# University of Groningen

## The values of prediction in criminal cases

Jellema, Hylke

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*
Publisher's PDF, also known as Version of record

[Link to publication in University of Groningen/UMCG research database](Link to publication in University of Groningen/UMCG research database)

# The values of prediction in criminal cases

## Hylke Jellema

University of Groningen, Groningen, Netherlands

## Abstract

Like scientists, investigators and decision-makers in criminal cases both explain known evidence and use the resulting explanations to make novel predictions. Philosophers of science have made much of this distinction, arguing that hypotheses which lead to successful predictions are—all else being equal—epistemically superior to those that merely explain known data. Their ideas also offer important lessons for criminal evidence scholarship. This article distinguishes three values of prediction over explaining known facts in criminal cases. First, witnesses who predict are—all else being equal—more reliable than those who do not because they are less likely to be biased or lying. Second, investigators who only explain known facts run the risk of 'fudging' the scenarios that they formulate. Predictions can protect us against this danger. Third, carefully constructed predictions may help investigators to avoid confirmation bias. This article ends with a case study of the murder of Hae Min Lee.

## Introduction

Like scientists, investigators and decision-makers in criminal cases both explain known evidence and use the resulting explanations to make new predictions. For instance, imagine that a detective suspects someone of being the perpetrator of a criminal act because he was near the scene of the crime. Alternatively, imagine that this detective comes to suspect him for some other reason and only later finds out that this person was near the crime scene. In both cases the detective ends up formulating the scenario that this person was the perpetrator. Furthermore, in both cases she has at least one piece of evidence supporting this scenario—namely that this person was near the crime scene. Yet this piece of evidence comes to support the scenario in different ways. In the first case, the evidence was *accommodated*—the detective formulated the scenario based on this known piece of evidence. In the second case, the evidence was *predicted*—the detective did not use it in constructing the scenario.

**Corresponding author:**
Hylke Jellema, University of Groningen, Faculty of Law, Oude Kijk in 't Jatstraat 26 Groningen, 9700 AB Netherlands.
E-mail: h.jellema@rug.nl

Philosophers of science have made much of the distinction between prediction and accommodation. In science, successful predictions are often seen as one of the hallmarks of a good theory and scientific theories that make no testable predictions are often seen as defective—even if they explain the known data. Because of this, many philosophers of science argue for *predictivism*—the thesis that successfully predicted facts provide stronger evidence for a theory than successfully explained known facts. In this article I argue that the predictivist debate from the philosophy of science can teach us valuable lessons about criminal evidence too.

In criminal law, predictions do not derive from scientific theories, but from *scenarios*—which are accounts of what happened in a case. For instance, in a murder case a typically scenario details who killed the victim, how they did so, and why they did it. Such scenarios both explain known evidence and produce further evidence that we might expect to find. Though it may happen less than in science, we also encounter predictivist intuitions regarding the predictions of scenarios. For instance, Josephson (2000) claims that checking whether a crime scenario's predictions are confirmed is one of the most important criteria on the basis of which we should assess such scenarios. The necessity of checking predictions is also repeatedly emphasised by van Koppen (2011: e.g. 76)—who mentions their importance in science as an illustration of why they are important in criminal cases too. A more explicit appeal to predictivism is made by Mackor (2017), who adopts a Lakatosian position on the value of predictions. She draws an analogy between the role of predictions for assessing scientific research programmes and their role in assessing scenario's. However, though she briefly alludes to the contemporary predictivist debate, she does not discuss the arguments from this debate in detail. Finally, Tuzet (2019) *does* discuss one argument from the contemporary predictivist debate, namely Peter Lipton's fudging argument.[1] Based on this argument he suggests that predictivism also holds with respect to legal evidence. Nonetheless, his discussion of this point is brief and in need of further elaboration. So, at least some authors consider predictions to have a value above and beyond a scenario merely accommodating the known facts. Nonetheless, we currently lack a systematic examination of what this value might be.

The goal of this article is to address this gap. I examine several existing arguments about the value of prediction developed by philosophers of science and show that analogous arguments can be made with respect to criminal cases. In particular, this article develops three arguments for the special value of predicted over accommodated criminal evidence. First, witnesses who predict are—all else being equal—more reliable than those who do not because they are less likely to be biased or lying. Second, investigators who only accommodate run the risk of 'fudging' the scenarios that they formulate: they sacrifice the quality of a scenario to make it fit the facts. Predictions can protect us against this danger. Third, carefully constructed predictions may help investigators to avoid confirmation bias. In order to show how these arguments play out in real, complex situations, I end this article with a case study of the murder of Hae Min Lee. This murder was the subject of the popular 2014 true crime podcast *Serial*.

## On predictions and evidential strength

This article examines various arguments for *predictivism*—the thesis that successful predictions yield stronger evidence than successful accommodations. As a preliminary, let's look at the notions of prediction, accommodation and evidential strength.

---

1. I discuss this argument under 'Fudged scenarios'.

## Predictions and accommodations

The term 'prediction' has a subtly different meaning in the context of the predictivism debate than how we normally use the term. When we talk about predictions in daily life, we typically mean that someone made an explicit statement about the future. For example, someone might say 'I predict that it will rain tomorrow'. However, while that is one example of a prediction, not all predictions are like that. First, predictions do not have to be explicitly stated by a person. Second, they do not have to be about the future. Predictions are empirical consequences that derive from hypotheses. In other words, if we assume that a given hypothesis is true, this creates certain expectations about the kind of evidence we should encounter, that we would not otherwise encounter. In science, such hypotheses often take the shape of models or theories—which in turn lead to predictions. In criminal trials the relevant hypotheses are often *scenarios*. We can think of such scenarios as stories about what happened in a case. For instance, imagine a murder case where a woman's dead body is found and where the woman's husband is the main suspect. The relevant scenario is then 'the husband killed his wife'. Certain facts will be more likely to be true if the scenario is true—these are the empirical consequences of that scenario. For instance, if the husband killed his wife, this implies that the wife is indeed dead.

However, not every empirical consequence counts as a prediction as I'm using the term here. In the predictivism debate, the term 'prediction' is often used as a shorthand for 'the prediction of a *novel fact*'. A fact is novel if it was not used to construct the explanation.[2] For instance, suppose that the husband became a suspect because he repeatedly threatened his wife in the past. The husband's threats are empirical consequences of the scenario (if we assume that they are more likely to have happened if the husband did indeed kill her than if he did not). But they are not *novel*, because they were used to arrive at the hypothesis that he is the killer. Facts that are used in the construction of a hypotheses are called *accommodations*. In contrast, if investigators did not know about the husband threatening his wife when they formulated the scenario that he killed her, then these threats are a novel fact. The discovery of these threats means that the scenario was confirmed by a successful prediction.

