Information Economics in the Criminal Standard of Proof

Christian Dahlman[1] & Anders Nordgaard[2,3]

1. Introduction

In the Bayesian approach to legal evidence, the criminal standard of proof (´beyond reasonable doubt´) is viewed as a rule that sets a threshold for the required probability of the *factum probandum* (Finkelstein & Fairley 1970; Kaye 1982; Schauer & Zeckhauser 1996; Hedden & Colyvan 2019). This threshold is formalized as the sufficient degree of probability (p) of the *factum probandum* (H) given the evidence (E), $P(H|E) \geq p$. A frequent critique of this approach is that the standard of proof cannot be reduced to a probability threshold, since proof beyond reasonable doubt not only requires that the evidence presented in court supports the *factum probandum* to a sufficient degree, but also requires a sufficient quantity of evidence. The latter requirement concerns the degree at which the relevant facts of the case have been investigated and presented in court, and we will therefore refer to it as a requirement of *sufficient informativeness*.

John Maynard Keynes famously talked about informativeness as 'the weight of the evidence' (Keynes 1921, 71), but this expression creates some confusion in a legal context since lawyers sometimes talk about the degree at which the evidence supports the *factum probandum* as the 'weight of the evidence'. Dale Nance has therefore suggested that the probability of the *factum probandum* given the evidence be referred to as 'Δ-weight' and the informativeness of the evidence as 'Σ-weight' (Nance 2008). We will avoid the term 'weight' in this article and discuss the two-fold standard of proof as a *requirement of sufficient probability of the factum probandum* and a *requirement of sufficient informativeness*.

Scholars who oppose the Bayesian approach to legal evidence have often claimed that sufficient informativeness cannot be modelled probabilistically (Brilmeyer 1986, 682-685; Cohen 1986, 648; Stein 2011, 241-42; Haack 2014, 61-62). We will show in this paper that they are wrong. In the following sections we will show how a requirement of sufficient informativeness can be modelled as a probabilistic decision rule. Decision theory has previously been used in the literature on the standard of proof to model sufficient probability of the *factum probandum* (Kaplan 1968; Cullison 1969; Tribe 1971; Lillquist 2002; Laudan & Sanders 2009). In this paper, we will use decision theory to model sufficient informativeness.[4]

[1] Faculty of Law, Lund University.
[2] Department of Computer and Information Science (IDA), Linköping University, and the Swedish National Forensic Centre (NFC).
[3] Thanks to Marcello Di Bello, Alex Biedermann, Hylke Jellema, Anne Ruth Mackor, Dale Nance, Franco Taroni and Lena Wahlberg for helpful suggestions on previous versions of this paper.
[4] Cf Nance 2016, 120 note 48.

The two-fold standard of proof can tentatively be formulated as follows.

> There is proof beyond reasonable doubt if the probability of the *factum probandum* given the evidence is equal or greater than p, and the facts of the case have been sufficiently investigated.

Both requirements must be met for the standard of proof to be fulfilled. If the latter requirement of sufficient informativeness is not met the defendant is acquitted, and there is no need to assess if the probability of the *factum probandum* meets the former requirement. In the following sections we will re-formulate this two-fold requirement in more specific terms and show how it can be modelled probabilistically. It should be made clear from the start that our purpose is not to investigate to what extent this two-fold interpretation of the criminal standard or proof correctly describes positive law in any particular legal system, and we do not argue that legal systems should adopt this standard of proof. The purpose of our model is only to explicate how such a standard of proof can be modelled probabilistically.

As we will show, sufficient informativeness is fundamentally about information economics, and can be formalized as a decision rule about the probability that additional evidence would lead to a switch from a sufficient probability of the *factum probandum* to an insufficient probability. This is a so-called *second order probability*.

The following sections are structured as follows. We will start with the concept of informativeness (section 2) and illustrate information deficiency in legal evidence with two court cases (section 3). We will use information economics to define sufficient informativeness (section 4), and use decision theory to model rules that optimize social utility for sufficient probability of the *factum probandum* (section 5) and sufficient informativeness (section 6). Since our model of sufficient informativeness includes a second order probability we will address the arguments against higher order probabilities that have been raised by some scholars (section 7). We will end the paper by discussing how informativeness should be handled in forensic reporting to help the fact-finder determine if the requirement of sufficient informativeness is met (section 8), and supply a practical example (section 9).

