BAYESIAN MODELLING OF CRIMINAL CASES AS A WHOLE
A PHILOSOPHICAL REFLECTION ON DUTCH CASE LAW

Anne Ruth Mackor[*]

ABSTRACT: The Bayesian model is gaining ground as a rational method for evaluating evidence in criminal law. The central question of this article is whether the Bayesian model can provide guidance for the judicial evaluation of evidence. More specifically, it focuses on the use of Bayesian modelling of complete criminal cases. Section 2 explicates the model, discussing an example of how experts from the Netherlands Forensic Institute (NFI) apply it in their reports. Section 3 addresses the question of whether non-forensic experts and judges can or even should use the Bayesian model as well. The focus of the rest of this article is on Bayesian modelling of criminal cases as a whole. In section 4, I discuss Dutch case law on the use of such integral modelling. The Netherlands appears to be internationally unique in that this question has been addressed in case law. Since the Bayesian model is worldwide gaining ground in forensic sciences and in theorizing about criminal evidence law, a discussion of Dutch case law can be relevant for legal scholars and legal philosophers from other, in particular from comparable continental legal systems. Section 5 offers a further legal-philosophical and epistemological reflection on the nature and functions of Bayesian modelling of criminal cases as a whole. The contribution concludes with recommendations for further research. Part of this further research is being carried out within the research project Preventing Miscarriages of Justice.

KEYWORDS: criminal law; case law; bayesian model; bayesian networks

---

SUMMARY: 1. INTRODUCTION: EVIDENCE IN CRIMINAL LAW.— 2. THE BAYESIAN
MODEL.— 3. THE SCOPE OF THE BAYESIAN MODEL IN CRIMINAL CASES.—
4. DUTCH CASE LAW ON BAYESIAN MODELLING OF CRIMINAL CASES AS A
WHOLE: 4.1. The case of 'The six of Breda'; 4.2. Four later court rulings; 4.3. The silence of
the Dutch Supreme Court.— 5. FURTHER CONSIDERATIONS REGARDING BAYESIAN
MODELLING OF CRIMINAL CASES AS A WHOLE.— 6. CONCLUSION AND RECOM-
MENDATIONS.— REFERENCES

1.   INTRODUCTION: EVIDENCE IN CRIMINAL LAW

Article 338 of the Dutch Code of Criminal Procedure requires that the judge
be *convinced* by the content of the lawful evidence presented during the trial that
the defendant committed the offence with which he or she is charged, before the
judge may proceed to convict the defendant. Because of the term lawful evidence,
Dutch criminal law is referred to as a *legal* system of proof. Because of the fact that
a judge is not obliged to convict if there is sufficient lawful evidence, it is referred to
as a *negative* system of proof. Accordingly it is a negative legal system of proof. How-
ever, as the law of evidence does not give further guidance to the judge's conviction,
it is usually considered a *free* system of proof. As a consequence, Dutch criminal law
does not so much leave the judge free, but rather helpless. It is unclear how the judge
can or should rationally evaluate evidence, i.e. how and to what extent the content
of the various lawful means of evidence can or should contribute to his conviction.

The Bayesian model is gaining ground in forensic sciences and in theorizing
about criminal evidence law as a rational method of evaluating evidence. This article
focuses on the question of whether the Bayesian model can provide guidance for the
judicial evaluation of evidence. More specifically, it focuses on the use of Bayesian
modelling of complete criminal cases. Section 2 explains the model, discussing an
example of how experts from the Netherlands Forensic Institute (NFI) apply it in
their reports. Section 3 addresses the question of whether non-forensic experts and
judges can or even should use the Bayesian model. We will distinguish between the
following questions. Can forensic and/or non-forensic experts use the Bayesian to
determine the probative force of individual evidence? Can the Bayesian model be
used for an integral modelling of a complete case as a whole? If so, should this
be done by an expert or by the court itself?

The focus of the rest of this article is on the latter question, i.e. Bayesian mo-
delling of criminal cases as a whole. In section 4, I discuss Dutch case law on the use
of such integral Bayesian modelling. The Netherlands appears to be ahead of other
countries, insofar as this question has been addressed in Dutch case law. Because the
Bayesian model is gaining ground as a rational method for evidence evaluation, a
discussion of Dutch case law can also be relevant for legal scholars and legal philoso-
phers from other, in particular from comparable continental, legal systems. Section 5
offers a legal-philosophical reflection on the nature and functions of Bayesian mod-

elling of criminal cases as a whole. The article concludes with recommendations for further research. Part of this further research is being carried out within the research project Preventing Miscarriages of Justice.[1]

2.   THE BAYESIAN MODEL

Since the 1970s, the Bayesian model has been gaining ground as a rational method of evaluating evidence in forensic sciences and in theorizing about criminal evidence law.[2] For the past 20 years, it has become widely accepted among European forensic scientists as a method for rationally evaluating evidence. Let me start with a simple example. It is known from statistics that it rains more than half of the days of the year in the Netherlands.[3] However, after looking outside and seeing bright sunshine ('evidence' 1) and hearing the weather forecast predicting "dry everywhere" ('evidence' 2), I will lower my estimate of the probability of rain for today. This is something we already do spontaneously and intuitively; the Bayesian model provides a precise method for this.

Put simply, the model prescribes that we first estimate the prior odds, that is, prior to the evidence, that it will or will not rain. In the example we can do this on the basis of official statistics.[4] When these are not available, we will base our estimate on our own experience or informally on 'facts of general knowledge'. The next step is to estimate which hypothesis (in the example: rain or no rain) makes the evidence (bright sunshine plus weather forecast) more likely. This second estimate is called the likelihood ratio (LR). The greater the LR, the stronger the evidence. Other words for likelihood ratio are therefore "evidential strength", "probative strenght", "evidential force" and "probative force". The Bayesian model offers a method to derive the posterior odds that it will or will not rain today from the prior odds (in this example: based on the statistics on the number of days of rain per year) and the LR of the evidence (in this example: bright sunshine plus weather forecast). In a 'formula': prior odds x likelihood ratio = posterior odds.[5]

---

[1]   See note *.

[2]   Note that the Bayesian model is not the only method for rational evaluation of evidence. H. Prakken, F.J. Bex & A.R. Mackor Editors' Review and Introduction: Models of Rational Proof in Criminal Law, *Topics in Cognitive Science* 12 (2020) 1053–1067 offer an overview of three main approaches, viz. argumentation-based approaches; explanation-based, more in particular story-based approaches; and probabilistic, in particular Bayesian, approaches.

[3]   *https://www.knmi.nl/over-het-knmi/nieuws/het-regent-niet-vaker-wel-harder*

[4]   An important question is which statistics this assessment should be based on. In the example: on the statistics for the entire year, for this month, or for this specific day? I will leave this problem, which is known as the reference class problem, aside in this paper.

[5]   NFI *Vakbijlage Waarschijnlijkheidstermen* [*Technical Supplement Probability Terms*], 2017. *https://www.forensischinstituut.nl/publicaties/publicaties/2017/10/18/vakbijlage-waarschijnlijkheidstermen*, p 3.

In a similar way the Bayesian model can be applied in criminal cases. The focus is usually on the question of the posterior probability that the defendant committed the alleged offence in light of the prior probability and the probative force of all the evidence (the LR's).[6] Since around 2005, the Netherlands Forensic Institute (NFI) writes its reports in accordance with the Bayesian model.[7] In 2017, the NFI published a technical supplement explaining the model and its application in NFI reports.[8] NFI researchers report on the LR, that is, on the probability of finding the research results in the light of at least two hypotheses. Consider the following example.[9]

> A researcher compares a shoe print with the suspect's shoe. He formulates the following hypotheses.
> H1 The suspect's shoe caused the print.
> H2 Another shoe with a similar profile and size caused the print.
> After forensic investigation, the conclusion is:
> The results of the investigation are much more likely if the suspect's shoe caused the mark than if another shoe with a similar profile and size caused the mark.

