 for the analysis of a case with evidence gaps). In conclusion, in the scenario approach it is stressed that the weight of the evidence is not only determined by the question whether it discriminates between a scenario and an alternative scenario

and by the question how well-established or reliable the evidence is, but also by the question whether the evidence was used to create or accommodate the scenario or whether it was unknown and predicted by the scenario.

### **3. Psychological and probabilistic critiques of explanation-based approaches**

Explanation-based theories have been criticized on several grounds. The most familiar critique is that they are underdeveloped since core notions such as coherence, inference to the best (causal) explanation and plausibility are insufficiently defined. In this section we discuss more specific points of critique that have been launched by empirical scientists, more in particular psychologists, and by Bayesian probabilists.

#### **3.1 Psychological critique of explanation-based approaches**

Explanation based and story model approaches to legal decision making in criminal cases can be criticised on psychological grounds, namely that they justify rather than counteract weaknesses and fallacies in human reasoning. First, decision makers typically tend to overvalue evidence that supports their favourite decision and undervalue evidence against the favourite decision. They simply do not evaluate competing scenarios. That tendency, commonly denoted as confirmation bias (Nickerson, 1998), is seen as a general human cognitive process. It is part of what is called the human tendency to keep our vision of the world without too much dissonance (Festinger, 1957). It is also referred to as a holistic evaluation of evidence (Schweizer, 2014). And, indeed, the story as presented by for instance the prosecution can be appealing to such an extent that evidence pointing in the other direction is ignored (Schweizer, 2014; Wagenaar, et al., 1993), for instance leading to miscarriages of justice (Gross, 1998, 2008; Gross, Jacoby, Matheson, Montgomery & Patel, 2005; Huff & Killias, 2008).

Second, decision makers usually rely on simple – fast and frugal – heuristics rather than careful consideration of alternatives (Gigerenzer, 2007; Gigerenzer & Engel, 2006). A normative model that directs decision makers to evaluate competing scenarios does not come naturally to humans.

Nevertheless, the story model (Pennington & Hastie, 1993) and, following that, the scenario approach (Van Koppen, 2011, 2013; Wagenaar, et al., 1993) and the theory of Allen and Pardo (Allen & Pardo, 2019b), have been built on how people actually reason in decision-making in criminal cases. There is, for instance, a host of research on how jury-members decide on cases (see for an overview Hans, 2008; Van Koppen, 2009; Vidmar, 2009, 2011; Vidmar & Hans, 2007). All of that research demonstrates that people think and decide about the world in the form of stories, of scenarios. However, would the story model be a descriptive model of actual decision making in criminal cases, it should allow for the confirmation bias (Kassin, Dror & Kukučka, 2013; O'Brien, 2007; Oswald & Grosjean, 2004) and the

coherence bias according to which, factfinders maximize coherence by discounting contradicting evidence, inflating supporting evidence and by interpreting ambivalent evidence in a way that is coherent with the emerging decision (Schweizer, 2014, p. 66; Simon, 2004, p. 522). People have the tendency to look for verification of their hypotheses and do not look for falsifying their hypotheses.

Another point of critique is that much of what is happening in criminal cases cannot be grasped in the form of a scenario or story. For instance, Simon (2019, p. 3) argues that in many instances in criminal cases there is not much of a story to tell. Simon gives the example of a neighbour who hears a dreadful shout in the night. However, Simon's critique misses the mark. Although a shout in the night may not be a full-blown story in the sense of 'human action sequences connected by relationships of physical causality and intentional causality between events' (Pennington & Hastie, 1993, p. 196), it is a story all the same. Through our general world knowledge, a shout in the night points at a crime taking place with everything that brings such an occurrence to mind, based on our knowledge of the world. It may cause somebody to have a look outside or phone the police.

A third point of critique is that, on any explanation-based approach, a hypothesis or a scenario is offered that can causally explain the evidence at hand. Thus, the argument runs from the hypothesis to the evidence and not the other way around. In everyday life, however, people reason in both directions.<sup>19</sup> That happens, for instance, in a police investigation, especially in the first stages, right after the crime is discovered. Based on what the police uncover, the evidence, they try to build scenarios of what may have happened (De Poot, Bokhorst, Van Koppen & Muller, 2004). In a later stage, often the very same evidence is used to assess the veracity of the scenarios built. From the point of view of confirmation bias that may seem an unsound practice. In actual criminal cases it is unavoidable.

The points just mentioned may be valid critique of how people are actually inclined to reason in criminal cases. However, they miss the mark with respect to normative theories. Take for instance van Koppen's scenario theory. On this approach fact finders should formulate scenarios that are resistant to additional why-should-I-believe-you-questions. Therefore, on the scenario theory decision makers are forced to consider formulating strong scenarios and cannot get away with sloppy work on that part. Second, since scenarios must always and explicitly be compared to other competing scenarios, this instruction functions as a hallmark prevention of confirmation bias. Especially the instruction to try to falsify the favourite scenario is excellent practice in preventing confirmation bias. Finally, the obligation to formulate a scenario as precisely and detailed as possible makes it possible to formulate precise and possibly risky predictions the outcomes of which might confirm or falsify the scenario.

---

19. Bex (2011) has developed a hybrid theory according to which factfinders reason causally within the scenario, but argue evidentially, from evidence to story, when they connect the scenario and the evidence for it.

### 3.2 Probabilistic critiques of explanation-based approaches

The approaches we have discussed so far all rely on inference to the best explanation. As we have seen, IBE is an inference from an explanation being best out of the available ones to the conclusion that the explanation is probably true. One important strand of criticism aimed at IBE, and thus at explanation-based approaches in criminal law, comes from Bayesian probabilism, a normative model for evidential reasoning view which we discuss in more detail in section 4.

Many Bayesian probabilists defend the view that factfinders who draw conclusions based on the evidence in a case, should do so in accordance with probability theory. That does not necessarily mean that factfinders should engage in actual probabilistic calculations in their decision process.<sup>20</sup> Rather, it means that factfinders' conclusions should be consistent with those they would reach if they would have (correctly) applied probabilistic calculations (Friedman, 1997, p. 289). However, some theorists have argued that explanation-based approaches are, at least potentially, in conflict with Bayesian probability theory. We examine five objections from adherents of Bayesian probability theory and the responses of adherents of explanation-based approaches. The first four objections target the idea that there is a necessary link between an explanation being the best explanation and that explanation being probably true. The last objection asks why a good, as for instance coherent, story would also be a probable story.

#### *A. The bad lot problem*

The bad lot problem is possibly the best-known argument against IBE in philosophy of science (Van Fraassen, 1989). It has also been raised against IBE in legal contexts. For instance, Amaya calls it 'the most serious problem that a model of IBE for law has to face' (Amaya, 2009, footnote 13). The problem is that on IBE we should choose the best<sup>21</sup> explanation out of the available ones. However, an explanation can be the best without being good. If all available explanations are bad ones, even the best explanation will be poor. In that case IBE would tell us to accept a bad explanation as true. In criminal law that would imply that a defendant could be convicted on a weak case.

---

20. Though it is suggested that it may be useful to do so when it comes to evidence that has a statistical basis such as DNA evidence (Allen, 1997, p. 258; Stein, 1996, p. 35). Probabilistic methods are also proposed for reconstructing and analysing cases by experts (cf. Fenton, et al., 2016; Kadane & Schum, 2011). See, however, the discussion below.