2. Informativeness and Information Deficiency

Every probability assessment is based on a specific set of information that has been 'gathered' (in some sense) by the person assessing the probability. Some assessments are based on more information than others, and it is commonly assumed that more information makes an assessment better, epistemically speaking. The gathering of information is viewed as an epistemic good, since we assume that more information will typically make the probability assessment converge to the frequency of whatever it is in the real world that is responsible for the probability under assessment.

This being said, it should also be kept in mind that more information does not make a probability assessment more 'correct' or 'true'. An assessment of subjective probability is a betting-preference, not a proposition about something in the external world, and therefore cannot be true or false (Ramsay 1926, 67; de Finetti 2017, 62-63). If a probability assessment is changed due to new information the earlier assessment should not be viewed as incorrect, but rather as a rational betting-preference given the information at the time (de Finetti 2017, 122). In terms of betting, the two-fold requirement in the criminal standard of proof can be described as follows. The requirement of sufficient probability of the *factum probandum* instructs the fact-finder how to bet, given that the fact-finder is betting on the evidence at hand, while the requirement of sufficient informativeness instructs the fact-finder whether to bet on the evidence at hand or not (Nance 2016, 171).

The literature on informativeness is scattered over many scientific disciplines and manifests itself in an archipelago of conversations that use different terminologies to discuss the same problem. The various ways to characterize the benefit of gathering more information as a basis for a probability assessment include 'completeness', 'weight', 'comprehensiveness', 'reliability', 'robustness, 'informativeness', 'accuracy', 'stability', 'confidence', 'sensitivity', 'resilience', 'tenacity' and 'reduced epistemic risk'. Some scholars discuss informativeness on a general epistemological level (e.g. Sahlin 1983; Haack 1993; Logue 1995), while others focus on a specific context. Informativeness in legal evidence has been explored by Jonathan Cohen (1977; 1986), Neil Cohen (1985), David Kaye (1986), Barbara Davidson & Robert Pargetter (1986; 1987), Richard Friedman (1997), Dale Nance (1998; 2008; 2016; 2021), Alex Stein (2005), Hock Lai Ho (2008), Susan Haack (2014), Christian Dahlman, Farhan Sarwar & Lena Wahlberg (2015), and Jordi Ferrer Beltrán (2018).

Decision makers are sometimes troubled by *information deficiency*, sensing that their probability assessment is based on insufficient information, and wishing to gather more information before committing to a decision. Such dilemmas arise in legal decision making when legal fact-finders in a criminal trial (judges or jurors) feel that a certain aspect of the case has an information gap that they wish could be remedied before they decide if the *factum probandum* is sufficiently probable given the evidence. Information gaps come in different forms. It can be a question that was not asked to a witness, or a witness that was not called to testify at all. It can be a forensic trace that was not investigated at the crime scene, or an alternative perpetrator that was never investigated by the police.

Information deficiency can be handled by the legal system in different ways. In some legal systems a case can be stopped from going to trial and sent back to the prosecutor for further investigation, and some systems allow for the court to fix information gaps by making its own inquiries. If sufficient informativeness is included as a requirement in the standard of proof the defendant can be acquitted due to information deficiency. As we explained in the introduction, this is the way of handling information deficiency that we investigate in this paper. We will model a two-fold standard of proof that requires a sufficient probability for the *factum probandum* as well as sufficient informativeness. As we shall see, these requirements are connected to each other, since the benefit of additional investigation depends on the potential of additional information to change the probability of the *factum probandum* from sufficient to insufficient.

3. Two Cases of Information Deficiency

To illustrate how a criminal standard of proof that includes a requirement of sufficient informativeness can lead to an acquittal due to information deficiency we will now introduce two hypothetical court cases. We refer to them as *The Case of the Tin Box* and *The Case of the Missing Fingers*. They are based on our experience of real trials in the Swedish legal system and serve as simplified versions of real cases, highlighting the dilemma of information deficiency.

The Case of the Tin Box unfolds as follows.