The report gives the LR of the shoe mark. In this example, the LR is "the results of the investigation are much more likely if H1 is true than if H2 is true". "Much more likely" means that the results are 100 to 10,000 times more likely.[10] In other words, the probability of finding the shoe mark with this profile and size is 100 to 10,000 times greater if the mark was made by the suspect's shoe than if it was made by any other shoe. In this example, the LR is qualitative, i.e. expressed in words. A quantitative LR, expressed as a specific number, is possible if relevant data is available from representative samples or experiments.

Next, it is up to the court to combine the LR from the NFI report with all other information. This is a complex task. First, the court must estimate the prior probabilities of both hypotheses. In other words, the court must estimate the probability, prior to the evidence, that the suspect's shoe and not another shoe made the mark. Sometimes that estimation can be based on statistics. We saw that this was possible when estimating the prior probability of rain. However, in criminal cases reliable statistics are often lacking, and then the estimation is necessarily subjective. That is also why experts do not offer estimates of the prior probability and leave that task to

---

[6] Another important probability assessment in criminal law is that of the risk of recidivism.

[7] H. Prakken & R. Meester, 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* 2017-5, 185-197.

[8] NFI 2017.

[9] NFI 2017, p. 2.

[10] NFI 2017, p. 6, table 2. When it comes to hypotheses about the source of DNA traces, the results are often "much more likely" (10,000 to 1,000,000 times) or even 'extremely much more likely' (more than 1,000,000 times) under the hypothesis that the trace comes from the suspect than under the hypothesis that it comes from any other person. However, much evidence has a much smaller LR. For example, 'slightly more likely' means that finding the evidence is 2 to 10 times more likely, and 'more likely' means 10 to 100 times more likely.

the court. After all, in such cases they do not have the specific knowledge or expertise on which to base such an estimation.[11] Secondly, the court must estimate the probative force of all other evidence. The NFI can provide a LR for some of the evidence, but there is also a lot of evidence for which scientific knowledge is lacking and on which experts are therefore unable to report. For that evidence, the court will have to estimate the probative force by itself.[12] Finally, the court must combine the NFI's LR with the probative force of all other evidence, such as witness statements, DNA, glass, fibre or gunpowder traces and the relevant priors.

It is for the latter task, the combination of priors and LR's, that the Bayesian model offers a precise mathematical method. The method can also help to prevent reasoning errors, such as the notorious prosecutor's fallacy.[13] Again, take the example of the shoe print. The conclusion of the above-mentioned NFI-report was: "The results of the investigation are much more likely if the suspect's shoe caused the mark than if another shoe with a similar profile and size caused the mark."

However, many people, including judges, read the conclusion as: "The probability that the shoe print came from the suspect's shoe (the hypothesis), given the match (the evidence), is much greater than the probability that the print was made by another shoe (the alternative hypothesis)." This interpretation is incorrect, however. The mistake is that hypothesis and evidence are reversed. In general terms: the NFI reports on the LR, which is the conditional probability of finding the evidence if the hypothesis is true. However, many people, including judges, read it the other way around as the posterior, the conditional probability that the hypothesis is true, given the evidence. A classic example can illustrate that this can be a major mistake: the

---

[11]   The Amsterdam Court of Appeal of 12 April 2021 (ECLI:NL:GHAMS:2021:986), quotes a forensic expert who explains that: "… the NFI deliberately refrains from making a priori assessments, because the NFI cannot determine those probabilities. The NFI never has access to all the information in a case, because a case always contains information that falls outside the NFI's area of expertise. The court must determine what the a priori probability is, given the established facts before and during the incident. The NFI's statements only relate to the discovery of … findings as to whether a hypothesis is correct, compared to an alternative hypothesis. In view of the above, the Bayesian method of reasoning is pure, because this method allows experts to remain within their own area of expertise."

[12]   For example, a judge might ask how much more likely it is that the defendant filled up with a small amount of petrol shortly before a certain fire took place (evidence) if the defendant is the perpetrator (H1) than if he is not the perpetrator (H2). If no statistics are available, judges would have to rely on their own experience or on so-called 'facts of general knowledge'. This example is taken from the conclusion of Advocate General P.M. Frielink of 25 May 2025 (ECLI:NL:PHR:2025:378).

[13]   The term is taken from William C. Thompson and Edward L. Shumann, "Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor's Fallacy and the Defense Attorney's Fallacy." *Law and Human Behavior* 2 (3): 167–87. Also see C. Dahlman, A systematic account of probabilistic fallacies in legal fact-finding, *The International Journal of Evidence & Proof* 2024/1, pp. 45-64 and Hylke Jellema and Christian Dahlman, Probabilistische drogredenen in de rechtszaal: een systematisch overzicht [Probabilistic fallacies in the courtroom: a systematic overview], *Expertise en Recht* 2025(3), 64-74 for an overview and explanation of the most important fallacies.

conditional probability that an animal has four legs if it is a cow is very high, whereas the conditional probability that an animal is a cow if it has four legs is very low.

Because of the complexity of the details of the Bayesian model and the scope of this article, the question of exactly how the Bayesian model should be applied in a mathematical manner cannot be addressed. However, the above explanation should be sufficient, also for legal scholars without knowledge of the Bayesian model, to understand the discussion in the rest of this article. [14]

3. THE SCOPE OF THE BAYESIAN MODEL IN CRIMINAL CASES

In the previous section, I gave an example of the application of the Bayesian model to forensic evidence. Because the Bayesian model can not only be used quantitatively, that is, in numbers (e.g. an LR of 950), but also qualitatively, that is, verbally (e.g. "much more likely"), it is in principle suitable for application to all evidence, including evidence for which no relevant data from representative samples or experiments are available.

It has been argued that legal psychologists should also report according to the Bayesian model. [15] For example, a psychologist could report that it is 'slightly more likely' or 'more likely' that this witness recognised the suspect as the perpetrator (H1) than that he is not the perpetrator (H2). Again, note that the psychological expert also reports on the LR, the conditional probability of the evidence given the hypotheses, not on the posterior, the conditional probability of the hypotheses given the evidence.

If, in principle, a qualitative or quantitative LR can be assigned to all evidence, the question arises as to whether judges themselves may or even should evaluate evidence according to the Bayesian model, particularly evidence that experts cannot report on because there is insufficient information available for a scientifically sound estimation. If the answer to these questions is affirmative, then a logical follow-up question is whether the Bayesian model can be applied to analyse a criminal case as a whole.

This follow-up question is the subject of debate. A first question is what is meant by a case as a whole. 'As a whole' does not mean that all the information in the file is modelled, but it does mean that at least the central or ultimate (alternative) hypotheses (in short: minimally the guilt and innocence hypotheses) and the most important evidence are analysed in relation to each other. An important question is who should

---

[14]   The Bayesian method is briefly explained in NFI 2017, in particular pp. 3-5 (in Dutch).

[15]   See Rassin, Arbiyah, Boskovic., Otgaar., & Merckelbach, (2022). Likelihood ratios in psychological expert opinion, and their reception by professional judges. *The International Journal of Evidence & Proof, 26*(4), 325-341. *https://doi.org/10.1177/13657127221119545*. For an example of the use of LR in a forensic psychological report, see the conclusion of A-G P.C. Vegter of 16 June 2020 (ECLI:NL:PHR:2020:560). More on this in section 4.3.

or may perform such an integral Bayesian analysis.[16] This presents a dilemma.[17] According to Dutch law it is the task of the court to select and evaluate the evidence, but judges are not experts in Bayesian modelling.[18] It is therefore, at least at present, unlikely that judges are able to carry out a Bayesian analysis on their own. An obvious alternative is for an expert to carry out such an analysis for the court.

However, there are at least two arguments against this. First, the expert would not only have to be an expert in Bayesian modelling, but also a 'material' expert in all types of evidence for which experts have not provided LR. Of course, no such 'omniscient' experts exist. A second objection is that even if we could find such 'omniscient' experts, they would be sitting in the judge's chair as a result of the choices they make in selecting, evaluating and combining the hypotheses and evidence. This would be unacceptable from a rule of law perspective.