21. Given a specific definition of explanatory goodness. See Ylikoski & Kuorikoski (2010) for an overview of how philosophers have interpreted this term. Furthermore, see the above accounts that define how well an explanation does in terms of criteria such as, for example, simplicity (Allen and Pardo), coherence (Amaya and van Koppen, or the absence of story gaps (van Koppen).

One response to that objection has been to amend the definition of IBE by adding the demand that the best explanation should also be sufficiently good on its own (Lipton, 2004, pp. 63, 154). However, even if the best explanation is good on its own, it may still not be enough to guarantee its probable truth, if we have failed to consider even better explanations. That worry is not unique to IBE. However, it does mean that factfinders have to ensure that they are reasonably sure that they have considered all relevant explanations. For instance, Amaya (2009, pp. 154-155) states that ‘we need to have some reason to believe that the set of hypotheses from which we have inferred to the best is “good enough”.’ Similarly, Van Koppen (2011, p. 52) emphasizes that the scenario-based approach requires that factfinders consider all ‘reasonable’ scenarios.

### *B. Dutch Book argument*

Another argument against IBE has also been formulated by Van Fraassen (1989). That argument starts with the well-known Dutch Book argument. According to that argument, whenever an agent’s degrees of belief in certain propositions (such as the proposition that the defendant is guilty) violate the maxims of probability theory, that agent risks having a Dutch book argument made against them (Vineberg, 2016).

A Dutch Book is a series of bets that guarantee that the agent will lose money. That argument presupposes a hypothetical betting scenario where the agent has to bet on the truth of some propositions. The betting scenario is used as an analogy for decision-making in real-life. The Dutch Book argument is too complicated to fully explain here. What is important, however, is that Bayesians often use the Dutch Book argument to support the idea that any agent who reasons in a way that deviates from the precepts of (Bayesian) probability theory is irrational. In section 4, however, we will show that the same argument has been launched against Bayesian probabilism, namely that following the precepts of Bayes can result in irrational conclusions.

Van Fraassen applies that argument to IBE. According to him there are only two options: either IBE is in accordance with Bayes’ rule, namely when it implies that we have to update our beliefs in accordance with Bayes’ rule, or IBE instructs us to update our beliefs in a different manner, in particular by giving a bonus to the explanation that best explains the evidence. Van Fraassen has argued that in the former case, IBE is superfluous, in the latter case it deviates from the precepts of Bayesian probability theory and is therefore irrational.

Most adherents of IBE adhere to the first view but claim that even though IBE is consistent with Bayes’ rule, nevertheless it is not superfluous. Lipton (2004), among others, has claimed that IBE is a heuristic for a Bayesian calculus which is both fruitful and feasible. We briefly discuss possible relations between Bayesian probabilistic and explanation-based approaches in section 5.

### *C. Disjunctive explanations*

Another reason why choosing the best explanation may not lead us to the most probable conclusion is the problem of disjunctive explanations. That objection has been raised in particular against the theory of Allen and Pardo (e.g., Clermont, 2017; Nance, 2019), but if valid, it would also seem to hold against Amaya's and Van Koppen's theory. In essence, the argument is that while one explanation may be much better than any competitor, it might still be less probable than the disjunction of all competitors. For instance, suppose that there is a criminal case in which 'the defendant murdered his wife' is the best explanation. However, suppose further that the disjunction 'either the victim committed suicide or it was an accident or someone else than the husband killed her' is jointly more probable. In such cases the one-to-one comparison that explanatory accounts seem to promote could lead us in the wrong direction.

As a response to that problem, Allen and Pardo (2019b, pp. 23-28) state that nothing in their theory prohibits factfinders from considering such 'disjunctive' explanations.

### *D. Are criminal trials about comparing explanations?*

IBE is at its core a comparative notion: it presupposes that we compare the quality of different explanations. However, it has been argued that criminal trials need not be about comparing explanations. Rather, the onus is on the prosecution to prove the defendant's guilt. While the defendant may offer his own version of the facts, he does not have a duty to do so (Clermont, 2015, p. 359). Apart from the fact that the defendant has the right to remain silent, he can also attack the prosecution's case by offering counter-arguments against the scenario and the evidence for it, without offering an alternative scenario or explanation. So, it seems that IBE is doing little work. On a probabilistic view, what is going on in criminal trials is that the prosecution needs to prove that it is probable beyond a reasonable doubt that the defendant committed the act of which he is accused. That account seems to fit better with the structure of criminal trials.

One response is to claim that proving beyond a reasonable doubt in fact means that the prosecution's scenario must be demonstrated to be better than any alternative reasonable scenario (Van Koppen, 2011; Wagenaar, et al., 1993), even if the defence does not come up with an alternative of its own. In that sense every criminal trial is comparative, although it must be noted that in practice the comparison is typically left implicit.

Another response has been that we should not take the name 'inference to the best explanation' too literally. Instead, IBE is often used to describe explanatory reasoning in general. Such explanatory reasoning does not have to be comparative. For example, according to Allen and Pardo's theory, a

defendant's guilt can only be proven beyond a reasonable doubt if there is a plausible explanation implying his guilt and no plausible explanation implying his innocence. That means that factfinders do not necessarily have to look at how well explanations perform compared to one another (Pardo & Allen, 2008). However, the approach would still be explanatory since factfinders would still be involved in causal explanatory reasoning.

#### *E. Good versus probable explanations*

Explanation-based accounts use qualitative criteria to evaluate explanations. For instance, both in the descriptive story model and in the normative scenario approach, how good an explanation is, depends, among other things, on the number of evidence gaps and story gaps (Bex, 2011, pp. 91-92; Pennington & Hastie, 1993, pp. 190-199; Van Koppen & Mackor, 2019). Similarly, Allen and Pardo mention criteria such as simplicity, accordance with background beliefs and lack of ad hoc-ness as marks of a good explanation (Pardo & Allen, 2008, p. 230). However, they have not argued why these criteria are truth-conducive. In other words, the question is what reason we have to believe that the better a story is, the more likely it is to be true.<sup>22</sup> In the absence of an argument, it is unclear why such criteria should have a normative status and showing a reliable connection between these criteria and the probability of explanations has proven to be difficult. Take the notion of coherence as an example. Epistemologists have given 'impossibility results' for coherence, i.e. mathematical proofs that suggest that more coherent theories are not necessarily more probable than less coherent theories (e.g., Bovens & Hartmann, 2003; Olsson, 2005, 2019). However, others have criticized these results.<sup>23</sup>

In response to that critique, defenders of explanation-based approaches have referred to epistemologists and philosophers of science who have done work on epistemic virtues such as simplicity, robustness and coherence, showing how these virtues can be truth-conducive (Cabrera, 2017; Douglas, 2009; McMullin, 1996; Thagard, 1978). These insights can also be applied in the context of criminal law (Mackor, 2017, Dahlman & Mackor, 2019). We will briefly return to this topic in section 5.

In conclusion, the link between inference to the best explanation and the probable truth of the best explanation is problematic for various reasons. Advocates of explanation-based approaches have partially responded by amending their theories to ensure fit with probabilism (e.g. by allowing for the possibility of disjunctive explanations and elaborating on epistemic virtues). However, that does not mean that explanation-based approaches are just Bayesian approaches in disguise. It does show,

---

22. In fact, research on lying defendants shows that a false story may sometimes be more believable than a true one (Granhag & Strömwall, 2004).

23. See Schippers (2016) for a recent contribution that summarizes the debate.

however, how probabilistic considerations can be an incentive to improve explanation-based approaches. Moreover, the opposite might also be true, namely that Bayesian approaches can be improved by adding insights from explanation-based ones. We elaborate on the latter claim in the next section.