> In 2005 a man walks into a Swedish police station and says that he wants to turn himself in. His name is AA and he says that he has just killed an elderly woman who lives by herself in an apartment nearby. The police rush to the apartment and find the woman stabbed to death. In his confession AA explains to the police that he had heard that the woman kept a huge amount of cash in a tin box, and had knocked on her door and tricked her to let him in by pretending to work for the local church. He says that his plan was to distract the woman and quickly grab some money from the tin box, but she caught him in the act, and he panicked and stabbed her. At the police station AA pulls out a switch blade knife from his pocket and puts in on the table. The knife is smeared in blood, and is sent to forensics, who quickly confirm that the blood belongs to the victim. The autopsy report is consistent with AAs confession. The angle of the stab wound suggests that the perpetrator is above medium height, which is somewhat odd since AA is shorter than medium, but can be explained if AA held the knife high. AA is prosecuted for murder. At the trial, AAs defense attorney says that he suspects that his client is giving a false confession to cover for someone else. AA has no criminal record, but he has two sons who both have previous convictions for burglary and assault. Both sons are above medium height, and are known to carry switch blade knives. AA insists that he did it. He claims that his sons have nothing to do with the murder, and gives the court a detailed and vivid story of how he committed it, that fits perfectly with the crime scene. During the trial the court learns that the tin box was found open at the crime scene, but was never examined for fingerprints or DNA. Apparently, the police did not consider this necessary, since AA had confessed and the victim's blood had been found on his knife. When the defense attorney tried to have the tin box examined for fingerprints or DNA, it was too late. The box had been wiped clean from the victim's blood, which had removed all potential traces from the perpetrator. The defense attorney argues that the police investigators committed a huge blunder when they missed to look for forensic traces on the tin box, since the results of this investigation could have produced evidence favorable to the defendant. If fingerprints or DNA from one of AAs sons had been found on the box, AAs confession would have been falsified. AA is acquitted. The court says in its verdict that the police should have examined the tin box for fingerprints or DNA, and explains that the evidence against AA would have been sufficient for a conviction if the tin box had been properly examined and this had not produced any evidence against the prosecution's case, but since this examination is now missing from the investigation, the evidence against AA is not sufficiently robust for the standard of proof in criminal cases.

The Case of the Missing Fingers is inspired by a Swedish case where one of the authors of this paper did consulting work for the defense.[5] It goes as follows.

> In 2013 a beheading video is spread on the internet. The video is made with a smartphone in Syria and shows in graphic detail how a British journalist is decapitated by ISIS. The face of the executor is masked but his hands are visible and two fingers are partly missing on his right hand. A couple of months later the Swedish police receives an anonymous tip from a woman who has seen the beheading video on the internet and says that she recognizes the hand. She believes that the executor is BB, a man of Syrian origin living in Sweden. The police investigate BB and find that he made two trips from Sweden to Syria in 2013 to support ISIS in its cause. BB admits that he has participated in ISIS activities in Syria, but denies that he is the executioner in the beheading video, and claims that he has never killed anyone. A forensic image analyst compares the hand in the video with BBs hand, and report that they match. The missing fingers are severed in the same places. To assess the probability of a random match, the forensic analyst consults reference data on the prevalence of missing fingers. Searching a data base with 20 000 pictures of hands collected from the general Swedish population the forensic analyst finds 20 hands (1 in 1000) with severed fingers. At closer scrutiny, two of them (1 in 10 000) are severed in the same way as the hand in the beheading video, and match it just as well as BBs hand. The forensic analyst therefore concludes that the probability of a random match is approximately 1 in 10 000. The two matching hands in the reference data base belong to men who died before 2013 and can therefore be ruled out as suspects. BB is prosecuted for murdering the British journalist. The case for the prosecution is based on BBs affiliation with ISIS and the expert testimony of the forensic image analyst. BBs defense attorney argues that the random match probability assigned by the forensic expert is too small, since it is based on the prevalence of missing fingers in the general Swedish population and it is reasonable to assume that such injuries are more common among men that are affiliated with ISIS. In the cross-examination of the forensic expert, the defense attorney asks if it is possible that the gathering of further reference data about people affiliated with ISIS could have shown that missing fingers are considerably more frequent in this reference class, for example that 1 in 1000 rather than 1 in 10 000 are disfigured in this way. The forensic expert replies that this possibility cannot be ruled out. BB is acquitted. The court says in the verdict that the prosecution should have backed their case with better reference data. The court explains that a random match probability of 1 in 10 000 would have been sufficient for proof beyond reasonable doubt, given the other circumstances of the case, if this probability had been robust, but in the absence of more reference data on people affiliated with ISIS it is not sufficiently robust for the standard of proof in criminal cases.