An alternative is for judges themselves to select and evaluate hypotheses and evidence, but to be assisted in their Bayesian modelling, for example by the court's forensic advisor.[19] However, this does not solve the problem, at least not entirely. After all, even then there is a risk that the forensic adviser's modelling proposals will exert undue influence on the judge's evidential decision.

All this brings us to the following four questions:

1. Can the Bayesian model be used to determine the probative force (the LR) of individual evidence?

a.   By forensic experts?

b.   By non-forensic experts, for example by legal psychologists?

---

[16] I take the term "integral" from the judgment of the Court of Appeal in The Hague of 14 October 2015 (ECLI:NL:GHDHA:2015:2860) that will be discussed in section 4.1.

[17] More on this dilemma in A.R. Mackor & R. Schutgens, De kansendeskundige als lekenrechter: Artikel 116 Grondwet, 2022, [The probability expert as lay judge: Article 116 of the Constitution], in B. Rijpkema & A. Ellian (eds.). *Een nieuw commentaar op de grondwet.* Boom Publishers Amsterdam, 2022, 451-466. See also A.R. Mackor, Risks of Incorrect Use of Probabilities in Court and What to Do about Them, in: *Risk and Responsibility in Context*, in A. Placani & S. Broadhead (eds.). New York and London: Routledge, 2024, pp. 94-108.

[18] In fact, research shows that judges regularly misunderstand Bayesian reports. See Jan de Keijser & Henk Elffers, 'Understanding of Forensic Expert Reports by Judges, Defense Lawyers and Forensic Professionals', *Psychology, Crime & Law* 2012, 18 (2): 191-207 for an experiment with, among others, Dutch judges. For examples of the prosecutor's fallacy in Dutch case law, see, for example, H. Prakken, Kansoordelen door deskundigen: over "logisch" rapporteren en wat daarbij mis kan gaan. [Probability assessments by experts: on "logical" reporting and what can go wrong.] *Ars Aequi* 67 (2018), 740-747; see also the NFI professional supplement pp. 7-8.

[19] Forensic advisors are appointed at all Dutch criminal courts. For more information about the role of forensic advisor, see Rosanne de Roo, Wanda Remijn, Jaimy Meeuwissen, Jeltje Kruithof-van Esch, Suzanne van der Heijden, Michelle Claushuis & Laura van Blijswijk-Kieftenbeld, Forensisch adviseurs actief bij alle gerechten in Nederland: tijd voor een terugblik [Forensic advisors active in all courts in the Netherlands: time for a review], *Expertise en Recht* 2021-6, 234-239.

2.   Can the Bayesian model be used for an integral modelling of a case as a whole?

   a.   By an expert?
   b.   By the court itself, possibly with the help of the court's forensic advisor?

In the following section, we discuss the response of Dutch case law to question 2a. Section 5 discusses question 2b from a broader philosophical perspective.

4.   DUTCH CASE LAW ON BAYESIAN MODELLING OF CRIMINAL CASES AS A WHOLE

4.1.   The case of 'The six of Breda'

At first glance, these legal-philosophical, constitutional and epistemological questions may seem purely theoretical. However, they have already been addressed several times in Dutch case law, viz in cases in which the Public Prosecution or the defence had submitted a report containing such an integral analysis. According to the Court of Appeal The Hague, this happened for the first time in its ruling of 14 October 2015 (ECLI:NL:GHDHA:2015:2860) in the review case of the six from Breda.[20] The ruling explicitly addresses the use of a report submitted by the Public Prosecution containing a integral Bayesian analysis of the case.[21] The Court did not take the conclusion(s) of the report into consideration in its assessment of the evidence. It justified its decision as follows. The Court states that it has

> " […] found that there is insufficient scientific consensus on the question of whether the method of (classical) Bayesian analysis used by [expert 1] can be reliably applied to and in complex criminal cases such as the present one, in which there are many findings that cannot always be regarded as independent of each other. […]"
> "Moreover, there appear to be considerable differences of opinion among experts as to which (and how) findings in complex criminal cases should be included in the analysis, how the (a priori) probabilities of the findings can be (reliably) determined or estimated, and what expertise is necessary for that determination or estimation. This leads the court to conclude that, given the current state of scientific knowledge, the reliability of the aforementioned method and, therefore, of the results obtained from its use, is too uncertain for actual application in a complex criminal case such as the one at hand."

Which of the four questions that were distinguished in section 3 does the Court answer in this ruling? The wording 'application of Bayesian analysis in complex criminal cases' is somewhat misleading. This case does not concern the application of

---

[20]   Under 2.2, the court of appeal writes that "... to the court's knowledge, such a more comprehensive Bayesian analysis has not yet been used in a criminal case."

[21]   The heading of 2.2 reads: 'The admissibility and reliability as evidence of the Bayesian report', whereas this concerns not evidence but an expert report on the evidence. This wording may be a consequence of the defence's position, cited in 2.2, that the report 'is completely unusable as either a guideline or as evidence in the assessment of the present criminal case'.

the Bayesian model by a forensic expert to individual evidence in a complex case (question 1a), but an integral Bayesian analysis of a complex case as a whole by an 'omniscient' expert (question 2a).

The Court of Appeal rightly notes that in (an integral analysis of) a complex case many choices have to be made and that this requires many different types of expertise. In this case, the expert must estimate the prior probabilities of the hypotheses and the LR's of all the evidence. The court of appeal seems to incorrectly conclude from this correct insight that the reliability of the Bayesian method as such is too uncertain for application.

However, as was already stated in section 2, the method itself is not disputed among forensic experts. The reliability of the Bayesian method comes into question in the selection and evaluation of the hypotheses, the evidence and their interdependencies; not in the method as such. The reliability of the application is particularly questionable when an expert performs an integral analysis (question 2a), because there is a high probability that his selection and evaluation will not, or at least not entirely, correspond to that of the court, thereby placing the expert on the court's chair. [22] [23]

However, this argument against an integral analysis does not hold if judges themselves apply the Bayesian method, for example during deliberations in the court chamber (question 2b). The argument is even less sound if judges do not rely solely on this method, but primarily use it as a means of inspection after they have arrived at a preliminary assessment of the evidence in a 'traditional' manner. They can compare their preliminary assessment of the evidence with the outcome of a more or less detailed integral Bayesian analysis. [24] If there are (significant) differences, this may

---

[22]   More on this topic in section 3.

[23]   In other words, if the inter-rater reliability of the application of the method is low, the expert takes the place of the judge. See also Prakken and Meester 2017. "Reliability means ... whether different people applying this method to the same case would arrive at the same outcome. The question is to what extent Bayesian network models of complex criminal cases can be reliable in this sense. ... In Bayesian modelling, at least the following choices must be made: 1. Formulating the hypotheses, 2. Selecting potentially relevant findings, 3. Determining statistical dependency relationships between the various findings given the hypotheses, 4. Estimating or positing the relevant probabilities, 5. Determining whether further potential findings could be relevant."

[24]   In the Netherlands, an over-simplification of an integral Bayesian analysis has been informally referred to as "Bayes on a beer mat". Marjan Sjerps, forensic statistician at the Dutch Forensic Institute NFI, writes about this in a personal capacity: 'I think it can be useful for legal decision-makers ... to take a very rough look at what the most important pieces of evidence are, what the whole thing is worth approximately and how likely it is that the Public Prosecution Service is right. For example, for criminal judges who—after the usual deliberations—make this calculation, it can be an extra check to see whether the decision is reasonable." (personal e-mail correspondence, 13 August 2025). Also note that in such cases, support from a forensic advisor also seems less problematic, because there is already a draft evidence assessment that has been established in the 'traditional' manner and the question of whether and why the Bayesian modelling corresponds to or deviates from the 'traditional' evidence assessment is explicitly addressed.

give rise to critical reflection on possible errors of reasoning in the 'traditional' and/or Bayesian analysis. In this way, a Bayesian analysis can contribute to the 'sharpening function' of the reasoning in the evidential decision. In section 5 this option will be discussed more extensively.