#### 4. Comparing Bayesian probabilistic and explanation-based approaches

In the previous section we discussed probabilistic critiques of explanation-based approaches. In this section we first briefly explicate the Bayesian probabilistic approach and then discuss several well-known points of critique of Bayesianism, in particular those which have been endorsed by Allen and Pardo (2019b), Amaya (2015, pp. 79 ff) and Van Koppen (2011, chapter 9).

##### 4.1. Bayesian probabilistic approach and its critiques

Bayesians claim that factfinder's evaluations of any hypothesis are degrees of belief, which express how confident the thinker is in the truth of the hypothesis (Kaye, 1979). Furthermore, they state that these degrees of belief, if rational, should obey the axioms of probability theory.

These axioms are most popularly expressed in the form of the Kolmogorov axioms which, stated informally, are:

- i. The probability of any event is equal to or greater than 0.
- ii. The probability that at least one of all the possible outcomes of a process will occur is 1.
- iii. If A and B are mutually exclusive outcomes, then the probability of either of them happening is the sum of the probability of A happening and the probability of B happening.

In this chapter we take Bayesian probabilists to be theorists who claim that when a thinker receives new evidence, he should update his degrees of belief in line with Bayes' theorem. The theorem states that the probability of a hypothesis given the new evidence (its posterior probability),  $P(H|E)$ , is a function of the likelihood of the evidence given the hypothesis,  $P(E|H)$ , the prior probability of the hypothesis,  $P(H)$ , and the marginal probability of the evidence,  $P(E)$ :

$$P(H|E) = \frac{P(E|H)*P(H)}{P(E)}$$

In its 'odds' form, in which one hypothesis ( $H_1$ ) is compared to another hypothesis ( $H_2$ ), it reads as follows:

$$\frac{P(H_1|E)}{P(H_2|E)} = \frac{P(E|H_1) * P(H_1)}{P(E|H_2) * P(H_2)}$$

Like IBE, Bayesian probabilism has been applied in legal contexts (Fenton, et al., 2016). However, its status as a normative epistemic theory, in particular for the evaluation and the assessment of legal evidence and legal proof is disputed.

The main point of critique concerns the values that Bayesians have to assign to the prior probability and to the likelihoods. The crucial question is where probabilists find these values. Roughly, two options are available. The first, and most generally chosen, option is to accept that those values are subjective, and thus are to be chosen by the individual. The other option is to search for objective values. These values might either derive from logic (for example, the probability that a fair coin will land head or tail is 0.5) or from statistics (an example would be the probability that a woman who has been killed, has been killed by her partner or ex-partner, which is said to be 0.52 (Nieuwbeerta & Leistra, 2007)). We now turn to several influential critiques of probabilism that revolve around the problem of assigning the required values.

#### **4.2. The problem of the prior**

Bayes' rule gives a way of updating one's prior probability. However, it says nothing about how to assign the initial probability, one's degree of belief in the truth of a hypothesis before evidence (Allen & Pardo, 2019b; Amaya, 2015, p. 82; Van Koppen, 2011, p. 214). For instance, it, as said, does not forbid incorporating prejudices for or against the defendant and neither does it prohibit a factfinder setting his or her prior degree of belief in the defendant's guilt to 0.9999, or even to 1. Such a factfinder could easily end up convicting the defendant, regardless of the evidence. That seems an unreasonable choice, but it would not be irrational from a probabilistic point of view which in itself allows for purely subjective degrees of belief, even though there are some rules one needs to adhere to, to make sure these beliefs follow the axioms.

Another worry about the prior has to do with the presumption of innocence. After all, if we set the prior probability of guilt at 0, we can never prove the guilt of the defendant. However, if we set the prior higher than 0, that seems to conflict with the presumption that the defendant is innocent until proven otherwise (Amaya, 2015, p. 85; Van Koppen, 2011, p. 216). There have also been discussions whether the prior probability for guilt should be set at a normatively fixed standard. However, there is no agreement what that fixed standard should be either. All suggestions made so far seem highly problematic (Dahlman, 2018; Fenton, Lagnado, Dahlman & Neil, 2017).

One response of Bayesians has been to claim that setting the prior probability is not very important. They state that, after a sufficient number of updatings, priors eventually 'wash out'. In other words,

the assignment of priors will converge on the ‘right’ posterior probability (Edwards, Lindman, & Savage, 1963). However, the problem is that such washing out requires a great deal of evidence, which we sometimes have in a mundane criminal case, but which we seldomly have in a criminal case where the evaluation of evidence is not clear-cut (Amaya, 2015, p. 83; Bunge, 2008; Dahlman, 2018; Godfrey-Smith, 2009; Van Koppen, 2011, pp. 214 ff).

#### **4.3. Absence of data**

We have seen that probabilities can be set either subjectively or objectively. Most Bayesians adhere to the view that the values of priors are set subjectively. The same seems to hold for the values that are assigned to the likelihoods. In scientific context at least some objective information will be available, but in legal cases we usually lack the required objective data to assign meaningful probabilities to either priors or likelihoods (Allen & Pardo, 2019b; Amaya, 2015, pp. 83, 85; Van Koppen, 2011, pp. 220-222, 226).

For example, when Bayesian probability theory is applied in statistical contexts, the prior is typically set using base rate information – information about how often the event that we are interested in (e.g., whether or not someone has a certain disease) occurs in the relevant population (e.g., people over 65 years old). This information might be available in e.g. medical research, but in criminal law we almost never have such base-rate information. For instance, we do not know how often defendants are falsely accused or how often witnesses lie. In the absence of such base rate information, it is unclear how the relevant probabilities can be meaningfully estimated. The same holds for the assessment of the evidence. In science we often have objective and reliable information, but in criminal cases there is rarely objective, precise, let alone quantifiable information either about the probability that evidence occurs  $P(E)$  or about the probability that the evidence occurs if the hypothesis is true  $P(E|H)$  (Van Koppen, 2011, p. 213). Take for instance the numbers attached to DNA evidence that are usually presented in rather precise terms. These numbers are based on a theoretical calculation in which important elements are left out. For instance, rates of laboratory errors and crime scene errors are not incorporated (Gill, 2016; Thompson, 2009, 2011). Furthermore, they hide severe problems with interpreting DNA mixtures (Butler, Kline, & Coble, 2018).

We have seen that one approach to tackling this problem is to say that the relevant probabilities are subjective degrees of probability, rather than frequencies. However, this solution comes with its own problems. Although these degrees of belief are constrained by the Kolmogorov axioms, these axioms hardly put any limits on the acceptable degrees of belief. This in turn means that “in the subjective-Bayesian model someone can, without being contradicted and unpunished, entertain the most outrageous prejudices” (Hofstee, 1980, p. 81, our translation). Earlier we mentioned the example of a

fact-finder who believes that the probability that a defendant is guilty is 100%. On a probabilistic account this implies that no amount of evidence can result in another conclusion. However, we would not call a factfinder who convicts all defendants, even when there is overwhelming evidence for their innocence, rational. Yet nothing in probability theory rules out such a dogmatic and irrational stance. Another response of Bayesians to the problem of lack of data has been to admit that Bayesian probabilism needs something else to assign reasonable values to the prior probability and of the likelihood. In doing so they distinguish between Bayes' theorem and a Bayesian approach to evidence which allows for more principles than Bayes' theorem. Fenton and colleagues for example (Fenton, et al., 2017; Lagnado, Fenton & Neil, 2013) propose to use specific legal idioms and causal structures. This seems to be a fruitful approach. However, one could argue that what they then do is to use an explanation-based, more specifically scenario approach, to solve the problem.