The Case of the Tin Box and *The Case of the Missing Fingers* illustrate two different kinds of information deficiency. In *the Case of the Tin Box* the information gap concerns *case specific evidence* (fingerprints or DNA on the tin box). In *the Case of the Missing Fingers* there is an information deficiency with regard to the *sampling of reference data* (people affiliated with ISIS). The two cases also illustrate another difference. In *the Case of the Tin Box* it is too late to remedy the information gap. In *the Case of the Missing Fingers* it is practically possible to gather more information.

[5] Hovrätten för Västra Sverige, 2016-03-30, B 5306-15.

4. Information Economics and Shift-Ability

It is obvious that a requirement of sufficient informativeness cannot include every possible way of gathering potentially relevant information. Gathering information comes with a cost and there are obviously situations where the cost of a certain investigation clearly outweighs the potential benefit of obtaining the information it can produce. This means that a requirement of sufficient informativeness must include a cost-benefit calculus. Information economics must be integrated in the criminal standard of proof. In information economics, informativeness is optimized when all lines of inquiry where the expected benefit exceeds the investigation cost have been exhausted (Lawrence 1999, 28-31). For the purpose of this investigation, we will assume that the requirement of sufficient informativeness in the criminal standard of proof that we want to model requires that informativeness is optimized.[6]

Our tentative formulation of the two-fold standard of proof in Section 1 can now be re-formulated in terms of the general definition of optimized informativeness, as follows.

> There is proof beyond reasonable doubt if the probability of the *factum probandum* given the evidence is equal or greater than *p*, and there is no gathering of evidence that has not been undertaken where the expected benefit of an investigation is greater than the investigation cost.

Neil Cohen has proposed that the legal standard of proof be modeled as a twofold requirement, where 1) the fact-finder's assessment of the probability of the prosecutor's/plaintiff's hypothesis must reach a certain threshold, and 2) the fact-finder's must have a certain 'level of confidence' in this assessment (Cohen 1985, 399), meaning that it must fall within a certain confidence interval (Cohen 1985, 410). The problem with this proposal is that it uses a measure of informativeness for something that requires a measure of sufficient informativeness. It fails to take information economics into account. It is important to distinguish between models of informativeness and models of sufficient informativeness. A model of informativeness, e.g. a confidence interval or credible interval, is a purely epistemic measure, but a model of sufficient informativeness is a question of practical optimization (Nance 2016, 262), that must take the costs and benefits of information gathering into account. There is no such thing as sufficient informativeness disconnected from information economics.

Barbara Davidson and Robert Pargetter have proposed a twofold requirement for the criminal standard of proof, where 1) the fact-finder's assessment of the probability of the prosecutor's hypothesis reaches a certain threshold, and 2) this assessment is 'very reliable' (Davidson & Pargetter 1987, 182). Unlike Cohen's confidence-requirement, their reliability-requirement has a cost-benefit-approach to information (Davidson & Pargetter 1987, 187).

Information economics in legal evidence needs to consider a number of factors. Some investigations are more costly than others, but some investigations are also more likely than others to produce valuable information. Davidson and Pargetter observe that the question

[6] For an equivalent assumption, see Nance 2016, 135.

of weight is fundamentally about *shift-ability* (Davidson & Pargetter 1986, 224-225). The value of further inquiry lies in the possibility that additional information could lead to a switch in the legal decision. Davidson and Pargetter do not develop this further, but their observation is the key to a precise understanding of sufficient informativeness in the criminal standard or proof. In the following sections we will present a model of sufficient informativeness that revolves around shift-ability.

An investigation leads to a switch from conviction to acquittal when it produces new evidence that pushes the probability of the *factum probandum* under the threshold set by the standard of proof. It is more likely that an investigation will have this effect when the probability of the *factum probandum* given the current information barely surpasses the threshold, compared with a situation where the probability surpasses the threshold by a significant margin (Nance 2016, 137). The value of a certain investigation is therefore not constant but depends on its potential to produce a switch at the current state of information.

The two-fold standard of proof can now be re-formulated to explicitly spell out shift-ability, as follows.

> There is proof beyond reasonable doubt if the probability of the *factum probandum* given the evidence is equal or greater than *p*, and there is no gathering of evidence that has not been undertaken where the expected benefit of a potential shift in the probability of the *factum probandum* to a probability smaller than *p* is greater than the investigation cost.