4.2. Four later court rulings[25]

In four later court rulings, other courts have also decided on the use of the Bayesian method by an expert in the analysis of a case as a whole. On 23 May 2016, the Zeeland West-Brabant District Court (ECLI:NL:RBZWB:2016:3060), like the Court of Appeal The Hague, did not take the expert's report into consideration.

The district court writes:

> "Derksen's [the expert, ARM] calculation, which would appear to show that the defendant is, in short, probably innocent, was made according to Bayes' theorem. In the opinion of the court, this means that a controversial rule from probability theory has been used for the purpose of establishing the truth in criminal proceedings, the outcome of which depends to a large extent on the selection and evaluation of the evidence made by Derksen. The court does not rule out that this method, if applied correctly, can be an aid to finding the truth, but the ultimate answer to the question of whether a charged offence can be proven legally and convincingly is not a question that lends itself to probability theory. The answer to that question is and remains reserved for the judge. In addition, in this case, the court makes a different selection and assessment of the evidence than Derksen."

Like the court of appeal of the Hague, the district court answers question 2a in the negative. Unfortunately, this court too fails to make a clear distinction between the application of the Bayesian method to individual evidence (question 1) and to a criminal case as a whole (question 2). Also, when the district court writes that Bayes' theorem is not undisputed, it confuses —like the court of appeal of The Hague— the validity of the method as such with the inter-rater reliability of its application by an expert in the specific case.[26]

The district court corrects itself by noting that the Bayesian method may be an aid to find the truth, but then states without further argument that the ultimate question of proving guilt does not lend itself to probability theory. The question

---

[25] The case law on the use of Bayesian modelling in criminal cases was selected by means of the search terms "Bayes" (17 hits), "Bayesian" (17 hits) and "Bayesian" (94 hits). I examined all court rulings and conclusions of Advocates General at the Supreme Court in the field of criminal law up to 1 October 2025 that have been published on the website of the judiciary (*https://uitspraken.rechtspraak.nl/*). The total number of rulings is not the sum of these numbers; some rulings contain more than one term. In criminal law, the term "Bayesian" first appears in the judgment of the Court of Appeal in 's-Hertogenbosch of 24 September 2008 (ECLI:NL:GHSHE:2008:BF2188). Incidentally, the first hit in case law is not in criminal law, but in tax law. It concerns the conclusion of Advocate General Niessen of 14 March 2008 (ECLI:NL:PHR:2008:AU0837). The vast majority of case law concern the application of the Bayesian model to individual forensic evidence.

[26] See note 23 on inter-rater reliability.

of how the judge can rationally reach a verdict of guilt without using some form of probability theory remains unanswered.[27] Finally, the court also fails to distinguish between rejecting an integral Bayesian analysis by an 'omniscient' expert (question 2a) an integral Bayesian analysis by the court itself, whereby the court retains control over the selection and evaluation of the evidence (question 2b).

On 5 February 2018, the Court of Appeal in 's-Hertogenbosch (ECLI:NL:GH-SHE:2018:421), in the appeal in this case, also did not take Derksen's reports into consideration. The Court of Appeal argued at length that Derksen's selection of facts was arbitrary and that his argument was incomprehensible in some respects. However, the arguments put forward by the court against the use of the Bayesian method in general are less convincing. Under the heading "Research reports by Prof. Dr T. Derksen", the court writes:

> "… the Bayesian method, a quantitative research method that views legal evidence as thinking about probabilities, has been subject to much criticism in the literature. First of all, this method assumes that, given the hypotheses, all means of evidence are statistically independent of each other. In addition, it is doubtful whether the probabilities of the components of the input in the Bayesian method can be expressed in reliable numbers. The use of numbers creates an appearance of objectivity. The a priori probability of a component to be introduced can only be estimated subjectively, and this subjective probability cannot be tested objectively. Finally, the question is whether degrees of belief must comply with the axioms of probability theory."

I will limit myself to a few observations. First, this court too fails to distinguish between question 1a, question 2a and question 2b. Second, is noteworthy that the court raises critical questions about the Bayesian method, but at the same time allows a Bayesian analysis of individual forensic evidence. Third, the court incorrectly states that the Bayesian method assumes that evidence is independent.[28] In fact, the strength of the Bayesian method is that it can deal with both dependent and independent evidence. More on this in section 5. Fourth, the court incorrectly equates the quantitative application of the Bayesian method with the Bayesian model perse. We have seen that the Bayesian model can also be applied qualitatively. Furthermore, the court refers to "a subjective probability that cannot be tested objectively", without explaining what it means by subjective and objective. The term "subjective" suggests arbitrariness, but that is not necessarily true.[29] Finally, the court ignores the fact that the current judicial 'free evaluation of evidence' is less transparent than a Bayesian analysis.

The court also makes two interesting remarks. The first is about the reliability of the numbers and the false appearance of objectivity numbers can create. I will discuss this topic in section 5. Second, the court poses the question whether degrees of belief must

---

[27]   At least two alternatives approaches could be used, viz an argumentative or an explanation-based or scenario approach. See note 2.

[28]   This argument only applies to a simplified method also known as "linear Bayes".

[29]   Subjective in the Bayesian sense means that probabilities are not properties of reality, but of our minds. However, these subjective Bayesian beliefs need not be subjective in the sense of irrational. They can be based on scientific information and they can be discussed and assessed intersubjectively.

comply with the axioms of (Bayesian) probability theory. The answer to this question is not necessarily affirmative; other approaches exist.[30] However, the court does not make it clear which approach it would consider as an alternative to the Bayesian model. Moreover, now that the court does allow the Bayesian model for individual evidence (question 1), it is unclear how, in the court's view, these different approaches could or should be combined in a Bayesian analysis of a case as a whole (question 2).

The next case also concerns a Bayesian analysis of a criminal case as a whole. On 30 May 2017, the Court of Appeal in 's-Hertogenbosch (ECLI:NL:GH-SHE:2017:2300) rejected the Public Prosecution's request to appoint Alkemade as an expert.[31] The court limited itself to a brief statement of reasons in line with both previous rulings:

> "As argued by the Public Prosecution and the defence, and as is also apparent from the literature on the subject, Bayesian analysis in complex criminal cases is not without controversy. The reliability of the Bayesian analysis method is still uncertain according to the current state of science."

This ruling also fails to make it clear that the problem with Bayes does not lie in the use of the Bayesian model in complex criminal cases as such. After all, it is precisely in complex cases that NFI experts regularly provide Bayesian analyses of forensic evidence (question 1a). Again, the problem in this case lies solely with an integral Bayesian analysis by an 'omniscient' expert (question 2a).[32]

Three years later, on 12 November 2020, the same court of appeal in 's-Hertogenbosch (ECLI:NL:GHSHE:2020:3483) cleared up one of the misunderstandings created by the rulings discussed above. The court wrote that:

> "[...] there is consensus on the use of the Bayesian method in calculating the probative force of forensic evidence alone at forensic institutes in the Netherlands and abroad. The defence counsel's argument apparently refers to the use of the Bayesian method to analyse the entire case. However, that situation does not apply in this case, as only the forensic evidence has been examined. The court therefore, like the district court, sees no reason to doubt the Bayesian method used by the NFI and the findings of the investigation carried out on the basis of that method."[33]

---

[30] One of the best known is the Dempster-Shafer theory of belief functions.

[31] In another case in which the same expert, again at the request of the Public Prosecution, had drawn up a Bayesian report, the Court of Appeal in 's-Hertogenbosch, 22 November 2016, ECLI:NL:GHSHE:2016:5165, did not pass judgement on the use of the Bayesian model as such. It merely states: 'In view of the objections raised by Prof. H. Prakken in the counter-investigation, the court did not have the report by Dr F.J.M. Alkemade contribute to the court's ruling on the evidence decision.' On these cases, also see H. Prakken, Rechters en Raadsheren over Bayes [Judges and Justices on Bayes]. In R. Horselenberg, V. van Koppen, & J. de Keijser (eds.), *Bakens in de Rechtspsychologie: Liber Amoricum voor Peter van Koppen,* Boom Uitgevers, 2020, 287-300.