#### **4.4. Problems of feasibility and complexity**

Some versions of probabilism require calculations which are, in the practice of criminal cases, almost never feasible to do. The first problem is that humans are typically not very good at making calculations with probabilities (Saks & Thompson, 2003, pp. 338-339). For example, it turns out hard to avoid making the notorious prosecutor's or defense attorney's fallacy. Another problem, already addressed in section 4.3, is that, in criminal law, there is no knowledge or belief about evidence that allows for a precise numerical assessment. Even though on a Bayesian approach one does not model the 'facts' or evidence about the facts, but rather our beliefs about them, the problem is that we often have no idea what probability we should attach to a certain belief about a piece of evidence.

Take for instance the number accompanying DNA-evidence. Even if we would accept that DNA-evidence is an example of evidence that allows for the most precise calculations, DNA is no more than a first step in a chain of evidence. If we have a solid match between the DNA of a suspect and the DNA in a specimen from the crime scene, it just moves the discussion. It moves to the question what the meaning is of the specimen at the crime scene. Can it be accepted as coming from the perpetrator? And with what probability (Gill, 2016; Van Koppen, 2011; Thompson, 2009, 2011)? These are the problems of feasibility.

Second, the more evidence there is in a case, the more calculations are required to calculate the posterior probability of whatever hypothesis we are interested in. In fact, the number of calculations needed increases with the amount of evidence. So, computing the relevant probabilities quickly becomes unmanageable, even for experts. That is the problem of complexity (Allen & Pardo, 2019b, pp. 38 ff; Amaya, 2015, pp. 83, 86).

One response to these worries, has been the introduction of Bayesian Network approaches, which express the dependencies between hypotheses and evidence graphically (Fenton, et al., 2016; De Zoete, Fenton, Noguchi, & Lagnado, 2019) and in which computers do the calculations. These Bayesian networks (BN) may be more insightful for laymen. However, the question remains whether these networks are insightful enough for non-experts to meaningfully engage with. Thus, it does not seem helpful for a factfinder to make decisions. For example, Fenton et al. (2019, p. 20) admit that a BN model is not as easily accessible as scenario-based approaches. They also admit that they use causal structures to place constraints on the feasible range of probabilities (Fenton, et al., 2019, p. 19). Building such a network still requires a great deal of effort and expertise. So, the Bayesian network approach only solves the problem of feasibility to the extent that the time and resources are available to construct them and they do not seem to solve the problem of complexity to such an extent that they may be helpful in decision-making.

#### **4.5. Paradoxes of proof**

Another strand of criticism against probabilistic approaches to criminal evidence is that they lead to so-called paradoxes of proof. We will restrict ourselves to discussing the two most commonly mentioned paradoxes, both of which have been used as an argument in favour of explanation-based approaches. These paradoxes relate to Bayesian interpretation of the criminal standard of proof, according to which guilt has to be proven beyond a reasonable doubt. That standard is sometimes interpreted as the demand that guilt has to be proven with a probability of 0.95 or more (Connolly, 1987; Dane, 1985; Dhimi, 2008).<sup>24</sup> The two paradoxes we discuss are the conjunction paradox and the paradox of naked statistical evidence.

##### *The conjunction paradox*

The conjunction paradox was formulated by Cohen (1977, pp. 58-67) and it is one of the central arguments of Allen and Pardo (2019b) against Bayesian models of legal proof.

The paradox begins with the observation that the beyond a reasonable doubt standard requires that all elements of a crime have to be individually proven. For instance, proving a murder might mean proving both 'killing' and 'intent'. Allen and Pardo claim that the probabilistic approach therefore

---

24. What the threshold should be, has been a subject of debate for some centuries (Volokh, 1997), of which the most well-known participant probably is William Blackstone with his proposition: '[...] all presumptive evidence of felony should be admitted cautiously: for the law holds, that it is better that ten guilty persons escape, than that one innocent suffer' (Blackstone, 1765-1769, book 4, chapter 27). The threshold varies, depending on the seriousness of the crime the defendant is charged with (De Keijser & Van Koppen, 2007; Van Koppen, 2011), the attitude of factfinders (De Keijser, De Lange & Van Wilsem, 2014), how the rule is explained to a jury (Stoffelmayr & Diamond, 2000), and what legal politics dictate (Dershowitz, 1996).

implies that the factfinder has to prove that the probability of each individual element exceeds the threshold (Allen & Pardo, 2019b, p. 13). However, suppose that the threshold for proof beyond a reasonable doubt is 0.95 and that both elements are proven with 0.96 probability. Assuming independence (which is, by the way, questionable as regards the relation between an intention and the subsequent act) the probability of the conjunction of these two claims would then only be 0.92 – below the threshold of 0.95. So, although all elements are proven beyond a reasonable doubt, the probability of the crime as a whole does not sanction convicting the defendant. Pardo and Allen see that as a serious problem for probabilistic interpretations of the proof standard because it violates one of two basic premises of the legal system.

First, a fundamental goal of the proof standard is to distribute errors in a way that leads to very little false convictions (at the expense of more false acquittals). For example, if we have a threshold of 0.95 for proof beyond a reasonable doubt, we would have a maximum of 1 in 20 cases be a false conviction. However, suppose that we allow conviction based on the above conjunction. That would mean that the number of false convictions could be higher, since a case where the probability of guilt is 0.92 could also be proven beyond a reasonable doubt. Second, suppose that the probabilist claims that proof beyond a reasonable doubt instead means that the probability of the conjunction of elements has to be at least 0.95. That would allow situations where weak proof of one element is compensated by very strong proof for the other elements (as long as the probability of the conjunction is high enough). That would violate the requirement that all elements of a crime have to be proven beyond a reasonable doubt. In other words, the probabilistic interpretation of the beyond a reasonable doubt standard either violates the legal system's requirement of a just error distribution or the requirement that all elements of the crime have to be proven beyond a reasonable doubt.

However, many probabilists deny that the conjunction problem is as problematic for them as Pardo and Allen make it out to be, for instance by offering probabilistic interpretations of the proof standards that do not face the problem (e.g., Schwartz & Sober, 2017; Spottswood, 2016; Wittlin, 2019).

#### *The paradox of naked statistical evidence*

The second paradox also relates to the probabilistic explication of the proof standards (Nunn, 2015). Take the following example (adapted from the so-called Prison Riot Hypothetical of Nesson, 1979, pp. 1192-1193): there were one hundred prisoners in a prison yard and we know (from camera footage) that 99 of them have participated in killing a prison guard and that only one of them did not. However, the camera footage is too grainy to make out the identity of the killers and of the single bystander. Proper investigations have been done, but no further incriminating or exculpating evidence has been found. The paradox that Bayesians are faced with is this: given a probability of 0.99, it seems

permissible to convict one and even all of the prisoners. However, many people have objected to that conclusion because it conflicts with their intuitions.

Now suppose that instead of the camera footage a witness, whom we know to be 99% reliable, claims that he saw prisoner X killing the guard. Can we now convict prisoner X? Many believe that these two cases differ: that we can convict based on the witness evidence, but not on the camera evidence. The question is why.