## Evidential strength and risky predictions

There are various ways in which we can spell out the notion of evidential strength. This article uses one of the most common approaches, namely the Bayesian notion of the likelihood ratio: How much a piece of evidence confirms a hypothesis is directly proportional to how expected that piece of evidence is given the truth of the hypothesis divided by how expected it is given the falsehood of that hypothesis.[3] For instance, how expected is it that we would find a defendant's fingerprints on the murder weapon if he killed the victim and how expected is this if he did not kill her? If predictions provide stronger evidence than accommodations, we should somehow be able to spell this out in terms of the likelihood ratio.

Let me finally turn to a closely related term that is worth spelling out for the ensuing discussion, namely 'riskiness'. A prediction that—if confirmed—would provide strong evidence for a hypothesis (in terms of the likelihood ratio) is a *risky prediction*. Such a prediction is likely to come true if the scenario from which it follows is true, but unlikely to come true if the scenario is false. For instance, take again the scenario that the husband was the killer. An example of a non-risky prediction resulting from this hypothesis is that we would find the husband's DNA on the victim. This prediction is likely to come true regardless of whether the husband actually is the killer. After all, partners often have physical

---

2. This is called the 'heuristic' or 'use novel' conception of novelty. The heuristic conception is by far the most widely used conception of novelty in contemporary philosophy of science (Barnes, 2018). Other notions of novelty also exist. The most well-known of these is the temporal conception, which defines a novel fact as one that was unknown to the scientific community at the time when the relevant theory was created. However, the temporal conception no longer has defenders within the philosophy of science for reasons that I will not go into here (Harker, 2008: 431).

3. Formally, $P(E \mid H) / (P(E \mid \neg H)$

contact, even when they are not killing one another. Contrast this with a riskier prediction. For example, suppose that the victim died from being beaten to death. We can then predict that if the husband did indeed kill her, he is likely to have sustained injuries to his fists. This is a risky prediction because such wounds are likely if he killed his wife by beating her to death (such an act often leaves injuries to fists) but unlikely if he did not kill her. After all, most of us do not typically have injuries on our hands.[4]

In contrast to predictions, accommodations can never be risky. After all, the definition of riskiness implies that the fact could turn out to be wrong. But in the case of accommodation we explain facts that we know to be true. Some have tried to ground predictivism in this observation (see Mayo, 1991 for a discussion). But we should be careful not to confuse riskiness with evidential strength. Accommodations can still yield strong evidence for a hypothesis, even if they are not risky. After all, the accommodated fact can have a high likelihood ratio. Consider the above example again. Suppose that the husband became a suspect in the killing of his wife *because* his hands were injured. This would be a case of accommodation, but this does not necessarily change how expected this evidence is given either the truth or falsehood of the hypothesis that the husband is the killer. So we need some further argument for why predictions yield stronger evidence than accommodations. In the remainder of this article I offer three such arguments.

## Predicting witnesses

The first benefit of prediction over accommodation concerns the reliability of witnesses. Witness testimony is one of the most important types of criminal evidence. It includes the testimony of the defendant, eyewitnesses, character witnesses and expert witnesses, such as forensic scientists and psychiatrists. The testimony of these witnesses can be supported by predicted and accommodated facts. However, I want to argue that testimony which is confirmed by predictions is—all else being equal—more strongly confirmed than testimony that is confirmed only by accommodations. The reason for this is that successful predictions make it less likely that the witness is biased or lying.

To begin with an example, imagine that a bank robbery was committed. Marcy, an eyewitness who was inside the bank during the robbery, testifies that: 'The robber was a bald man in a red sweater with a big scar across his face.' Furthermore, suppose that investigators obtain camera footage from the neighbourhood surrounding the bank. On the footage a man—Luke—can be seen a short distance from the bank, ten minutes before the robbery took place. Luke is bald, has a big scar across his face and is wearing a red sweater in the footage.

That the footage showed someone matching Marcy's description of the robber close to the bank, shortly before the robbery occurred, seems to provide strong support to her testimony. When cast in terms of the accommodation/prediction distinction, this counts as a successful, risky prediction. After all, Marcy's testimony was not based on the camera footage. Furthermore, the likelihood ratio of the evidence is high—which makes the prediction risky. To see why, imagine that Marcy's description of the robber is entirely wrong. In that case it is very unlikely that we would observe a bald man in a red sweater and a scar over his face near the bank during that time. But if her description is accurate, we would expect the robber to have been near the bank around the time that the robbery occurred.

However, now suppose that Marcy accommodated the camera footage. Recall that an accommodated fact is one that is used in the construction of a hypothesis. For instance, the police might have first shown Marcy the footage and she could have filled in the details of her hazy memory based on Luke's appearance in the camera recordings. If we know that this is the case, we should assign less credibility to Marcy's testimony than in the case where she made a risky prediction. In particular, the likelihood

---

4. Of course, such a prediction is not risky if there is a plausible alternative explanation for why he has injuries to his fists that implies his innocence. For instance, if the husband is a professional, bare-knuckled boxer, such injuries are expected, even if he did not kill his wife. Whether a prediction is risky therefore also depends on the specifics of the case and particularly on the competing scenarios involved in it.

ratio of the testimony is then lower, because we would expect Marcy to report that the robber was a bald man, with a red sweater and a scar over his face, even if this was not true.

Obviously, whether a witnesses' testimony was influenced by certain information is usually a question that we seek to answer, not something that we can know with certainty. Nonetheless, the surer we are that certain facts were predicted, the more these facts confirm the testimony. If we are certain that a witness did not use specific information in the construction of her testimony, then this information confirms her testimony more strongly than if we suspect that her testimony only fits with these facts *because* she fitted her testimony to them.

To give another example, suppose that a suspect is guilty. It can then be in his best interest to call upon his right to remain silent and only offer an alternative scenario at a late stage, when all the evidence has been presented in court. Or he can continuously change his story to fit any counterevidence. In either case, he fits his testimony to the evidence in ways that make it seem well-supported when it is not. Because criminal evidence often deteriorates quickly, it may be impossible to further test such a scenario by gathering new data. Furthermore, the fact that this scenario is constructed to fit with the evidence can mean that it can be more coherent (Vredeveldt et al., 2014) and better supported by the evidence (Gunn et al., 2016) than a true explanation. After all, the suspect then knows what evidence to accommodate. However, if a suspect offers an explanation that yields checkable predictions then this alleviates the worry that he is lying by fitting his story to the available evidence (Jellema, 2019).

The above is an example of consciously fitting one's testimony to the evidence. However, such this process will often happen subconsciously. For instance, it is well known that the reliability of witness memories can be negatively influenced by 'post-event information'—i.e. information provided to a witness after her observation (Zaragoza et al., 2007). To give an example, police providing information about a potential suspect to a witness is a well-known cause of witnesses identifying the wrong suspect (Wise et al., 2014). Such post-event information then plays a causal role in the construction of the testimony and hence counts as an accommodation—one that lowers the witness's reliability.