[32] The court of appeal also explicitly states however: "The defence's assertion that the Public Prosecution should not be allowed to introduce a report by Dr Alkemade into the proceedings is not supported by law." The Court of Appeal reiterates this view in its ruling of 9 October 2018 (ECLI:NL:GHSHE:2018:4181). Accordingly, the admissibility of a report with an integral Bayesian analysis is not under discussion.

[33] Other courts had previously ruled in a similar vein. On 25 February 2015 (ECLI:NL:RBLIM:2015:1588*),* the Limburg District Court ruled in the first instance in the same case

Thus, the court explicitly states that the Bayesian method is generally accepted for the evaluation of forensic evidence (question 1a) and—or so I would add—also in complex criminal cases. The court also states that this method is only open to debate when it is used to analyse a case as a whole (question 2). However, the ruling leaves open the question of whether the Bayesian method is acceptable when non-forensic experts, such as legal psychologists, apply it to individual pieces of evidence (question 1b). Nor does the court rule on the question of whether there is scope for an integrated Bayesian analysis when this is carried out by the judge himself (question 2b).

4.3.   The silence of the Dutch Supreme Court

Unfortunately, so far only this ruling by the Court of Appeal in 's-Hertogenbosch, which has removed some of the ambiguities, has been issued. Two other cases in which the defence submitted critical questions about the use of the Bayesian method to the Dutch Supreme Court were disposed of by the Supreme Court on the basis of Article 81(1) of the Judicial Organisation Act.[34] Also neither of the two Advocates General at the Supreme court provided clarity in their conclusions.

In the first case (Supreme Court, 25 August 2020, ECLI:NL:HR:2020:1317), the defence criticises the use of the Bayesian method by a legal psychologist (question 1b). In the conclusion of 16 June 2020 (PHR, ECLI:NL:PHR:2020:560), the Advocate General dismisses this criticism by claiming that an evaluation of the reliability of the method is factual in nature and does not fit within the cassation proceedings. This view may not be legally contestable, but it can be critized on more fundamental grounds. The question of whether a method has been applied reliably in a specific case may be a question that must be answered by the trial court, but answering the question of whether a method as such is suitable for the specific type of application does not seem to be reserved exclusively for the trial court. In any case, the Advocate

---

that "...the method of investigation used by the NFI, namely the Bayesian method, is generally accepted in science and is widely regarded as a scientifically sound method of investigation". On 16 February 2016 (ECLI:NL:RBGEL:2016:825), the Gelderland District Court stated: "The experts have stated that the Bayesian method was used in the "Breda Six" case cited by the defence. However, according to the experts, in the present case, a Bayesian evaluation of the scientific probative force of a reported DNA match was used. The evaluation of the probative force therefore only concerns the DNA evidence. The comparison with the "Breda Six" is therefore incorrect.' On 19 December 2017, the District Court of East Brabant stated (ECLI:NL:RBOBR:2017:6556): 'Bayesian statistics have been used for years, both nationally and internationally, to analyse forensic evidence. It is a logically correct way of forming an expert opinion, explicitly avoiding making statements outside the area of expertise. The use of this form of statistics in, for example, the field of forensic DNA research has long been considered standard practice and is generally uncontroversial."

[34]   Article 81(1) of the Dutch Code on the Judicial Organisation reads: "If the Supreme Court finds that a complaint lodged cannot lead to cassation and does not require a response to legal questions in the interest of legal uniformity or legal development, it may limit itself to this finding when stating the grounds for its decision."

General could have taken the opportunity to clarify that the application of Bayes by legal psychologists (question 1b) does not differ in principle from its application by forensic investigators (question 1a) and therefore—in line with case law, more specifically the above-mentioned ruling of the Court of Appeal in 's-Hertogenbosch of 12 November 2020—is in principle unproblematic.

In the second case (Supreme Court, 15 June 2021, ECLI:NL:HR:2021:788), the defence criticises the use of the Bayesian model, but does not make a clear distinction between the generally accepted use of Bayes by forensic experts on individual evidence (question 1a) and the controversial use of the Bayesian model for an integrated analysis of a case as a whole (question 2).[35] Here too, the Advocate General in his conclusion of 20 April 2021 (PHR, ECLI:NL:PHR:2021:391), fails to take the opportunity to provide clarity on this distinction and on the use of the Bayesian model in criminal cases. The Advocate General could minimally have made it clear that the use of the Bayesian method by forensic experts to determine the probative force of individual forensic evidence is undisputed.

5.   FURTHER CONSIDERATIONS REGARDING BAYESIAN MODELLING OF CRIMINAL CASES AS A WHOLE

Partly in response to the case law from 2015, 2016 and 2017 that was discussed in section 4, a brief but heated debate has taken place in Dutch literature on the question of whether there should be room for Bayesian analyses of cases as a whole by experts. Prakken and Meester have explicitly spoken out against it.[36] At least for now, they only see a role for the Bayesian method in the analysis of individual (forensic) evidence and for qualitative and global use to recognise and prevent probabilistic

---

[35] The case in question only concerns the application by forensic experts to individual evidence, not an integral analysis of a case as a whole. Another and more important complaint by the defence concerns the fact that the DNA traces were only examined at source level (the question of whether the DNA originated from the suspect) and not also at activity level (the question of how what activity caused the DNA trace). See section 5 on this distinction.

[36] Prakken and Meester 2017. In their article, the authors discuss whether Bayes' rule can be applied in a scientifically reliable and useful manner in court to analyse complex criminal cases as a whole and whether such use of Bayes' rule in court is practically possible and, if so, under what conditions. Their article has also been noted in practice; the conclusion of Advocate General Hofstee of 20 April 2021 (ECLI:NL:PHR:2021:391) shows that the defence referred to their article. Incidentally, Prakken was a counter-expert in two cases in which Alkemade acted as an expert, and in that capacity he sharply criticised the content of both of Alkemade's reports and pointed out probabilistic errors in reasoning therein. See Frans Alkemade, Makkers, staakt uw wild geraas. Een oproep tot de Nederlandse deskundigen in forensic reasoning om een consensus te formuleren [Friends, cease your wild clamour. A call to Dutch experts in forensic reasoning to formulate a consensus], *Expertise en Recht* 2018 (1), 26-35 for a response to Prakken and Meester 2017.

errors in reasoning.[37] However, Prakken and Meester do not address the question of who could perform an integral Bayesian analysis, at what stage, and for what purpose. In order to arrive at a balanced judgement on the use of Bayesian analyses of criminal cases as a whole, we must minimally distinguish between the following questions.[38]

1   Who performs the Bayesian analysis?
At a minimum, a distinction must be made between:
A   Expert or forensic advisor
B   Judge(s)
C   Judge(s) with the support of a forensic advisor

2   At what stage is the analysis used?
At a minimum, a distinction must be made between:
A   Hearing and proceedings[39]
B   Court chamber
C   Judicial ruling

3   What is the function of the analysis?
At a minimum, a distinction must be made between:
A   Contributing to the investigation at the hearing
B   Sharpening the judicial thought process and, related to this, internal check of the evidence decision during the court chamber's meeting
C   Contributing to the justification of the evidence decision in the ruling[40]

As part of the research project Preventing Miscarriages of Justice[41], our research group is studying the added value of a combined use of the scenario approach and the Bayesian model in the evaluation of evidence in criminal cases.[42] Here I will limit myself to a discussion of the research of our group into the use of the Bayesian model. Dahlman has proposed a new systematisation of probabilistic thinking errors.[43] Based