Probabilists can reply in several ways to the problem of statistical evidence. They can reject common intuitions about statistical evidence as irrational. Second, they can agree with the intuition, but argue that the difference between statistical and non-statistical evidence can be accounted for within the probabilist framework. For one thing they can point out that the cases differ as regards their likelihood ratio. Interestingly, from an explanation-based perspective there also seems to be a relevant difference between the two cases, namely that in the latter case, the hypothesis that the prisoner participated in the killing seems to offer a much better causal explanation of the fact that the witness stated he saw prisoner X kill the guard, than the hypothesis that the witness was mistaken or intended to frame the prisoner. In the former case, on the other hand, it seems that the hypothesis that the prisoner is innocent can explain the camera images, and thus fact that there is a 0.99 probability that defendant participated in the killing, roughly equally well (Di Bello, 2019). So, explanation-based approaches seem to fit with common intuitions about when to convict or not based only on statistical evidence.

#### **4.6. Conclusion**

So far, we have not discussed arguments against using the Bayesian probability theory in general.<sup>25</sup> Instead, we have discussed arguments specifically directed against using the Bayesian model for evidence in criminal cases. The problem with Bayesianism in criminal cases seems to be threefold. First, in itself it lacks a way of meaningfully assigning or even delineating priors and likelihoods. It needs something else, such as causal structures and idioms, to set constraints (Fenton, et al., 2019; Lagnado, et al., 2013). Second, it may not be a psychologically feasible way of reasoning and in any case, it is less feasible than explanation-based approaches. That seems especially pressing when Bayes' rule is used to model all of the evidence in a criminal case, rather than to analyse a part of the evidence (Prakken & Meester, 2017). Third, it may lead to paradoxes of proof, although recent work on Bayesian approaches that overcome these paradoxes looks promising (de Zoete et al., 2019).

---

25. Furthermore, we have not discussed alternative probabilistic approaches. In particular we did not discuss the Dempster-Shafer theory, which has also called the theory of belief functions. See Shafer (1976) and for an application to criminal law Meester and Kerkvliet (2016).

In conclusion, it seems that both explanation-based approaches (see section 3.2) and Bayesian probabilistic approaches (this section) are confronted with several problems and the question has been raised whether the approaches can be combined to overcome some or all of the problems on both sides. We briefly discuss the relation between Bayesian and explanation-based approaches in the next and final section.

## 5. Evaluation and conclusion

In this chapter we have discussed three different explanation-based approaches to reasoning about evidence and proof in criminal trials: Allen and Pardo's relative plausibility theory, Amaya's theory of inference to the most coherent explanation and Van Koppen's scenario theory. Whereas Allen and Pardo aim to explain the American legal system, in particular the different standards of proof, Amaya and Van Koppen offer explicitly normative theories, aimed at improving legal reasoning.

We have shown that causal explanation, more specifically inference to the best explanation (IBE), is at the heart of these theories and that coherence is a core criterion to assess the quality of explanations. We have also pointed out that all three theories stay close to how people actually reason, both in everyday life and in criminal trials. Since Allen and Pardo's theory is explanatory, it obviously stays close to actual reasoning, but the same holds for Amaya's and Van Koppen's normative theories. In respect to these three issues in particular, explanation-based approaches differ from Bayesian approaches. Bayesian approaches do not aim to stay close to how people actually reason and probability rather than causal explanation and coherence is their core notion.

In this final section we briefly address the question how Bayesian and explanation-based approaches to reasoning about evidence and proof in criminal trials might relate to each other. First it should be noted that on an abstract level the structure of analysis of Bayesian and explanation-based approaches seems to be the same. They both instruct fact-finders to assess two things. A fact finder both has to assess the quality (the prior probability or the plausibility) of the hypothesis or the scenario in itself and he has to assess how probable or how well explained the evidence is in the light of the hypothesis or the scenario. However, they use different concepts (causal explanation versus probability) and explanation-based approaches propose a holistic and qualitative approach, whereas Bayesians propose an atomistic approach that is necessary to allow for a quantitative analysis.

Lipton (2004) has offered a well-known view about the relation between Bayesian and explanation-based or IBE approaches. Even though Lipton primarily focuses on the comparison of Bayes and IBE as descriptive theories, his classification is also useful for the comparison of the normative theories. In this section we confine ourselves to brief discussion of the relation between Bayesianism and explanationism as normative theories.

Lipton has distinguished three possible views of the relation between explanation-based and probabilistic approaches, namely that the Bayesian, or the explanation-based approach, or both, are incorrect as normative theories; that they are correct and compatible; or that they are correct and complementary.

In section 3.2 we discussed Van Fraassen's argument according to which the explanation-based approach is either incompatible with Bayesianism and therefore irrational or compatible but accordingly superfluous (Van Fraassen, 1989). However, several defenders of explanation-based approaches have argued that there is a middle way between irrationality and redundancy. They defend the view that Bayesianism and explanation-based approaches are not just compatible, that can exist next to each other, but are complementary, that is, they can be an aid to each other (Hitchcock 2007, Lipton 2004, Okasha 2000, Poston 2014, Psillos 2007).

On most of these complementary views, explanation-based approaches should conform to Bayes' rule. At the same time, these views hold, firstly, that epistemic virtues such as simplicity, robustness and coherence can be truth-conducive and thus that an assessment of a scenario and the relation between scenario and evidence in terms of virtues, which is deemed to be cognitively more feasible than a Bayesian analysis, can be a heuristic to a Bayesian assessment of a scenario. Next to that, these complementary views hold that explanation-based considerations can or even should be used help to determine 1) which hypotheses are to be tested and 2) which evidence is relevant, and to help to assign weights to 3) the priors and 4) the likelihoods. The general idea is that the weight of the prior can be assessed by determining how well background knowledge would explain the hypothesis and similarly that the likelihood can be assessed precisely by determining how well the hypothesis would explain the evidence.

On a weak complementary view, Bayesianism would allow for explanation-based considerations to fulfil this role. On a strong complementary view, these explanation-based considerations are indispensable and thus mandatory to turn subjective beliefs into something more objective.

This strong complementary view is closely allied to a yet stronger view, namely that explanation-based approaches are correct and Bayesian approaches are not applicable to criminal cases. In section 4 we have discussed several points of critique of applying Bayes' rule in criminal cases. One of these was that Bayes' rule does not say anything about the choice of the hypotheses and the evidence nor about the weight that can or should be assigned to them. Summarizing this view quite bluntly, it has been claimed that in itself Bayes' rule allows for 'garbage in, garbage out'. For that reason, even subjective Bayesians might want to defend the view that Bayes' rule should only be used if the priors and likelihoods can be determined with some precision and objectivity or if they can be expected to 'wash out' (Douven, 2017). In most criminal cases, however, priors and likelihoods can neither be determined with sufficient precision and objectivity, nor 'wash out'. That is an important difference between

applying Bayes' rule in a legal and a scientific context. Accordingly, this critique of Bayesianism is not so much that it fails in general as a normative approach, but rather that it should not be used in criminal cases.

Another argument for this critical view of Bayesianism is that complex cases can only be analysed with the help of Bayesian Networks to do the complex calculations. Although the fact that a network can do the calculations for us can be regarded a strength of Bayesianism, it can also be seen as a drawback since these calculations are likely to be opaque to legal fact finders. In particular in a legal context, that is a serious drawback since legal decisions should not only and not even primarily result in reaching (probable) truth, but first of all in a decision that is justified in a manner which is understandable to the court, to the parties involved and preferably also to society at large. Even though it might be acceptable that some part of a judicial decision is technical and difficult to understand, even for judges and parties of a case, this seems questionable when almost all of the decision becomes opaque to them. However, that might happen if a complete case is modelled in Bayesian terms.