What goes for eyewitnesses also goes for expert witnesses. Experts are prone to make erroneous judgments when they receive too much information about a case. This is known as 'contextual bias' (Kassin et al., 2013; Thompson, 2017). Receiving information about a case can lead experts to develop expectations about the outcome of an examination. For instance, fingerprint examiners are less likely to report a match between a latent print from a crime scene and a suspect when they are told that the suspect had a solid alibi (Dror et al., 2006). Similar types of contextual bias occur in several other types of forensic science, such as document examination, bite mark analysis, bloodstain pattern analysis, forensic anthropology and DNA analysis (Thompson, 2017). Doing the initial assessment without knowledge of the target eliminates such potential influences.

The point is not that explaining known information inherently makes a witness less reliable. In the case of eyewitness testimony, there is at least one type of data that we *do* want the witness to fit their testimony to—namely their own sensory observations during the events that they are reporting on. Similarly, in the case of the expert witness, there will be certain facts that the expert *should* base her testimony on. However, other information should be excluded from the expert's judgment. If such information did not play a causal role in the construction of the expert testimony, but does support her judgment, then it is stronger evidence than if it *did* play such a causal role. Consider a fingerprint examiner who did not know that the suspect had an alibi and who reports a non-match between a print found on the crime scene and that of the suspect. Contrast this to a second expert who did know about the alibi. In the second case, the alibi might have been a subconscious reason for the examiner to come to the conclusion that the fingerprints did not match. So we should assign a higher reliability to the former expert's conclusion than to the latter upon learning that the witness had an alibi.

My claim is similar to that of various philosophers of science who argue that predictions tell us something about the competence of a scientist (e.g. Barnes, 2008; Kahn et al., 1996; Maher, 1988). For instance, Kahn et al. (1996) claim that scientists may choose to predict because they have confidence in their abilities, whereas those who accommodate may do so because they lack such confidence. They

argue that predictions therefore tell us something about the *private knowledge* of the scientist, meaning the non-public reasons that the scientist has for making his predictions.

Some authors regard such arguments with suspicion. For example, as Lipton (2005) points out, in science we are after the evaluation of hypotheses, not the evaluation of scientists. Therefore, even if predicting scientists tend to be more reliable theory-constructors than those who accommodate, the fact that they predicted should ideally not influence how we evaluate their *theories*. Nonetheless, even if we grant this point in the context of science, when it comes to witness evaluation, we obviously *are* interested in how reliable they are as sources of testimony. Such witnesses typically testify during the investigation or the trial *because* decision-makers expect them to have knowledge that is not easily accessible to others. This can either be because the witness is attesting to their personal experiences or because they report on their field of expertise. In such situations, decision-makers often cannot readily evaluate the credibility of the claims made by the witness. Instead, they have to evaluate their credibility as a person—i.e. how reliable such a witness is.

## Fudged scenarios

The argument from the previous section established that witnesses who predict are—all else being equal—more reliable than those who do not. However, the scope of this argument is limited. In the philosophy of science, predictivism is the position that we should assign a higher degree of confirmation to scientific theories that successfully predict certain facts than to theories that only explain known facts. This epistemic advantage is not because we consider the scientists who formulated the theory to be better, but because we believe that the theory is better supported by the evidence in some sense. The analogous claim in criminal cases would be that a scenario which is confirmed by predictions is—all else being equal—more strongly confirmed than one that only explained known data. Let's turn to an argument supports this broader claim.

My argument is an analogous application of Peter Lipton's (2005) 'fudging' argument for predictivism from the philosophy of science. This argument connects accommodation to a specific kind of biased hypothesis construction. According to Lipton, scientists who only explain known evidence when they formulate a hypothesis can be prone to 'fudging their theories', i.e. proposing weak hypotheses that explain as much of the evidence as possible. A fudged hypothesis seems well-supported by the evidence, but this support is an illusion. I propose that fudging is also a danger in criminal cases and that the act of prediction can protect against this danger. But before I argue for this, let's first look at Lipton's argument in the context of science.

When scientists formulate a hypothesis, they will usually try to make it fit with as much of the available data as possible. This is generally a good approach to hypothesis construction. However, fitting a hypothesis as closely as possible to the existing data is not *always* a reliable method. In particular, such a fit may come at the cost of weakening the hypothesis. As Lipton and others, such as Lange (2001), note, explaining known data can occasionally lead to scientific theories that lack coherence, are overly complex or fit poorly with our background assumptions. Such theories explain the evidence well, but lack internal plausibility. To illustrate, Lipton gives the example of Ptolemaic astronomy, which had to accommodate contradicting astronomical data by adding more and more complex epicycles to the theory.[5] In such cases, the scientist is overly focused on making sense of the evidence and as a result (subconsciously) makes 'unnatural' modifications to her theory. To put it differently, in order to have it fit with the data, she changes the hypothesis in ways that make it less internally coherent, more complex and/or decrease the theory's fit with our background knowledge.

---

5. Lipton's example may not be well chosen, however. For example, Kuhn (1957: 169) argues that Copernicus' system is neither simpler nor more accurate (observationally) than the Ptolemaic system.

Like scientists, when criminal investigators formulate a scenario, they typically also want this scenario to fit as closely to the evidence as possible. After all, most people consider a scenario that explains more of the evidence to be better than one that explains less (Pennington and Hastie, 1991; Pardo and Allen, 2008). But the danger of fudging lurks here as well. The goal of fitting the scenario to the data can lead to sacrifices in the quality of that scenario. There are several reasons why such fudging might occur. Lipton mostly refers to situations where the data is varied—which an overly complex theory can easily account for. However, I believe that there are at least two other causes of fudging that are more relevant for criminal investigations.

The first of these causes is the attachment to existing hypotheses in the face of contradicting evidence. For instance, Lipton's own example of Ptolomaic astronomy is strictly speaking not about heterogeneous data, but about *unwanted* data. The reason that this theory became overly complex is because it had to explain away problematic data that contradicted the theory. Similarly, when investigators become attached to a specific explanation they sometimes keep amending that explanation to explain away further, conflicting data—at the cost of making that scenario more incoherent or fit less well with certain background assumptions about how the world works. For example, a contradicting witness statement might be explained away by arguing that this witness is lying for no reason.

Another possible source of fudging is that in criminal cases, some of the data may be unreliable or irrelevant. Trying too hard to fit one's scenario to such data can result in a weaker scenario. Consider a situation in which multiple eyewitnesses report on a robbery. We know that eyewitnesses commonly misremember many of the details of any event (Wise et al., 2014). So the stories that these witnesses tell will often diverge both from the truth and from each other in many ways. However, suppose that we want our scenario to match up exactly with the available testimony. This may require a complex narrative of how the robbery went down. For instance, if witnesses misreport the precise timing and location of the events, then a narrative that fits with their testimony is likely to contain illogical jumps in time (see, for instance, the prosecution scenario in the Hae Min Lee case below).[6]

The problem with fudging is that it is sometimes difficult to notice. As Lipton (2005: 221) points out, evaluating whether a hypothesis was fudged is always not always straightforward: '[T]his may be to exaggerate scientists' abilities or equivalently to underestimate the complexity of the factors that determine the degree to which the hypothesis is supported by data.' Similarly, it is not always straightforward to determine how good a crime scenario is—especially in hard cases. Is a given scenario overly complex, does it fit poorly with our background assumptions? These questions may not be easy to answer. For instance, the case study at the end of this article illustrates how all involved in a criminal trial may fail to spot fudging and that careful analysis may be needed to bring it to light (see 'Fudging and the call record').