---

[37] For examples of a qualitative and global analysis of several court rulings, see R. Meester and L. Stevens, Correct redeneren: wat een Bayesiaanse analyse zegt over vier recente strafrechtelijke uitspraken [Correct reasoning: what a Bayesian analysis says about four recent criminal court rulings]. *Expertise en Recht*, 2021(3), 112-121.
[38] These distinctions are not intended to be exhaustive. I limit myself to the function of Bayesian analysis for the judiciary. I do not discuss its role for investigation, prosecution and defence.
[39] According to Dutch case law, expert reports containing a Bayesian analysis of a case as a whole may be brought into the proceedings (2a). See note 32.
[40] Above we have seen that the Court of Appeal in 's-Hertogenbosch states that it did not have the report on the integral Bayesian modelling contribute to its ruling on the evidence decision. See note 31.
[41] See note *.
[42] On the scenario approach, see A.R. Mackor and P.J. van Koppen, The scenario theory about evidence in criminal law. In: Christian Dahlman, Alex Stein, and Giovanni Tuzet (eds.), *Philosophical Foundations of Evidence Law*, Oxford University Press, 2021, pp 2015-230. DOI: 10.1093/ oso/ 9780198859307.003.0016. For a recent proposal on how to combine the scenario approach with the Bayesian method in a Bayesian Network, see C. Dahlman and H. Jellema, Scenario-Based Bayesian Networks for Legal Evidence, submitted.
[43] Dahlman 2024.

on this systematization he investigated whether and to what extent training in Bayesian reasoning helps lawyers to recognise these probabilistic thinking errors.[44] Jellema has translated Dahlman's overview of errors in reasoning into Dutch and made it accessible to practising lawyers, among others by providing examples from Dutch case law.[45] We plan to convert this overview into a checklist with 'critical questions' that lawyers can ask when analysing evidence. This checklist should help prevent thinking errors such as the prosecutor's fallacy which was discussed in section 2.

Our research group also investigates whether judges themselves can model court cases as a whole using the Bayesian model. We focus in particular on the phase of the court's deliberation chamber in which, as already suggested in section 4.1, judges arrive at an evidential judgement in a 'traditional' manner and compare this judgement with the outcome of their own Bayesian analysis. In our view, this analysis should consist of setting up a Bayesian Network with the help of a forensic advisor. In terms of the distinctions mentioned above, we focus on the combination of questions 1c, 2b and 3b.

What is a Bayesian Network (BN)? A BN looks like a mind map. However, it is a mind map whose nodes and connections are precisely defined. It is used to visualise and analyse the often complex dependencies between hypotheses and evidence.[46]

When drawing up a BN, the primary focus is not on the quantitative outcome that a BN can deliver, but rather on the fact that the activity of drawing up a BN jointly can support both the individual thought process and the discussion between judges. It should be emphasised that there is no single correct, true or best BN, but rather a BN on which the drafters have agreed or disagreed after consultation and discussion and which they can explain and justify to others.[47] Drawing up a BN can help the drafters to think in a structured way about complex relationships between hypotheses and evidence, making their assumptions transparent and open to discussion and potentially preventing probabilistic errors in reasoning (see Figure 1).

All this might still sound rather abstract, so let me give two examples. Suppose there are two witness statements in a criminal case. Drawing up a BN in combination with the aforementioned checklist of critical questions 'forces' the authors to think about the 'critical question' of whether the two witnesses spoke to each other or made their statements completely independently of each other. The answer to this critical question ('Is the evidence dependent or independent?') is important because

---

[44] C. Dahlman, Training legal fact-finders to recognise probabilistic fallacies, *Law, Probability and Risk* Volume 24, issue 1, 2025, https://doi.org/10.1093/lpr/mgaf010.

[45] Jellema and Dahlman 2025. The use of a checklist is in line with the view of Prakken and Meester 2017, who write: 'The power of Bayesian thinking lies primarily in its therapeutic and pedagogical value: with small examples, it is relatively easy to show that certain seductive patterns of thinking about probabilities are incorrect.' and: 'This rule arms us against errors in reasoning.'

[46] For an accessible introduction to Bayesian Networks, see David A. Lagnado, *Explaining the Evidence,* Cambridge, Cambridge University Press, 2022, chapter 3.

[47] It can be said, however, that some models are logically incorrect.

independent evidence has greater probative force than dependent evidence. This also shows in the different modelling of the witness statements (see Figure 2).
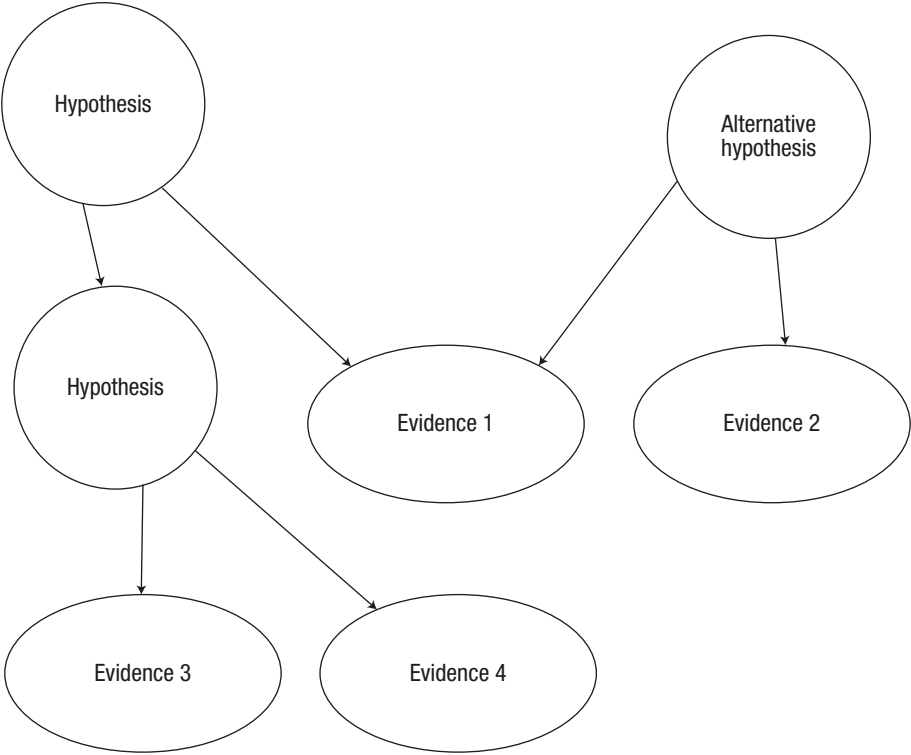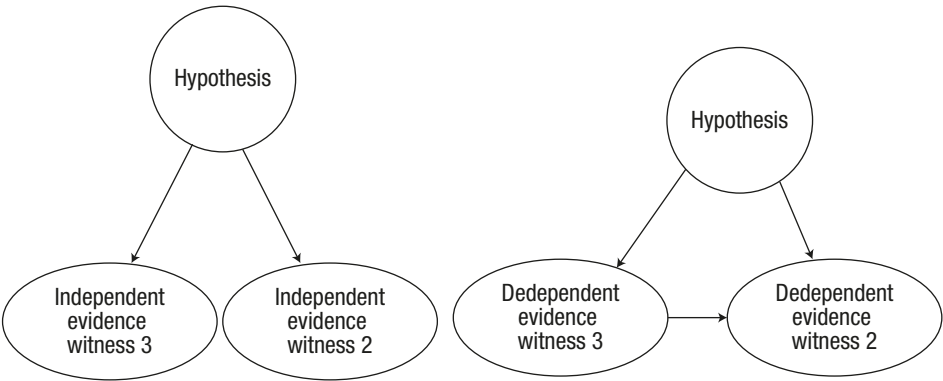
Figure 1. Bayesian Network

Figure 2. BN's with dependent and independent evidence

Another example: in a court case, the NFI has written a report on DNA evidence. Creating a BN helps the judges to consider the 'critical question' of exactly which question the forensic report answers. Does the report provide a LR about the source of the DNA, i.e. whether the the defendant is the source of the DNA, or does it provide a LR about the activity that caused the DNA to end up at the crime scene, for example through an act related to the crime such as strangulation? This distinction is important because the probative force of DNA at the activity level is often much lower than at the source level. Ignoring this distinction can therefore lead to an overestimation of the probative force. Here too, drawing up a BN in combination with asking the 'critical question' ("To which hypothesis does the LR relate?", "Source or activity?") helps to raise awareness of this crucial distinction which is also visible in the two BNs. [48]