Finally, on a slightly less critical view Bayes' rule should not be applied with respect to criminal cases as a whole. However, this view acknowledges that it can be used and indeed be useful with respect to those parts of the scenario and the evidence about which the priors and likelihoods can be established with sufficient precision and objectivity (in some cases DNA evidence might be an example; Prakken & Meester 2017). Moreover, on this view, a Bayesian approach can also be useful to protect fact-finders against several fallacies, such as base rate neglect, causal fallacies and the prosecutor and the defence attorney fallacy. Finally, on such a view, a comparison of an explanation-based analysis and a Bayesian Network analysis of the same case might be useful to analyse differences and agreements in the analysis. In this manner, a Bayesian analysis might work as a kind of triangulation of the explanation-based analysis.

In conclusion: explanation-based approaches stay close to how people reason in everyday life and in criminal cases, but they offer tools to reason in a structured way about evidence and proof in criminal cases and therewith to avoid certain fallacies and biases such as the confirmation bias. As regards the possible relation between Bayesian and explanation-based approaches, there does not seem to be agreement at the moment. As far as our own view is concerned, we do not reject a Bayesian approach in general. However, we agree with the critical view that Bayes' rule can be used to analyse parts of the scenario and to prevent fact finder from committing certain fallacies, but that – at least for now – it should not be used to analyse legal cases as a whole.<sup>26</sup>

---

26. The authors thank Christian Dahlman, Jaap Hage and Ronald Meester for their comments on an earlier version of this chapter.

## References

- Allen, R.J. (1997). Rationality, algorithms and juridical proof: A preliminary inquiry. *International Journal of Evidence and Proof* 1, pp. 254-275.
- Allen, R.J. & Leiter, B. (2001). Naturalized epistemology and the law of evidence. *Virginia Law Review* 87, pp. 1491-1550.
- Allen, R.J. & Pardo, M.S. (2019a). Clarifying relative plausibility: A rejoinder. *International Journal of Evidence and Proof*. Advanced online publication.
- Allen, R.J. & Pardo, M.S. (2019b). Relative plausibility and its critics. *International Journal of Evidence and Proof*. Advanced online publication.
- Amaya, A. (2009). Inference to the best legal explanation. In H. Kaptein, H. Prakken & B. Verheij (Eds.), *Legal evidence and proof: Statistics, stories, logic*. Farnham: Ashgate, pp. 135–160.
- Amaya, A. (2013). Coherence, evidence and proof. *Legal Theory* 19, pp. 1-43.
- Amaya, A. (2015). *The tapestry of reason: An inquiry into the nature of coherence and its role in legal argument*. Oxford: Hart.
- Anderson, T.J., Schum, D.A. & Twining, W.L. (2005). *Analysis of evidence* (2nd ed.). Cambridge: Cambridge University Press.
- Ask, K. & Granhag, P. A. (2005). Motivational sources of confirmation bias in criminal investigations: The need for cognitive closure. *Journal of Investigative Psychology and Offender Profiling*, 2, 43-63.
- Bennett, W. L. & Feldman, M. S. (2014). *Reconstructing reality in the courtroom: Justice and judgement in American culture* (2nd ed.). New Orleans: Quid Pro (1st ed. 1981).
- Bex, F.J. (2011). *Arguments, stories and criminal evidence: A formal hybrid theory*. Dordrecht: Springer.
- Bex, F.J., Van Koppen, P.J., Prakken, H. & Verheij, B. (2010). A hybrid formal theory of arguments, stories and criminal evidence. *Artificial Intelligence and Law* 18, pp. 123-152.
- Bex, F.J., Prakken, H., Reed, C. & Walton, D. (2003). Towards a formal account of reasoning about evidence: Argumentation schemes and generalisations. *Artificial Intelligence and Law* 11, pp. 125-165.
- Blackstone, W. (1765-1769). *Commentaries on the laws of England*. Oxford: Clarendon.
- Bovens, L. & Hartmann, S. (2003). *Bayesian epistemology*. Oxford: Oxford University Press.
- Butler, J.M., Kline, M.C. & Coble, M.D. (2018). NIST interlaboratory studies involving DNA mixtures (MIX05 and MIX13): Variation observed and lessons learned. *Forensic Science International: Genetics* 37, pp. 81-94.
- Cabrera, F. (2017). Can there be a Bayesian explanationism? On the prospects of a productive partnership. *Synthese* 194, pp. 1245-1272.

- Clermont, K.M. (2015). Trial by traditional probability, relative plausibility, or belief function? *Case Western Reserve Law Review* 66, pp. 353-391.
- Clermont, K.M. (2017). Common sense on standards of proof. *Seton Hall Law Review* 48, pp. 1057-1080.
- Cohen, L.J. (1977). *The probable and the provable*. Oxford: Clarendon.
- Connolly, T. (1987). Decision theory, reasonable doubt, and the utility of erroneous acquittals. *Law and Human Behavior* 11, pp. 101-112.
- Dahlman, C. (2018). Determining the base rate for guilt. *Law, Probability and Risk* 17, pp. 15-28.
- Dahlman, C. & Mackor, A.R. (2019). Coherence and probability in legal evidence. *Law, Probability and Risk* 18, pp. 275–294. doi:10.1093/lpr/mgz016
- Dane, F.C. (1985). In search of reasonable doubt: A systematic examination of selected quantification approaches. *Law and Human Behavior* 9, pp. 141-158.
- De Keijser, J.W. & Van Koppen, P.J. (2007). Paradoxes of proof and punishment: Psychological pitfalls in judicial decision making. *Legal and Criminological Psychology* 12, pp. 189-205.
- De Keijser, J.W., De Lange, E.G.M. & Van Wilsem, J.A. (2014). Wrongful convictions and the Blackstone ratio: An empirical analysis of public attitudes. *Punishment and Society* 16, pp. 32-49.
- De Poot, C.J., Bokhorst, R.J., Van Koppen, P.J., & Muller, E.R. (2004). *Rechercheportret: Over dilemma's in de opsporing* [Detectives' portrait: On dilemma's in police investigations]. Alphen aan den Rijn: Kluwer.
- De Zoete, J., Fenton, N., Noguchi, T. & Lagnado, D. (2019). Resolving the so-called “probabilistic paradoxes in legal reasoning” with Bayesian Networks. *Science and Justice* 59, pp. 367-379.
- Dershowitz, A.M. (1996). *Reasonable doubt: The criminal justice system and the O.J. Simpson case*. New York: Simon and Schuster.
- Dhami, M.K. (2008). On measuring quantitative interpretations of reasonable doubt. *Journal of Experimental Psychology: Applied*, 14, pp. 353–363.
- Di Bello, M. (2019). Plausibility and probability in juridical proof. *International Journal of Evidence and Proof*, advanced online publication.
- DiMaggio, P. (1997). Culture and cognition. *Annual Review of Sociology* 23, pp. 263-287.
- Douglas, H. (2009). *Science, policy, and the value-free ideal*. Pittsburgh: University of Pittsburgh Press.
- Douven, I. (2017). Abduction. E.N. Zalta (Ed.) *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/abduction/>.
- Edwards, W., Lindman, H. & Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological Review* 70, pp. 193-242.