This brings us to the argument for predictivism – i.e. the benefit of prediction over accommodation. In science, fudging happens because the scientist has an incentive to fit her theory to the data. As Lipton (2003: 170) puts it: '[t]he scientist knows the answer she must get, and does whatever it takes to get it.' However, this incentive is not present when the scientist makes predictions. After all, in that case she does not know what the outcome of the experiment will be and cannot fit her hypothesis to it. So, if the hypothesis and the evidence fit well together, this is less likely to be due to fudging. The conclusion of this argument is that, all else being equal, a hypothesis that predicted a fact is therefore better supported than a hypothesis that accommodated that same fact. Similarly, in the criminal trial context, fudging may happen because investigators know the evidence that they want to explain and may sacrifice the quality of their scenario to achieve this. Therefore, if we know that a fact was used in the construction of the

---

6. This second cause is similar to the point made by philosophers of science Hitchcock and Sober (2004). They link accommodation to the worry of 'overfitting'—where a scientist wants to fit her model too closely to noisy data, thereby making it overly complex. While they mainly use this in the context of statistical curve-fitting, they argue that the same notion can be applied to scientific theories in general.

scenario then there is a possibility that the scenario and the fact only fit due to fudging. But now suppose that we know that a fact was *not* used in this construction process. The investigator then did not have an incentive to fudge—i.e. to bend the scenario until it fit the fact. And therefore, we can be relatively certain that the fit between the scenario and the fact offers genuine support to that scenario. After all, they did not know all of the evidence in advance that their scenario had to explain.

Of course, whether a fact was predicted is sometimes also difficult to judge. But suppose that we are in a situation where we know with a high degree of certainty that a scenario made successful, risky predictions. In such situations the evidence provides strong support for the scenario. In contrast, if a scenario only accommodates known data, we have more reason to worry about fudging. Therefore, we should assign a higher degree of probability to this scenario. Not because predictions are better in some logical sense than accommodations. Rather, they are indicative of another epistemic virtue—namely non-fudging.[7]

## Fudged evidence, confirmation bias and the argument from choice

Let's now turn to the final benefit of prediction over accommodation, one that is similar to the benefit described in the previous section. The fudging argument that I just discussed concerned investigators fitting their scenario to the evidence. But the opposite is also possible—criminal investigators sometimes consciously or subconsciously fit their *evidence* to a preconceived scenario by selectively gathering or interpreting it. In other words, they *fudge their evidence*. Just like a fudged scenario, this may lead to the illusion of the evidence supporting a scenario well. I want to argue that checking a scenario's prediction is an important tool in preventing fudged evidence.

Fudged evidence is a well-known problem in criminal cases. Criminal investigators may fit the facts to a preconceived scenario in ways that leads to weak support. In those situations, they are blind to some of the evidence that the scenario should explain, or they misinterpret evidence to make it fit with the scenario. Such fudging may happen intentionally. However, it usually happens subconsciously, in which case we call it *confirmation bias*—'the unconscious tendency to seek out, select, and interpret new information in ways that validate one's pre-existing beliefs, hopes, or expectations' (Nickerson, 1998). Confirmation bias is pervasive in all of human affairs. Yet it is especially dangerous in criminal cases, where it is one of the leading causes of judicial errors (Gross et al., 2004). Investigators, judges and jurors are typically not aware of their own confirmation bias (Kassin et al., 2013). In criminal cases, we often encounter such bias in the form of *tunnel vision*—'a rigid focus on one suspect that leads investigators to seek out and favour inculpatory evidence, while overlooking or discounting any exculpatory evidence that might exist' (Findley and Scott, 2006).

In this article, I defend the epistemic superiority of prediction. However, the possibility of fudged evidence might—at first sight—seem to throw a spanner in the works of this project. After all, the fudging argument for predictivism given in the previous section relies on the observation that first gathering the facts can bias the way we then construct a hypothesis based on those facts. When we predict, the process is reversed: we first formulate a hypothesis and then gather the facts. But this provides an incentive to fudge the evidence so that it supports our preconceived hypothesis. In both cases the danger is that what comes first in the process may bias our second step. If this reversed fudging argument holds, it seems to support the notion that accommodation is superior to prediction when it comes to preventing fudged evidence. So let me begin by addressing this potential objection, before I move on to the value of prediction in counteracting confirmation bias.

---

7. This in contrast with the argument from last section, about witness evidence. Given that the argument from that section holds, there is a logical difference between predicted and accommodated evidence. In other words, predictions should make even a perfectly rational evaluator assign a higher degree of reliability to the witness.

Admittedly, there is some truth to the aforementioned argument: if we are sufficiently certain that the accommodating investigator gathers her evidence without *any* preconceptions, then we can safely assume that she did not fudge the evidence. In such a situation, there is no preconceived hypothesis to fit the evidence to. In contrast, a predicting investigator will have a motive to fudge their evidence much in the same way that an accommodating investigator will have a motive to fudge their scenario. However, the problem with this argument is that it relies on an unrealistic view of accommodation. The process of explaining known data can involve preconceptions just as much as the process of prediction. Even if we are creating a scenario from scratch, preconceptions do not need to take the form of a fully formed scenario. Investigators can also have vague suspicions and implicit biases. This may lead them to ignore some of the facts or to interpret them in a biased way when accommodating. To give an example, investigators may construct their scenario around the testimony of a select group of witnesses who agree on how a certain a certain event took place, while ignoring the testimony of other witnesses, who offer a different version of the events.

There is, therefore, no fundamental difference between prediction and accommodation in this regard. In both cases the danger of confirmation bias lurks. The above argument does not establish that accommodation is, in some sense, superior over prediction. Instead, confirmation bias poses a problem for both predictors and accommodators. Apparent success at explaining known data or at predicting novel facts may be due to the facts having been fitted to the explanation, which could lead us to overvalue how much our evidence supports a scenario. So, regardless of whether we accommodate or predict, we want some assurance that the evidential support for the scenario is genuine and not the result of fudged evidence.