This re-formulation is more exact than our previous versions, and we will refer to it as the *shift-ability formulation of the standard of proof*. We will now use decision theory to formalize it.

5. Applying Decision Theory to the Criminal Standard of Proof

Many legal scholars have used decision theory to model the requirement of sufficient probability of the *factum probandum* in the criminal standard of proof (Kaplan 1968; Tribe 1971; Lillquist 2002; Laudan & Sanders 2009), but decision theory has not been used to model the requirement of sufficient informativeness in the criminal standard of proof.

Decision theory has been used by forensic scientists to model optimal gathering of information in forensic decision making (cf Biedermann et al. 2020; Taroni et al. 2021). These models are related to the standard of proof since investigations carried out by forensic scientists derive their social utility from supplying information that ultimately goes into the legal fact-finder's decision whether the standard of proof is met (Gittelson et al. 2013, e43), but a model for forensic decision making is one thing and a model of the criminal standard of proof is another. The latter is an explication of a legal rule and must therefore model the requirements of that rule. In this and the following section we will use decision theory to model the two-fold standard of proof consisting of a requirement of sufficient probability of the *factum probandum* and a requirement that there is no gathering of evidence that has not been undertaken where the expected benefit of a potential shift in the probability of the *factum probandum* is greater than the investigation cost.

According to Nance, the two requirements in the standard of proof 'ordinarily do not interact' (Nance 2008, 273). This is not entirely true. Sufficient informativeness is always connected to sufficient probability of the *factum probandum* since shift-ability involves a hypothetical prediction about the sufficiency of the probability of the *factum probandum* given additional information. We will therefore start our decision analysis of the criminal standard of proof by modeling the requirement of sufficient probability of the *factum probandum*, and then move on to the requirement of sufficient informativeness.

A decision analysis takes account of the possible outcomes of a decision. With regard to the decision in a criminal trial to *convict* or *acquit*, there are four possible outcomes: 1) *correct conviction* (cc), 2) *incorrect conviction* (ic), 3) *incorrect acquittal* (ia) and 4) *correct acquittal* (ca), where 'correct' and 'incorrect' refers to whether the *factum probandum* is true or false ('correct' conviction means that the defendant actually did it, and so forth). The decision analysis must take account of the utility (U) of each outcome, which we will denote U_{cc}, U_{ic}, U_{ia} and U_{ca}. The exact values of these utilities are debatable, but most scholars agree that the correct outcomes have positive utility while the incorrect outcomes have negative utility.[7] The expected utility depends on the probability (p) of the *factum probandum*. Maximizing expected utility gives the following *decision rule* for when to convict (Tribe 1971, 1378-1381; Lillquist 2002, 109; Laudan & Saunders 2009, 3-6).

$$pU_{\mathrm{cc}} + (1-p)U_{\mathrm{ic}} > pU_{\mathrm{ia}} + (1-p)U_{\mathrm{ca}}$$

This gives the following threshold-probability for conviction.

$$p = \frac{1}{1 + \frac{U_{\mathrm{cc}} - U_{\mathrm{ia}}}{U_{\mathrm{ca}} - U_{\mathrm{ic}}}}$$

If, for example, U_{cc} = 5, U_{ic} = –500 U_{ia} = –5 and U_{ca} = 1, the threshold for conviction is 1/(1+10/501) ≈ 98%.

[7] It could be argued that a correct acquittal does not have a positive expected utility, but should be assigned zero utility, since it maintains status quo. The defendant is an innocent person who was viewed as innocent by the legal system before the decision due to the presumption of innocence and the acquittal only prolongs this status. On the other hand, it could be argued that a correct acquittal does have some positive utility, due to the double jeopardy clause, since it prevents the innocent person in question to be wrongfully prosecuted in the future.