- Fenton, N., Berger, D., Lagnado, D., Neil, M. & Hsu, A. (2014). When 'neutral' evidence still has probative value (with implications from the Barry George Case). *Science & Justice* 54(4), pp. 274-287.
- Fenton, N.E., Lagnado, D.A., Dahlman, C. & Neil, M. (2017). The opportunity prior: A simple and practical solution to the prior probability problem for legal cases *Proceedings of the 16th edition of the International Conference on Artificial Intelligence and Law*. ACM Online Library, pp. 69-76.
- Fenton, N.E., Neil, M. & Berger, D. (2016). Bayes and the law. *Annual Review of Statistics and Its Application* 3, pp. 51-77.
- Fenton, N.E., Neil, M., Yet, B. & Lagnado, D. A. (2019). Analyzing the Simonshaven case using Bayesian networks. *Topics in Cognitive Science*. Advanced internet publication.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.
- Fitelson, B. (2003). A probabilistic theory of coherence. *Analysis* 63, pp. 194-199.
- Friedman, R. D. (1997). Answering the Bayesioskeptical challenge. *International Journal of Evidence and Proof*, 1, pp. 276-291.
- Gigerenzer, G. (2007). *Gut feelings: Short cuts to better decision making*. London: Penguin.
- Gigerenzer, G. & Engel, C. (Eds.). (2006). *Heuristics and the law*. Cambridge, MA: MIT Press.
- Gill, P. (2016). *Misleading DNA-evidence: Reasons for miscarriages of justice*. Amsterdam: Academic Press.
- Godfrey-Smith, P. (2009). *Theory and reality: An introduction to the philosophy of science*. Chicago: University of Chicago Press.
- Granhag, P.A. & Strömwall, L.A. (Eds.). (2004). *The detection of deception in forensic contexts*. Cambridge: Cambridge University Press.
- Gross, S R. (1998). Lost lives: Miscarriages of justice in capital cases. *Law and Contemporary Problems* 61, pp. 125-149.
- Gross, S.R. (2008). Convicting the innocent. *Annual Review of Law and Social Sciences* 4, pp. 173-192.
- Gross, S.R., Jacoby, K., Matheson, D.J., Montgomery, N. & Patel, S. (2005). Exonerations in the United States: 1989 through 2003. *Journal of Criminal Law and Criminology* 95, pp. 523-560.
- Hans, V.P. (2008). Jury systems around the world. *Annual Review of Law and Social Science* 4, pp. 275-297.
- Harman, G.H. (1965). The inference to the best explanation. *Philosophical Review* 74, pp. 88-95.
- Hitchcock, C. (2007). The lovely and the probable. *Philosophy and Phenomenological Research* 74, pp. 433-440.
- Hofstee, W.K.B. (1980). *De empirische discussie: Theorie van het sociaal-wetenschappelijk onderzoek* [The empirical discussion: Theory of social science research]. Amsterdam: Boom.

- Huff, R. & Killias, M. (Eds.) (2008). *Wrongful conviction: International perspectives on miscarriages of justice*. Philadelphia: Temple University Press.
- Kadane, J.B. & Schum, D.A. (2011). *A probabilistic analysis of the Sacco and Vanzetti evidence*. New York: Wiley.
- Kahneman, D., Slovic, P. & Tversky, A. (Eds.) (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- Kassin, S. M., Dror, I.E. & Kukučka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition* 2, pp. 42-52.
- Kaye, D.H. (1979). The paradox of the gatecrasher and other stories. *Arizona State Law Journal*, pp. 101-109.
- Kleider, H.M., Pezdek, K., Goldinger, S.D. & Kirk, A. (2008). Schema-driven source misattribution errors: Remembering the expected from a witnessed event. *Applied Cognitive Psychology* 22, pp. 1-20.
- Lagnado, D.A., Fenton, N.E., & Neil, M. (2013). Legal idioms: A framework for evidential reasoning. *Argument and Computation* 4, pp. 46-63.
- Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge*. London: Cambridge University Press, pp. 91-196.
- Laudan, L. (2007). Strange bedfellows: Inference to the best explanation and the criminal standard of proof. *International Journal of Evidence and Proof* 11, pp. 292-306.
- Lewicka, M. (1998). Confirmation bias: Cognitive error or adaptive strategy of action control? In M. Kofta, G. Weary & G. Sedek (Eds.), *Personal control in action: Cognitive and motivational mechanisms*. New York: Plenum, pp. 233-258.
- Lipton, P. (2004). *Inference to the best explanation* (2nd ed.). London: Routledge.
- Lipton, P. (2007). Alien abduction: Inference to the best explanation and the management of testimony. *Episteme* 4, pp. 238-251.
- Lipton, P. (2009). Understanding without explanation. In H.W. De Regt, S. Leonelli & K. Eigner (Eds.), *Scientific understanding: Philosophical perspectives*. Pittsburgh: University of Pittsburgh Press, pp. 43-63.
- Lycan, W. G. (1988). *Judgment and justification*. Cambridge: Cambridge University Press.
- Lycan, W.G. (2002). Explanation and epistemology. In P.K. Moser (Ed.), *The Oxford Handbook of Epistemology*. Oxford: Oxford University Press, pp. 408-433.
- Mackor, A.R. (2017). Novel facts: The relevance of predictions in criminal law. *Strafblad* 15, pp. 145-156.
- McMullin, E. (1996). Epistemic virtue and theory-appraisal. In I. Douven & L. Horsten (Eds.), *Realism in the sciences*. Leuven: University of Leuven Press, pp. 1-34.

- Meester, R. & Kerkvliet, T. (2016). Assessing forensic evidence by computing belief functions. *Law, Probability and Risk* 15, pp. 127-153.
- Meissner, C.A., & Kassin, S.M. (2004). "You're guilty, so just confess!" Cognitive and behavioral confirmation biases in the interrogation room. In G. D. Lassiter (Ed.), *Interrogations, confessions, and entrapment*. New York: Kluwer Academic, pp. 85-106.
- Nance, D.A. (2001). Naturalized epistemology and the critique of evidence theory. *Virginia Law Review* 87, pp. 1551-1618.
- Nance, D.A. (2019). Belief functions and burdens of proof. *Law, Probability and Risk* 18, pp. 53-76.
- Nesson, C.R. (1979). Reasonable doubt and permissive inferences: The value of complexity. *Harvard Law Review* 92, pp. 1187-1225.
- Nickerson, R.S. (1998). Confirmation bias: An ubiquitous phenomenon in many guises. *Review of General Psychology* 2, pp. 175-220.
- Nieuwebeerta, P. & Leistra, G. (2007). *Dodelijk geweld: Moord en doodslag in Nederland*. Amsterdam: Balans.
- Nunn, G.A. (2015). The incompatibility of due process and naked statistical evidence. *Vanderbilt Law Review* 68, pp. 1407-1433.
- O'Brien, B.M. (2007). *Confirmation bias in criminal investigations: An examination of the factors that aggravate and counteract bias*: Unpublished diss. University of Michigan.
- Okasha, S. (2000). Van Fraassen's critique of inference to the best explanation. *Studies in History and Philosophy of Science, Part A* 31, pp. 691-710.
- Olsson, E.J. (2005). *Against coherence: Truth, probability, and justification*. Oxford: Oxford University Press.
- Olsson, E.J. (2019). Dahlman and Mackor on coherence and probability in legal evidence: a commentary. *Law, Probability and Risk* 18, pp. 295–303. doi:10.1093/lpr/mgz017
- Oswald, M.E., & Grosjean, S. (2004). Confirmation bias. In R.F. Pohl (Ed.), *Cognitive illusions: A handbook on fallacies and biases in thinking, judgment and memory*. Hove: Psychology Press, pp. 79-96.
- Pardo, M.S. (2013). The nature and purpose of evidence theory. *Vanderbilt Law Review* 66, pp. 547-613.
- Pardo, M.S. (2019). The paradoxes of legal proof: A Critical Guide. *Boston University Law Review* 99, pp. 233-290.
- Pardo, M.S. & Allen, R.J. (2008). Juridical proof and the best explanation. *Law and Philosophy* 27, pp. 223-268.