There are different ways to counteract confirmation bias. For instance, investigators can (and arguably should) consider multiple scenarios during a case (O'Brien, 2009; van Koppen, 2011). This prevents them from becoming overly focused on a single possibility. However, I want to emphasise a different method here. To prevent confirmation bias, investigators can adopt a *falsificationist mindset*— where they explicitly seek out evidence that might disconfirm their scenario, rather than implicitly trying to confirm that scenario (Nickerson, 1998; van Koppen, 2011). When we adopt such a mindset, we ask the question: 'suppose that our scenario is false, which evidence would we then expect to find (or not expect to find)?' Obviously we can only answer this question when we have already formulated a scenario. It does not make sense for an accommodator, who gathers evidence without having a scenario, to ask what could prove this scenario wrong. Nonetheless, it is possible that this question can be answered by referring to the available evidence, which was used in the construction of the scenario. If the accommodator has done a thorough and fair search for evidence before constructing their scenario it might be that she has already searched for—but failed to find—any contradicting evidence.[8] However, if this is not the case, the falsificationist mindset entails that we have to check on the scenario's risky predictions—which are those predictions that are likely to fail if the scenario is untrue. This role of predictions was famously emphasised by Popper (1959), who considered falsification attempts of theories crucial to the scientific enterprise. Within the predictivism debate Mayo (1991) argues that our intuitions about the value of predictions are best explained by the fact that testing a hypotheses' risky predictions put it through a 'severe test'—i.e. tests that a hypothesis is unlikely to pass if it is false.

This brings us to the final argument for the value of predictions over accommodations, which Lipton (2005) calls *the argument from choice*. When we accommodate, we have to make the best of whatever information is available. However, this will not always include the most telling evidence—i.e. the evidence with the most discriminating likelihood ratio. In contrast, when we predict, we can pick which predictions to test.[9] This means that we can choose to check the predictions that test our scenario the

---

8. Though investigators should be careful even when this appears to be the case. Research suggests that people often prematurely stop looking for further evidence once they have formulated a scenario that explains the known facts well enough (Hoch, 1984).

9. One caveat to this argument is that criminal investigators may not always have the opportunity to carefully test a scenario's predictions (cf. van Koppen, 2011: ch. 3). After all, criminal cases usually involve limited and deteriorating evidence.

most severely. If a scenario passes these tests it will thereby be more strongly supported than if it did not. This strong support does not arise because predicted evidence is inherently better; the resulting evidence would be equally strong if it were used to construct an explanation. Rather, when we test predictions, we have more options for choosing what evidence to gather and how to gather this evidence. For example, suppose that there is a murder case. Once we have a detailed scenario of when, where and how the murder took place, we can start to look for witnesses that were in that area at that time and find out whether their testimony matches this scenario.

Of course, as mentioned above, the danger with testing predictions is that investigators may be susceptible to fudging the results. When they test a scenario's predictions, they might subconsciously choose to look for only confirming evidence or to misinterpret the results to fit with their preconceived scenario. For instance, they might be inclined to ask witnesses leading questions. Nonetheless, when investigators adopt a falsificationist mindset, where they are cognisant of the possibility of confirmation bias and deliberately seek out evidence that might contradict their scenario, they diminish this worry. So checking on predictions does not inherently prevent confirmation bias. Rather, carefully testing predictions is a vital part of adopting a falsificationist attitude, which in turn counteracts confirmation bias. When our scenario only explains known facts, we run the risk of not testing our scenario as severely as possible.

## Case study: The murder of Hae Min Lee

So far, I have illustrated the special value of predictions in criminal cases with simple examples. Let's now consider the above arguments in the context of a real-life case—the murder of Hae Min Lee. The podcast *Serial* (2014) gained fame by exploring the intricacies of this case in its first season. Given the scope and complexity of this case, I invite the reader to listen to this podcast for a more extensive discussion of the investigation and subsequent trial. The goal of this case study is not to reach a verdict on whether the conviction of Adnan Syed—the suspect in this case—was legitimate. Rather, I discuss this case because it provides vivid, practical examples of predicting and non-predicting witnesses, severe testing by means of checking predictions, and fudged scenarios and evidence. Hence, it shows that the aforementioned arguments have practical relevance for real-life cases.

### The case

Hae Min Lee, an American high school student, was murdered in early 1999. Her body was found in a park four weeks after she was last seen alive. The cause of death was manual strangulation. Adnan Syed, Lee's ex-boyfriend, was arrested, charged and convicted for her murder. Though the decision was upheld in appeal, critics argued that it was unclear whether the evidence truly supported the prosecution's scenario. The conviction was primarily based on two items of evidence. First, the testimony of Jay Wilds, who claimed that Adnan had murdered Hae. Second, call records of Adnan's mobile phone.

*Jay's story.* Jay was an acquaintance that Adnan occasionally smoked weed with. He testified multiple times, both in police interviews and in court, that Adnan killed Hae. His story was as follows: They met up to go shopping in the morning. Adnan then told him that he was going to 'kill that bitch', referring to Hae who had recently started seeing someone else. Afterwards Jay dropped him off at school and kept both Adnan's car and phone, so that he could pick him up later. In the afternoon, Adnan called Jay from a payphone to come pick him up. When he arrived, Adnan showed him Hae's body in the trunk of her own car. After the incident, they left Hae's car elsewhere. Jay then dropped off Adnan at track practice. Later, Adnan called Jay to come pick him up again and after eating together they buried Hae's body in a nearby park.

Critics questioned Jay's credibility, especially because he kept changing his story between testimonies. For instance, he gave differing reports about where he first saw Hae's dead body. At one point he admitted having lied to the police on previous occasions. However, according to the prosecution, Jay's story was ultimately credible because it was corroborated by the call records.

*Adnan's story.* According to Adnan, he did indeed lend his car and phone to Jay to be picked up after track practice. During the time period when Hae was most likely killed he claims that he was first in the library and then went to track practice. After track practice, Jay did pick him up. But according to Adnan, they went to a friend's house together. In the evening, he went to the mosque alone.

*The cell phone records.* The call record was a list of over thirty calls that contained the phone numbers that called or were called by Adnan's cell phone, the time at which these calls took place, the duration of the calls and—most importantly—which cell phone tower the call was routed through.

According to both Jay and Adnan, it was Jay who had the phone on him that fateful day. The prosecution argued that the cell phone records corroborated Jay's testimony. First, they gave a rough location of where Adnan's phone was—which prosecutors claimed matched up with the locations in Jay's story. They also connected Adnan to the location of the phone because of one particular call. This was the 3:32 p.m. call made to Nisha, a friend of Adnan. She was the only person on the call list who did not know Jay. Hence, the reasoning was, Adnan must have been near the phone that afternoon, because Jay wouldn't have called her since he didn't know Nisha. This in turn would mean that Adnan was lying when he said that he and Jay were not together that afternoon until Jay picked him up from track practice.

## Predicting versus non-predicting witnesses

As I discussed earlier, predicting witnesses tend to be more reliable than those who accommodate. This case featured the testimony of two key witnesses: Jay and Adnan. Interestingly, the testimony of Jay was confirmed by a successful prediction, while that of Adnan was not. Let's consider what that means for the degree of reliability that we should ascribe to them.