6. Modeling Sufficient Informativeness

Information economics is always a question about further investigation, and it is always about the cost and benefit of a specific 'investigation' (**i**), a specific way of gathering more information. The investigation cost ($C_{\mathbf{i}}$) includes every disutility of conducting the investigation in question, and is not understood in a strictly pecuniary sense (de Finetti 2017, 499). The benefit is the expected utility ($\overline{U}_{\mathbf{i}}$) of the investigation. The difference between the expected utility and the investigation cost, $\overline{U}_{\mathbf{i}} - C_{\mathbf{i}}$ is the *net benefit* of the investigation. An investigation is economically *efficient* if the net benefit is positive (Kaplow 1994, 318).

$$\overline{U}_{\mathbf{i}} - C_{\mathbf{i}} > 0$$

In most cases there are several different ways to gather further information. There is a number of possible investigations, where each investigation has a different investigation-cost and a different expected utility. If more than one available investigation is economically efficient, expected utility is maximized by giving priority to the investigation with the highest net benefit (Kaplow 1994, 350; Nance 2016, 155; de Finetti 2017, 500). It should be noted that each investigation that is conducted can alter the net benefit of other investigations that have not yet been conducted, and thereby make previously efficient investigations in-efficient, or vice-versa. The cost-benefit analysis of various further investigations must therefore be up-dated after every investigative step. As an example, an investigation can rule out someone as the perpetrator, and therefore reduce the expected utility of further investigations regarding that person to zero. An example of the opposite effect occurs when an investigation leads to a new hypothesis about a possible perpetrator, and thereby increases the potential value of investigations that previously appeared irrelevant to the case (van Koppen & Mackor 2020, 1141).

An investigation reaches sufficient informativeness when all cost-efficient investigations have been exhausted, and all remaining options for further inquiry are in-efficient. Continued investigations beyond this point diminishes expected utility.[8] The requirement of sufficient informativeness can therefore be modeled as the non-existence of some further investigation that is cost-efficient:

$$\nexists \{\mathbf{i} : \overline{U}_{\mathbf{i}} - C_{\mathbf{i}} > 0\}$$

It should be mentioned that, since we are talking about a legal requirement, the legal system can make legal constraints on the kind of investigations that are obligatory for sufficient informativeness. Consider, for example, investigations that involve illegal interrogation

[8] It should be kept in mind that a probability assessment under optimized informativeness is not more accurate than probability assessments given other sets of evidence. Optimized informativeness only means that it is the right time to 'bet' on the probability assessment.

techniques or illegal invasions of privacy. If such an investigation is cost-efficient the requirement of sufficient informativeness dictates that it must be conducted, unless a constraint is added to this requirement saying that only legal investigations count.

Another possible constraint concerns investigations that the police could have done but failed to conduct and are no longer possible. Searching the tin box for fingerprints and DNA in *The Case of the Tin Box* serves as an example. Such situations raise an interesting question. Should the requirement of sufficient informativeness be restricted to investigations that are still possible at the trial, when the requirement is applied as a part of the criminal standard of proof? Or should it include all investigations that were available to the police and prosecution at one time or other? If the purpose of the requirement of sufficient informativeness is only to minimize the risk that investigations after a conviction produces new evidence that re-opens the case and exonerates the defendant, sufficient informativeness should be restricted to investigations that are still possible to conduct. If, on the other hand, the purpose is also to incentivize the police and prosecution to optimize informativeness by acquitting the defendant as a 'punishment' for sloppy police-work in cases where a cost-efficient investigation was overlooked (Nance 2011, 16), sufficient informativeness should include investigations that are no longer possible. The acquittal in *The Case of the Tin Box* is an example of the latter view.

As we have seen, sufficient informativeness depends on the expected utility of further investigations. So far we have treated expected utility as a black box. We will now move on and explore what factors go into $\overline{U}_{\mathbf{i}}$. Following the pioneering work of I.J. Good we will compare hypothetical 'states' (Good 1967, 319-320; Wendt 1969, 431; Lawrence 1999, 13; Buchak 2010, 97). In legal fact-finding, there are four possible states (S) that a decision in a criminal trial can lead to: 1) *correct conviction* (S_{cc}), 2) *incorrect conviction* (S_{ic}), 3) *incorrect acquittal* (S_{ia}) and 4) *correct acquittal* (S_{ca}). The potential utility of further investigations in a state where the probability assessment of the *factum probandum* (H) given the existing evidence (E_0) has reached the threshold for sufficient probability, $P(H|E_0) \geq p$, lies in the possibility that the information produced by further investigation will decrease the probability of the *factum probandum* under the threshold. The additional information could be case specific evidence (e.g. *The Case of the Tin Box*) but it could also be additional reference data (e.g. *The Case of the Missing Fingers*). For simplicity, we will denote the additional information $E_{\mathbf{i}}$ in either case. As we have seen, the potential utility of further investigations lies in a potential *switch* from conviction, $P(H|E_0) \geq p$, to acquittal, $P(H|E_0, E_{\mathbf{i}}) < p$ (Davidson & Pargetter 1985, 224-225).[9] We only need to consider a potential switch in one direction when we are dealing with the criminal standard proof. Since the burden of proof is on the prosecution, insufficient informativeness is a reason to acquit, but can never be a reason to convict. When the probability of the *factum probandum* is below the threshold, the prospect that additional information could produce evidence against the defendant, pushing the probability over the threshold, can never be a reason to