Figure 3. BN's of DNA analysis at source and activity level

<hr>
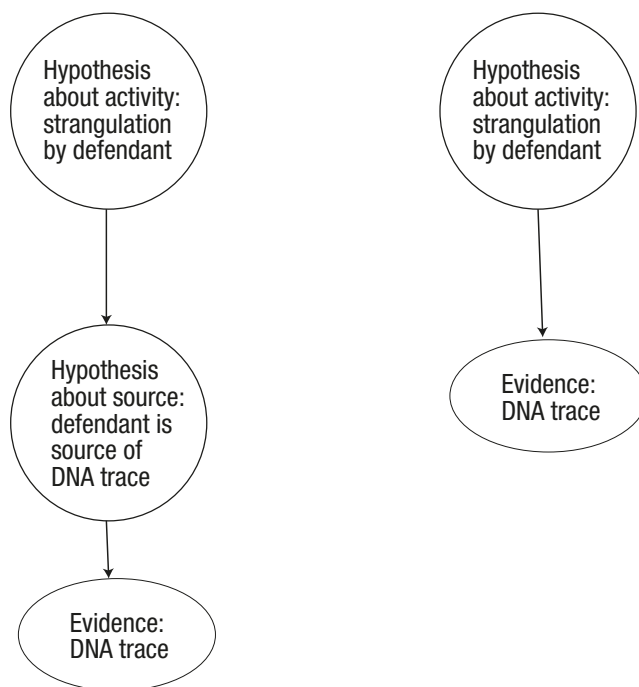
[48] Note that a similar error can also occur in the example of shoe prints that was discussed in section 2. The NFI reports on the suspect's shoe, not on the hypothesis that it was the suspect who wore the shoe and made the print. The LR that the suspect made the print is possibly much smaller than the LR that the print was made with his shoe.

Let us return to the use of Bayesian Networks by judges in the court's deliberation chamber. As mentioned above, the primary focus is not on the end product, but on the process of jointly creating a BN. Drawing up a BN makes both choices and errors in reasoning transparent, so that they can be discussed. The creation of a BN ultimately also leads to a product, a network. This network can be qualitative (as in the images above) or quantitative. When creators assign numbers to the prior probabilities and the LRs, the software used to create BNs can perform the often highly complex calculations for the user.[49]

In section 4.2, we saw that the Court of Appeal in 's-Hertogenbosch in its ruling of 5 February 2018 (ECLI:NL:GHSHE:2018:421) argues that quantifying subjective probability estimates gives the calculations an unjustified appearance of objectivity. This has been referred to as 'garbage in, garbage out' (Prakken and Meester, 2017, p. 191). This critique certainly has some merit. It is a legitimate question to ask what value numbers add if they are not scientifically substantiated. The main answer is that judges have to make a probability assessment anyway. Assigning a number to verbal expressions such as "probable" or "very probable" can ensure that judges use these terms in a more uniform manner. Like the NFI, judges too can indicate a range of numbers, for example a lower limit of 10 and an upper limit of 100. Another important function of quantification is that, in practice, it proves difficult to intuitively estimate the effect of combining prior probabilities with several LRs. By quantifying these probabilities and varying the numbers within a range which is deemed reasonable by judges, the creators of the BN can gain insight into the effect of the estimates they make. If the quantitative outcomes of the BN differ significantly from the qualitative outcome of the 'traditional' legal deliberations, this may be a reason to critically compare the 'traditional' evidence reasoning and the BN. But even if there is no reason to change the evidentiary decision, drawing up a BN can contribute to better evidential reasoning because it leads to clarification and careful consideration of choices.[50]

6.   CONCLUSION AND RECOMMENDATIONS

The Bayesian model is internationally accepted for estimating the likelihood ratio (LR), i.e. the probative force of forensic evidence. The Court of Appeal in 's-Hertogenbosch adopted this position in its ruling of 12 November 2020 (question 1a). However, it is unclear whether and under what conditions the method is also accepted in Dutch case law for analyses by non-forensic experts such as legal psychologists (question 1b). Bayesian analyses of criminal cases as a whole by an 'omniscient' ex-

---

[49] Software to create BN's such as Hugin and AgenaRisk is comparable to the software of a calculator that performs calculations for us based on our input. Of course, this quantification step assumes that the judge has quantified the values of priors and evidence.

[50] More on the constitutional importance of sound reasoning based on evidence in Mackor 2024.

pert have not been used been by lower courts (question 2a). However, there is still no clarity about the use of an integral Bayesian analysis drawn up by the court itself (question 2b). It is to be hoped that the Dutch Supreme Court will provide clarity on these topics.

I have argued that the lower courts have rightly rejected the modelling of criminal cases as a whole by experts. The main argument is not that 'the method is still too uncertain', but that this would allow experts to sit on the seat of the judge, which is unacceptable from a rule of law perspective. However, we must avoid throwing the Bayesian baby out with the bathwater. Further research is needed to determine whether and how the Bayesian model can help judges. In particular we want to investigate whether the use of a checklist of 'critical questions' can help prevent probabilistic errors in reasoning and thus improve the quality of judicial evidence evaluation. In addition, it must be investigated whether the creation of Bayesian Networks by judges in the court's deliberation chamber can improve the quality of judicial evidence evaluation. First, it is important to investigate whether it is feasible for judges to create a BN themselves. Feasibility in this context concerns both the time it takes to create a BN and whether judges are capable of creating a sound BN on their own or in collaboration with a forensic advisor. It is also, from a rule of law perspective, necessary to investigate the influence of the forensic advisor on the modelling. Theoretical and empirical research is needed to answer these questions. Over the next three years, we hope to be able to provide answers to these and other questions as part of our NWO research project Preventing Miscarriages of Justice. [51] [52]

---

[51]   See note *. For recent work and work in progress of our research group, see Thomas Boersma, Bayesiaanse netwerken in bewijsminimumzaken: toepassing op de zaak Ali B [Bayesian Networks in minimal evidence cases: an application to the case of Ali B], in progress and Thomas Boersma, Logische en correcte bewijswaardering: Bayesiaans redeneren in zaken met seksueel grensoverschrijdend gedrag, [Logical and correct evidence evaluation: Bayesian reasoning in cases of sexual transgressive behaviour], Tijdschrift Recht en Arbeid, forthcoming (the first Dutch Bayesian article on labour law). C. Dahlman and H. Jellema, submitted, on combining Bayesian and scenario approaches. Leya Hampson and Ludi van Leeuwen, Investigating the value of qualitative Bayesian networks of complete cases as "double-check" tools on traditional judicial reasoning: An exploratory study, forthcoming in the *Proceedings of the AI4EVIR Workshop on AI for evidential reasoning,* December 9, 2025, Turin, Italy. A.R. Mackor and H. Prakken On Reporting Likelihood Ratio's of Exhaustive and Non-Exhaustive Hypotheses about Rare Events in Criminal Case, forthcoming in the *Proceedings of the AI4EVIR Workshop on AI for evidential reasoning, December 9, 2025, Turin, Italy*.

[52]   For other relevant recent research on BN's see Nikolay Babakov, Ehud Reiter and Alberto Bugarín, Explaining Bayesian Networks Reasoning to the General Public: Insights from the User Study, *Multimodal, Affective and Interactive eXplainable AI*, 2025 for a large user study (124 participants) evaluating the interpretability of BN. In our research group, however, we do not only investigate the passive understanding of BN's, but focus on active co-creation of BN's.

REFERENCES

Alkemade, F. (2018). Makkers, staakt uw wild geraas. Een oproep tot de Nederlandse deskundigen in forensic reasoning om een consensus te formuleren. *Expertise en Recht*, *2018*(1), 26-35.

Babakov, N., Reiter, E., & Bugarín, A. (2025). Explaining Bayesian Networks Reasoning to the General Public: Insights from the User Study. *Multimodal, Affective and Interactive eXplainable AI*.

Boersma, T. (forthcoming). Bayesiaanse netwerken in bewijsminimumzaken: toepassing op de zaak Ali B.