Much of Adnan's scenario was unverifiable—it was low on details and did not lead to any clear, novel predictions that the police could follow up on. In contrast, Jay did make a successful prediction. His story included him and Adnan getting rid of this vehicle together and he successfully predicted the location of Hae's car. When Jay was first interviewed by the police, he correctly pointed the police to a hill near the city where it was parked. The police discovering the car there is a novel prediction. Imagine that the police had already found Hae's vehicle and had told Jay about this. If Jay's story then included its location, we might be worried that he was (deliberately or subconsciously) fitting his story to the information that the police gave him—i.e. that he was accommodating. Based on this successful prediction, the police considered his story believable enough to arrest Adnan.

What does this mean for their credibility? The answer to this question depends on the likelihood ratio of the testimony. For instance, what is the probability of Jay successfully predicting the car's location, given that he is telling the truth, versus its probability given that he is not telling the truth? If Jay is telling the truth, then he is very likely to get the car's location right. He might get it wrong if he misremembers, or if the car had, say, been towed. But on the balance, his successful prediction of the car's location is unsurprising if we assume that he is telling the truth. Now imagine that Jay is not telling the truth, how likely is his successful prediction then? It certainly makes Jay's story more credible than if he had made no novel predictions—for instance if he had included the car's location in his scenario *after* the police had told him about it. After all, this proves that he knew *something*. He was not merely fitting his testimony to publicly known evidence. The prediction therefore takes away one worry that we might have about his story. Nonetheless, there are other scenarios, in which Adnan is innocent, where Jay would also have knowledge of the car's whereabouts. Such scenarios include

Jay killing Hae himself or someone else than Adnan killing Hae and telling Jay about it. How much this prediction tells us depends on how plausible these alternative explanations for his knowledge of the car's location are.

Now let's consider Adnan, whose testimony did not lead to clear, novel predictions. Is this lack of prediction itself evidence that he is not telling the truth? Again, this depends on the likelihood ratio. How probable is this if we assume that he is not telling the truth? I would suggest that this is quite likely. The reason for this is that if Adnan is lying, then it is in his best interest to assure that investigators do not discover evidence that contradicts his story. However, his lack of predictions does not—by itself—mean that we should assign a low degree of credibility to his testimony. After all, suppose that he is telling the truth. The likelihood of him making no predictions is then, arguably, also high. After all, it will often be difficult to produce exonerating, novel evidence for where you were on any particular day, even if you did not do anything illegal that day. This is especially the case when you are first interviewed about that day over a month after the fact, as Adnan was.

Nonetheless, now suppose that Adnan *had* made a novel prediction. In that case the likelihood of his testimony would be much lower under the assumption that he was lying. After all, it would thereby have taken away the worry that he is lying by fitting his testimony to the facts provided by the police. By not predicting any novel evidence, Adnan fails to take away this worry. In other words, though his lack of predictions is not necessarily a sign of unreliability, if he *had* predicted, it would have been a sign of reliability.

## Route talk and severe tests

Let's turn to the argument from choice (I'll get to the notion of fudging in the next subsection). According to this argument predictions can sometimes be designed to provide as severe tests as possible, in ways that accommodations cannot. The reasoning is that when we predict, we can first identify what the most telling evidence would be and we can then deliberately search for that evidence. This option may not be available when we accommodate, where we have to make do with whatever evidence is available. Let's consider an example to illustrate this argument.

In episode 5 of the podcast, the presenters took up a challenge posed to them by Adnan to test a prediction following from the prosecution's timeline.[10] The prediction in question is about the time it would take to drive from Hae and Adnan's school to the parking lot where Hae was allegedly murdered. Hae's last class ended at 2:15 p.m. that day and multiple people remembered seeing her heading to her car afterwards. Furthermore, Adnan supposedly called Jay at 2:36 p.m. to come pick him up after he had killed her. According to the prosecution, in those 21 minutes, Hae drove to Best Buy and Adnan strangled her there.

The podcast makers wanted to test this prediction—that it was possible to make this drive in 21 minutes, and still leave time to strangle someone. They made several attempts, starting at the high school, right after classes ended. They concluded that it was indeed possible, as long as the driver made no errors and there were no delays because of traffic or other sources. Yet Adnan claimed that such conditions were unlikely, because leaving the school after classes ended meant that 1,500 students were exiting the building, and driving off the parking lot usually meant having to wait for buses. Furthermore, even without delays, Adnan would have had only three minutes to strangle Hae, put her body in the trunk of his car (in broad daylight) and call Jay from a public pay phone. Hence, the podcast deemed it unlikely that the state's timeline was correct.

The relevance of this example is that it illustrates the key role of predictions in severely testing a scenario. Such severe testing involves asking: 'Which facts could prove this scenario wrong?' When we

---

10. Note that the prediction as I discuss it here is therefore made by the prosecution scenario, not by Adnan. Adnan merely points out that this prediction is implausible.

predict we can choose to focus on the most promising predictions that could realistically falsify the scenario. For instance, in this case the prediction was that it should be possible to drive from the school to the Best Buy in 21 minutes, given similar traffic conditions to those Hae would have faced that day, while still leaving time for Adnan to strangle her. This is a very specific prediction, which the podcast makers checked on because its importance had been pointed out by Adnan. He suggested that this was a way to show that the scenario was wrong. To put it differently, this predicted fact arguably had a high likelihood ratio—if the prosecution's scenario was correct, it was very likely to be proven true. But, given Adnan's remarks, if the scenario was false then this would be one of the facts which would most likely show it.

To see why only accommodating may often leave a scenario non-severely tested, imagine that the podcast had only checked on information that was already known and that had already been used to construct the scenario. This would mean restricting themselves to re-checking the information that was used to arrive at the very conclusion that Adnan was the killer. Such information *could* turn out to contradict the scenario. I will discuss examples of this in the next subsection. However, if there was any information that could prove his innocence, it was likely to be found outside the set of evidence that first led investigators to suspecting Adnan.

## Fudging and the call record

I have discussed two kinds of fudging—fudged scenarios and fudged evidence. A scenario is fudged when it is designed to fit well with the evidence at the cost of its inherent plausibility. Evidence is fudged when it is selectively chosen or interpreted to fit with a preconception, such as a preconceived scenario. The Hae Min Lee case arguably featured instances of both a fudged scenario and fudged evidence, namely with respect to the fit between Jay's story and the call record. A look at this example will help to get a better grip on what fudging looks like in practice.

To begin with the apparent fit between the call records and Jay's scenario: Adnan's mobile phone pinged specific cell towers during the presumed day of the murder, whenever it was used to make calls. Because cellphones normally ping the closest tower, this meant that rough estimates could be made about the phone's location during specific times of the day. The prosecution claimed that Jay's story fit the call record perfectly. Furthermore, they argued that Jay could not possibly have known which towers were getting pinged when he told his story. In other words, they claimed that he had made a number of successful, and risky predictions, which offered strong evidence for his story. Furthermore, Adnan's story did not seem to fit with these records. So, if the prosecution was correct in their assertion that the records and Jay's story fit perfectly, then it would be very probable that Adnan was guilty.