- Pennington, N. & Hastie, R. (1993). The story model for juror decision making. In R. Hastie (Ed.), *Inside the jury: The psychology of juror decision making* (2nd ed.). Cambridge: Cambridge University Press, pp. 192-221.
- Popper, K.R. (1935. In reality 1934) *Logik der Forschung: Zur Erkenntnistheorie der modernen Naturwissenschaft*. Wien: Springer.
- Popper, K.R. (1963). *Conjectures and refutations: The growth of scientific knowledge*. London: Routledge & Keagan Paul.
- Poston, T.L. (2014). *Reason and explanation: A defense of explanatory coherentism*. New York: Palgrave Macmillan.
- Prakken, H. & Meester, R. (2017), Bayesiaanse analyses van complexe strafzaken door deskundigen. Betrouwbaar en zo ja: nuttig? [Bayesian analyses of complex criminal cases by experts. Reliable and if so: useful?] *Expertise en Recht* 5, pp. 185-197.
- Prakken, H. (2019). An argumentation-based analysis of the Simonshaven Case. *Topics in Cognitive Science*, advanced online publication.
- Psillos, S. (2007). The fine structure of inference to the best explanation. *Philosophy and Phenomenological Research* 74, pp. 441-448.
- Saks, M.J. & Thompson, W.C. (2003). Assessing evidence: Proving facts. In D. Carson & R.H.C. Bull (Eds.), *Handbook of psychology in legal contexts* (2nd ed). Chichester: Wiley, pp. 329-345.
- Shafer, G. (1976). *A mathematical theory of evidence*. Princeton: Princeton University Press.
- Schippers, M. (2016). The problem of coherence and truth redux. *Erkenntnis* 81, pp. 817-851.
- Schum, D.A. (1994). *The evidential foundations of probabilistic reasoning*. New York: Wiley.
- Schwartz, D.S. & Sober, E. (2017). The conjunction problem and the logic of jury findings. *William and Mary Law Review* 59, pp. 619-692.
- Schweizer, M. (2014). Comparing holistic and atomistic evaluation of evidence. *Law, Probability and Risk* 13, pp. 65-89.
- Shogenji, T. (1999). Is coherence truth conducive. *Analysis* 59, pp. 338-345.
- Siebel, M. (2011). Why explanation and thus coherence cannot be reduced to probability. *Analysis* 71, pp. 264-266.
- Simon, D. (2004). A third view of the black box: Cognitive coherence in legal decision making. *University of Chicago Law Review* 71, pp. 511-586.
- Simon, D. (2019). Thin empirics. *International Journal of Evidence and Proof*, advanced internet publication.
- Spottswood, M. (2016). Unraveling the conjunction paradox. *Law, Probability and Risk* 15, pp. 259-296.

- Stein, A. (1996). Judicial fact-finding and the Bayesian method: The case for deeper scepticism about their combination. *International Journal of Evidence and Proof*, 1, pp. 25-47.
- Stoffelmayr, E., & Diamond, S. S. (2000). The conflict between precision and flexibility in explaining beyond a reasonable doubt. *Psychology, Public Policy, and Law*, 6, pp. 769-787.
- Taylor, S.E., & Crocker, J. (1981). Schematic bases of social information processing. In E.T. Higgins, C.A. Herman & M.P. Zanna (Eds.), *Social cognition: The Ontario Symposium on Personality and Social Psychology*. Hillsdale, NJ: Erlbaum, pp. 89-134.
- Thagard, P.R. (1978). The best explanation: Criteria for theory choice. *The Journal of Philosophy* 75, pp. 76-92.
- Thagard, P.R. (2000). *Coherence in thought and action*. Cambridge, MA: MIT Press.
- Thompson, W.C. (2009). Painting the target around the matching profile: The Texas sharpshooter fallacy in forensic DNA interpretation. *Law, Probability and Risk* 8, pp. 257-276.
- Thompson, W.C. (2011). What role should investigative facts play in the evaluation of scientific evidence? *Australian Journal of Forensic Sciences* 43, pp. 123-134.
- Tillers, P. & Green, E.D. (Eds.). (1988). *Probability and inference in the law of evidence: The uses and limits of Bayesianism*. Dordrecht: Kluwer Academic.
- Van Fraassen, B.C. (1989). *Laws and symmetry*. Oxford: Oxford University Press.
- Van Koppen, P.J. (2009). Jury trials: Oposed. *E-Journal USA* 14(7), pp. 18, 23-24.
- Van Koppen, P.J. (2011). *Overtuigend bewijs: Indammen van rechterlijke dwalingen* [Convincing evidence: Reducing the number of miscarriages of justice]. Amsterdam: Nieuw Amsterdam.
- Van Koppen, P.J. (2013). *Gerede twijfel: Over bewijs in strafzaken* [Reasonable doubt: On evidence in criminal cases]. Amsterdam: De Kring.
- Van Koppen, P.J., & Mackor, A.R. (2019). A scenario approach to the Simonshaven case. *Topics in Cognitive Science*. Advanced online publication.
- Vidmar, N.J. (2009). Jury trials: In favor. *E-Journal USA* 14(7), pp. 18-22.
- Vidmar, N.J. (2011). The psychology of trial judging. *Current Directions in Psychological Science* 20, pp. 58-62.
- Vidmar, N.J., & Hans, V.P. (2007). *American juries: The verdict*. Amherst, NY: Prometheus.
- Vineberg, S. (2016). Dutch Book Arguments. In E.N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2016). Metaphysics Research Lab, Stanford University. Retrieved from <https://plato.stanford.edu/archives/spr2016/entries/dutch-book/>.
- Vlek, C.S. (2016). *When stories and numbers meet in court: Constructing and explaining Bayesian networks for criminal cases with scenarios*. Groningen: Unpublished doctoral dissertation Groningen University.

- Vlek, C.S., Prakken, H., Renooij, S. & Verheij, B. (2016). Building Bayesian networks for legal evidence with narratives: A case study evaluation. *Artificial Intelligence and Law* 24, pp. 285-324.
- Volokh, A. (1997). n Guilty men. *University of Pennsylvania Law Review* 146, pp. 173-211.
- Wagenaar, W.A., Van Koppen, P.J. & Crombag, H.F.M. (1993). *Anchored narratives: The psychology of criminal evidence*. London: Harvester Wheatsheaf.
- Walton, D.N., Reed, C.A. & Macagno, F. (2008). *Argumentation schemes*. Cambridge: Cambridge University Press.
- Williamson, J. (2010). *In defence of objective Bayesianism*. Oxford: Oxford University Press.
- Williamson, T. (2016). Abductive philosophy. *Philosophical Forum* 47, pp. 263-280.
- Wittlin, M. (2019). Common problems of plausibility and probabilism. *International Journal of Evidence and Proof*, online advanced publication.
- Ylikoski, P. & Kuorikoski, J. (2010). Dissecting explanatory power. *Philosophical Studies* 148, pp. 201-219.

PREPRINT