Boersma, T. (forthcoming). Logische en correcte bewijswaardering: Bayesiaans redeneren in zaken met seksueel grensoverschrijdend gedrag. *Tijdschrift Recht en Arbeid*.

Dahlman, C. (2024). A systematic account of probabilistic fallacies in legal fact-finding. *The International Journal of Evidence & Proof*, *2024*(1), 45-64.

Dahlman, C. (2025). Training legal fact-finders to recognise probabilistic fallacies. *Law, Probability and Risk*, *24*(1). *https://doi.org/10.1093/lpr/mgaf010*

Dahlman, C., & Jellema, H. (forthcoming). Scenario-Based Bayesian Networks for Legal Evidence.

De Keijser, J., & Elffers, H. (2012). Understanding of forensic expert reports by judges, defense lawyers and forensic professionals. *Psychology, Crime & Law*, *18*(2), 191-207.

De Roo, R., Remijn, W., Meeuwissen, J., Kruithof-van Esch, J., Van der Heijden, S., Claushuis, M., & Van Blijswijk-Kieftenbeld, L. (2021). Forensisch adviseurs actief bij alle gerechten in Nederland: tijd voor een terugblik. *Expertise en Recht*, *2021*(6), 234-239.

Gerechtshof Amsterdam [Tribunal de Apelación de Ámsterdam]. (2021, 12 de abril). ECLI:NL:GHAMS:2021:986.

Gerechtshof Den Bosch [Tribunal de Apelación de 's-Hertogenbosch]. (2008, 24 de septiembre). ECLI:NL:GHSHE:2008:BF2188.

Gerechtshof Den Bosch [Tribunal de Apelación de 's-Hertogenbosch]. (2016, 22 de noviembre). ECLI:NL:GHSHE:2016:5165.

Gerechtshof Den Bosch [Tribunal de Apelación de 's-Hertogenbosch]. (2017, 30 de mayo). ECLI:NL:GHSHE:2017:2300.

Gerechtshof Den Bosch [Tribunal de Apelación de 's-Hertogenbosch]. (2018, 5 de febrero). ECLI:NL:GHSHE:2018:421.

Gerechtshof Den Bosch [Tribunal de Apelación de 's-Hertogenbosch]. (2018, 9 de octubre). ECLI:NL:GHSHE:2018:4181.

Gerechtshof Den Bosch [Tribunal de Apelación de 's-Hertogenbosch]. (2020, 12 de noviembre). ECLI:NL:GHSHE:2020:3483.

Gerechtshof Den Haag [Tribunal de Apelación de La Haya]. (2015, 14 de octubre). ECLI:NL:GHDHA:2015:2860.

Hampson, L., & Van Leeuwen, L. (en prensa). Investigating the value of qualitative Bayesian networks of complete cases as "double-check" tools on traditional judicial reasoning: An exploratory study. *Proceedings of the AI4EVIR Workshop on AI for Evidential Reasoning*, Turín, Italia.

Hoge Raad [Tribunal Supremo de los Países Bajos]. (2020, 25 de agosto). ECLI:NL:HR:2020:1317.

Hoge Raad [Tribunal Supremo de los Países Bajos]. (2021, 15 de junio). ECLI:NL:HR:2021:788.

Jellema, H., & Dahlman, C. (2025). Probabilistische drogredenen in de rechtszaal: een systematisch overzicht. *Expertise en Recht*, *2025*(3), 64-74.

Koninklijk Nederlands Meteorologisch Instituut [KNMI]. (s.f.). *Het regent niet vaker, wel harder. https://www.knmi.nl/over-het-knmi/nieuws/het-regent-niet-vaker-wel-harder*

Lagnado, D. A. (2022). *Explaining the evidence*. Cambridge University Press.

Mackor, A. R. (2024). Risks of incorrect use of probabilities in court and what to do about them. En A. Placani & S. Broadhead (Eds.), *Risk and responsibility in context* (pp. 94-108). Routledge.

Mackor, A. R., & Schutgens, R. (2022). De kansendeskundige als lekenrechter: Artikel 116 Grondwet. En B. Rijpkema & A. Ellian (Eds.), *Een nieuw commentaar op de grondwet* (pp. 451-466). Boom Publishers.

Mackor, A. R., & Van Koppen, P. J. (2021). The scenario theory about evidence in criminal law. En C. Dahlman, A. Stein, & G. Tuzet (Eds.), *Philosophical foundations of evidence law* (pp. 215-230). Oxford University Press. *https://doi.org/10.1093/oso/9780198859307.003.0016*

Meester, R., & Stevens, L. (2021). Correct redeneren: wat een Bayesiaanse analyse zegt over vier recente strafrechtelijke uitspraken. *Expertise en Recht*, *2021*(3), 112-121.

Nederlands Forensisch Instituut [NFI]. (2017). *Vakbijlage Waarschijnlijkheidstermen. https://www.foren-sischinstituut.nl/publicaties/publicaties/2017/10/18/vakbijlage-waarschijnlijkheidstermen*

Niessen, R. E. C. M. (Advocaat-Generaal). (2008, 14 de marzo). Conclusión ECLI:N-L:PHR:2008:AU0837.

Parket bij de Hoge Raad [PHR]. (2020, 16 de junio). Conclusión ECLI:NL:PHR:2020:560.

Parket bij de Hoge Raad [PHR]. (2021, 20 de abril). Conclusión ECLI:NL:PHR:2021:391.

Parket bij de Hoge Raad [PHR]. (2025, 25 de mayo). Conclusión ECLI:NL:PHR:2025:378.

Prakken, H. (2018). Kansoordelen door deskundigen: over "logisch" rapporteren en wat daarbij mis kan gaan. *Ars Aequi*, *67*, 740-747.

Prakken, H. (2020). Rechters en Raadsheren over Bayes. En R. Horselenberg, V. van Koppen, & J. de Keijser (Eds.), *Bakens in de Rechtspsychologie: Liber Amicorum voor Peter van Koppen* (pp. 287-300). Boom Uitgevers.

Prakken, H., Bex, F. J., & Mackor, A. R. (2020). Editors' Review and Introduction: Models of Rational Proof in Criminal Law. *Topics in Cognitive Science*, *12*, 1053-1067.

Prakken, H., & Mackor, A. R. (en prensa). On reporting likelihood ratios of exhaustive and non-exhaustive hypotheses about rare events in criminal case. *Proceedings of the AI4EVIR Workshop on AI for Evidential Reasoning*, Turín, Italia.

Prakken, H., & Meester, R. (2017). Bayesiaanse analyses van complexe strafzaken door deskundigen. Betrouwbaar en zo ja: nuttig? *Expertise en Recht*, *2017*(5), 185-197.

Rassin, E., Arbiyah, N., Boskovic, I., Otgaar, H., & Merckelbach, H. (2022). Likelihood ratios in psychological expert opinion, and their reception by professional judges. *The International Journal of Evidence & Proof*, *26*(4), 325-341. *https://doi.org/10.1177/13657127221119545*

Rechtbank Gelderland [Tribunal de Distrito de Güeldres]. (2016, 16 de febrero). ECLI:NL:RB-GEL:2016:825.

Rechtbank Limburg [Tribunal de Distrito de Limburgo]. (2015, 25 de febrero). ECLI:NL:R-BLIM:2015:1588.

Rechtbank Oost-Brabant [Tribunal de Distrito de Brabante Oriental]. (2017, 19 de diciembre). ECLI:NL:RBOBR:2017:6556.

Rechtbank Zeeland-West-Brabant [Tribunal de Distrito de Zelanda-Brabante Occidental]. (2016, 23 de mayo). ECLI:NL:RBZWB:2016:3060.

Thompson, W. C., & Schumann, E. L. (1987). Interpretation of statistical evidence in criminal trials: The prosecutor's fallacy and the defense attorney's fallacy. *Law and Human Behavior*, *11*(3), 167-187.

Marcial
Pons

Càtedra
de Cultura
Jurídica

Universitat
de Girona

Diputació de Girona

www.quaestiofacti.com