However, there were several reasons to doubt that the fit between the call record and Jay's testimony was as close as the prosecution claimed. Take the 21 minutes between Hae leaving class and Adnan allegedly calling Jay to come pick him up for example. What I did not mention earlier is that neither Adnan nor Jay claim that the 'come pick me up' call took place at the time that the prosecution's time line said it did—2:36 p.m. According to Jay the call happened around 3:40 p.m., much later. But this was a problem for the prosecution, because the call record shows no call from the Best Buy location around that time. So they concluded that the 2:36 call—the only call that feasibly matched both the time and location—was the 'come pick me up' call (Serial, 2014: 199). However, this time line led to the implausible 21-minute window for Hae's murder. In other words, the prosecution tried to fit their time line to the evidence—Hae being spotted leaving class, Jay's testimony about Adnan calling him to be picked up, and the call records. But this also made the resulting time line implausible, as it meant that the alleged events—Hae driving to Best Buy, Adnan strangling Hae, Adnan calling Jay—would have to have taken place in an unrealistically brief amount of time. The prosecution could have opted for a different, more realistic time line. However, then their scenario would have conflicted with either Jay's testimony, sightings of Hae leaving class or the call record. For instance, if they had taken Jay's claim that the 'come pick me up' call took place at 3:40 p.m., then there would be a clear disconnect with the

call record—which showed no such call from Best Buy at 3:40 p.m. Instead the prosecution ended up with an implausible time line, though one that—at the very least appeared to—fit well with the evidence. In other words, they *fudged their scenario,* i.e. sacrificed the plausibility of their scenario to get it to fit with the evidence.

Susan Simpson (2015), a legal associate, pointed out another possible instance of fudging, in her online article 'Evidence that Jay's Story Was Coached to Fit the Cellphone Records'. She focused on one of the changes in Jay's story between police interviews. At first, Jay claimed that he was at home when Adnan called him to be picked up from track practice. This matched the location data on the phone records. However, he later changed his story and said that he was at a friend's house. The friend denied this and it also did not match the location data. According to Simpson, this change is best explained by the fact that the police were, at first, working with an inaccurate map of the phone location data. On this map, the tower that Jay's phone pinged was displayed in a different location than its actual location and Jay's new story matches that location. So, Simpson argued, the changes in Jay's story are likely due to the police coaching Jay to make his story fit with the cell records.

These are both instances of the *scenario* being changed to fit the phone records. However, there were also signs of the prosecutor fudging the *evidence*. For example, during the trial the prosecutor only cited four of the fourteen locations that the phone pinged, because only those four matched Jay's accounts. The pings that conflicted with his account were swept aside. Furthermore, prosecutors and investigators may have misinterpreted a crucial piece of evidence, in order to make it fit with their preferred scenario. This piece of evidence was the Nisha call—which tied Adnan to the location of the phone, because it was the only call to a person that only Adnan knew. This call is often treated as the smoking gun, which definitively disproves Adnan's story and proves Jay's. However, there are reasons for skepticism. Most importantly, in her description of the call, Nisha says that Adnan and Jay wanted her to come to a video store where Jay worked. But he did not start that job until two weeks after this call. Furthermore, while the phone tower matches the time of Jay's story, it does not match the location. In fact, none of the calls from around that time match where Jay says that they were at the time. Many have since pointed out that the prosecution did not disprove the possibility that Jay accidentally 'pocket dialled' Nisha's number. The podcast looked into this option extensively and concluded that it was indeed possible (Serial, 2014: 275). According to this theory, Nisha could be confusing this call with a later call later in the month, when Jay already worked at the video store. So this call may have been (subconsciously) misinterpreted to make it appear as if it strongly supports the prosecution scenario. If that was the case, there would be no call tying Adnan to the phone's location.

To summarise, there are several reasons to believe that the apparent match between Jay's testimony and the call records was due to fudging. Yet this fudging was not at all apparent at first. For example, much of it escaped the attention of the defence during the trial. Presumably the prosecution also missed it (assuming that they were not deliberately fudging). Furthermore, even after a careful examination by the podcast and others, such as Simpson, we cannot say with certainty that the prosecutors and investigators did indeed fudge—there is simply too much unclarity on this issue.

Let me make one final point about fudging that is potentially interesting. I mentioned earlier that predictions can play a key role in preventing confirmation bias. In particular, checking on predictions is often a vital part of adopting a falsificationist mindset. However, as I also discussed, the phenomenon of confirmation bias can also give rise to an *illusion* of predictive success. Consider the Nisha call. Jay mentions the call during a police interview, seemingly without knowing that it shows up on the call record. Nisha independently confirmed that she talked to Adnan who briefly put Jay on the phone. This therefore appears to be an instance of a successful, risky prediction by Jay. However, this appearance might be caused by the aforementioned fudging—where Nisha's call is re-interpreted to fit with Jay's story. So Jay's apparent predictive success may be due to investigators fudging the evidence. Again, such confirmation bias can be counteracted by carefully considering what evidence could disconfirm the scenario. For instance, the investigators could have considered whether there was any evidence that contradicted their interpretation of the Nisha call.

# Conclusion

Predictions are a vital element of science. It is therefore unsurprising that philosophers of science have thought about their value deeply. Though predictions are less central to criminal cases, I have argued that predicted evidence also has special value in that context. In particular, I distinguished three ways in which predicted facts can provide stronger criminal evidence than accommodations. First, witnesses who predict are—all else being equal—more reliable than those who do not. Second, investigators who only accommodate run the risk of 'fudging' their scenario—predictions can protect us against this danger. Third, carefully constructed predictions can be designed to provide severe tests of our criminal scenarios. This in turn is a useful tool in preventing confirmation bias. I showed how these arguments can feature in real, complex situations by discussing the murder case of Hae Min Lee.

One conclusion that follows from this article is therefore that a careful look the predictivist arguments developed by philosophers of science can offer valuable lessons for criminal evidence scholars. However, the reverse may also be true—the above discussion might shine some light on the scientific predictivism debate. First, I opted for a pluralist approach where the value of prediction over accommodation does not reduce to a single value, but to multiple. There are also some philosophers of science who also adopt a pluralist view (Barnes, 2005; Douglas and Magnus, 2013). My proposal further shows the value of such an approach—because it allows for prediction to matter in different ways in different situations. In particular, the three values of prediction that I distinguished are categorically different: For the evaluation of eyewitness evidence, predictions are—all other things being equal—inherently better evidence than accommodations. When it comes to fudging, the value of prediction is not inherent. A perfectly rational being should assign the same degree of confirmation to a scenario that was supported with predictions as they should to one that was supported with accommodated evidence (all other things being equal). Yet given the fact that those involved in criminal cases are not perfect rational beings, a piece of evidence that was predicted should give a confirmatory boost for the relevant scenario. Finally, the argument from choice does not establish such a confirmatory boost. However, it does suggest that predicted evidence is often better evidence than accommodated evidence. But this is only because of the practical possibilities of carefully designing such predictions. This distinction between an inherent epistemic value, a weaker epistemic value and a practical value of prediction can shine new light on the relation between different predictivist arguments in the philosophy of science.

A further point that I believe is worth considering, is the potential use of case-studies for the predictivism debate. I offered one such case study in this article. In general, I believe that we should not be content with thinking about predictivism using only simplified thought experiments, as many authors do. A look at actual cases—whether they are criminal or scientific—are vital to examine the nuts and bolts of predictivism. In particular, only considering simplified thought experiments has a tendency to lead to simplified philosophical accounts. This is arguably an issue in the predictivism debate. For instance, some claim that famous historical cases of telling predictions do not support the broad and strong predictivist claims that various philosophers of science derive from them (e.g. Scerri and Worrall, 2001).

Let me end by rephrasing my conclusions. First, witnesses sometimes fit their story to the wrong facts. That is why we want to make sure that they predict, rather than accommodate these facts. Second, investigators sometimes fit their story to the facts wrongly. Prediction yields less incentive for this kind of overfitting. Third, investigators sometimes fit their facts to the story wrongly. That is why it is useful for them to adopt a falsificationist attitude. This usually means testing a scenario's risky predictions. For these three reasons, those involved in criminal cases should care about whether evidence was predicted or accommodated.

## Acknowledgements

## Declaration of Conflicting Interests

## Funding

## ORCID iD

Hylke Jellema ![ORCID] https://orcid.org/0000-0003-1155-5581

## References

Barnes EC (2005). Predictivism for pluralists. *The British Journal for the Philosophy of Science* 56(3): 421–50.

Barnes EC (2008). *The Paradox of Predictivism*. Cambridge: Cambridge University Press.

Barnes EC (2018). Prediction versus accommodation. In Zalta EN (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2018). Available at: https://plato.stanford.edu/archives/fall2018/entries/prediction-accommodation/ (accessed 4 March 2021).

Douglas H and Magnus PD (2013). State of the field: Why novel prediction matters. *Studies in History and Philosophy of Science Part A* 44(4): 580–589.

Dror IE, Charlton D and Péron AE (2006). Contextual information renders experts vulnerable to making erroneous identifications. *Forensic Science International* 156(1): 74–78.

Findley KA and Scott MS (2006). The multiple dimensions of tunnel vision in criminal cases. *WIS. L. REV* 291(2): 312–313.

Gross SR, Jacoby K, Matheson DJ and Montgomery N (2004). Exonerations in the United States 1989 through 2003. *Journal of Criminal Law & Criminology* 95(2): 523.

Gunn LJ, Chapeau-Blondeau F, McDonnell MD, Davis BR, Allison A and Abbott D (2016). Too good to be true: When overwhelming evidence fails to convince. *Proc. R. Soc. A*, 472: 20150748.

Harker D (2008). On the predilections for predictions. *The British Journal for the Philosophy of Science* 59(3): 429–453.

Hitchcock C and Sober E (2004). Prediction versus accommodation and the risk of overfitting. *The British Journal for the Philosophy of Science* 55(1): 1–34.

Hoch SJ (1984). Availability and interference in predictive judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10(4): 649–662.

Jellema H (2019). Case comment: responding to the implausible, incredible and highly improbable stories defendants tell: a Bayesian interpretation of the Venray murder ruling. *Law, Probability and Risk* 18(2–3): 201–211.

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

Kahn JA, Landsburg SE and Stockman AC (1996). The positive economics of methodology. *Journal of Economic Theory* 68(1): 64–76.

Kassin SM, Dror IE and Kukucka J (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition* 2(1): 42–52.

Kuhn TS (1957). *The Copernican Revolution: Planetary Astronomy in the Development of Western Thought*. Cambridge, MA: Harvard University Press.

Lange M (2001). The apparent superiority of prediction to accommodation as a side effect: A reply to Maher. *The British Journal for the Philosophy of Science* 52(3): 575–588.

Lipton P (2005). Testing Hypotheses: Prediction and Prejudice. *Science* 307(5707): 219–221.

Mackor AR (2017). Novel facts: The relevance of predictions in criminal law. *Strafblad* 15(2): 145–156.

Maher P (1988). Prediction, accommodation, and the logic of discovery. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1988: 273–285. Philosophy of Science Association.

Mayo DG (1991). Novel evidence and severe tests. *Philosophy of Science* 58(4): 523–552.

Nickerson RS (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology* 2(2): 175–220.

O'Brien B (2009). Prime suspect: An examination of factors that aggravate and counteract confirmation bias in criminal investigations. *Psychology, Public Policy, and Law* 15(4): 315–334.

Pardo MS and Allen RJ (2008). Juridical proof and the best explanation. *Law and Philosophy* 27(3): 223–268.

Pennington N and Hastie R (1991). A cognitive theory of juror decision making: The story model. *Cardozo Law Review* 13(2–3): 519–557.

Popper K (1959). *The Logic of Scientific Discovery*. London: Hutchinson.

Scerri ER and Worrall J (2001) Prediction and the periodic table. *Studies in the History and Philosophy of Science* 32(3): 407–452.

Serial (2014). Complete transcripts for serial podcast season one. Available at: www.adnansyedwiki.com/wp-content/uploads/2018/06/Serial-Podcast-Transcripts-of-All-Episodes-with-ToC.pdf (accessed 4 March 2021).

Simpson S (2015). Serial: Evidence that Jay's story was coached to fit the cellphone records. In: ViewfromLL2.com. Available at: http://viewfromll2.com/2015/01/13/serial-evidence-that-jays-story-was-coached-to-fit-the-cellphone-records/ (accessed 4 March 2021).

Thompson WC (2017). Determining the proper evidentiary basis for an expert opinion: What do experts need to know and when do they know too much? In: Robertson C and Kesselheim A (eds) *Blinding as a Solution to Bias: Strengthening Biomedical Science, Forensic Science, and Law*. London: Academic Press, pp.133–150.

Tuzet G (2019). Abduction, IBE and standards of proof. *The International Journal of Evidence & Proof* 23(1–2): 114–120.

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

Vredeveldt A, van Koppen PJ and Granhag PA (2014). The inconsistent suspect: A systematic review of different types of consistency in truth tellers and liars. In: Bull R (ed) *Investigative Interviewing*. New York: Springer pp.183–207.

Wise RA, Sartori G, Magnussen S and Safer MA (2014). An examination of the causes and solutions to eyewitness error. *Frontiers in Psychiatry* 5: 102.

Zaragoza MS, Belli RF and Payment KE (2006). Misinformation effects and the suggestibility of eyewitness memory. In: Garry M and Hayne H (eds) *Do Justice and Let the Sky Fall: Elizabeth F. Loftus and Her Contributions to Science, Law, and Academic Freedom*. Mahwah, NJ: Lawrence Erlbaum, pp.35–63.