[9] A complete format for denoting a probability would include the conditioning on all information, including background information (I), e.g. $P(H|E_0, I)$. For simplicity, we omit background information from the formulas.

convict. The prosecution must produce the evidence. Since the burden of proof is on the prosecution, the potential effect of additional investigations can only benefit the defendant.

A switch from conviction to acquittal ($S_{c\to a}$) promotes utility if the prosecutor's hypothesis is false (defendant is innocent), since it substitutes an incorrect conviction (S_{ic}) with a correct acquittal (S_{ca}), but this is by no means a sure outcome. We must also consider the possibility that, although the additional investigation decreased the probability of the *factum probandum*, the hypothesis is actually true (defendant is guilty), and the switch from conviction to acquittal therefore substitutes a correct conviction (S_{cc}) with an incorrect acquittal (S_{ia}). In the same way a further investigation that does not lead to a switch from conviction to acquittal ($\sim S_{c\to a}$) can lead to a state where the fact-finder holds on to a correct conviction (S_{cc}), but we must also consider the possibility that, in spite of being based on more information, the decision is actually an incorrect conviction (S_{ic}). See figure 1.

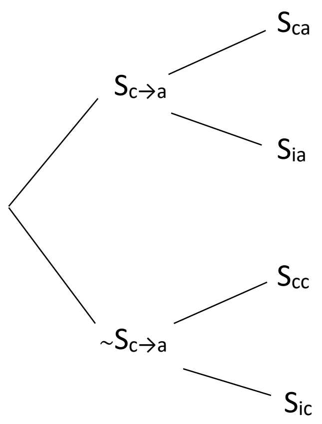

Figure 1. Possible outcomes of switch and non-switch from conviction to acquittal.

The expected utility of an investigation depends on the probability that it will lead to a switch from conviction to acquittal, $P(S_{c\to a}|\mathbf{i})$. If the probability of a switch is zero, the expected utility is also zero.[10] The expected utility if a certain investigation is conducted ($\overline{U}_{\mathbf{i}}$) can therefore be calculated as follows.

$$\overline{U}_{\mathbf{i}} = P(S_{c\to a}|\mathbf{i}) \times ((P(\sim H|E_0, E_{\mathbf{i}})U_{ca} + P(H|E_0, E_{\mathbf{i}})U_{ia}) - (P(H|E_0, E_{\mathbf{i}})U_{cc} + P(\sim H|E_0, E_{\mathbf{i}})U_{ic}))$$

The probability that an investigation will lead to a switch from conviction to acquittal, $P(S_{c\to a}|\mathbf{i})$, is the probability that the probability of the *factum probandum* given the evidence produced by the further investigation does not meet the required threshold. This is

[10] Nance has suggested that such situations could still have some positive utility, since more information gives the public a greater sense of confidence in the legal decision (Nance 2016, 264-267). In our view, it ought to decrease the public's confidence in the legal system to watch resources being wasted on investigations that have zero chance of changing a decision.

a probability on a probability, and it is therefore referred to in probability theory as a *second order probability* (or 'second order belief').

$$P(\mathrm{S}_{\mathrm{c}\to\mathrm{a}}|\mathbf{i}) = P(P(H|E_0, E_\mathbf{i}) < p|\mathbf{i})$$

The second order probability represents uncertainty about the potential evidence that the investigation will produce. So far we have simply represented the evidence produced by $\mathbf{i}$ with $E_\mathbf{i}$, but the outcome is actually a *set space* with a number of possible outcomes in terms of evidence, $\{E_\mathbf{i}^1, E_\mathbf{i}^2, \dots\}$. For simplicity, we assume a finite set of *n* outcomes, and denote this set of outcomes $\mathcal{E}_\mathbf{i}$, i.e., $\mathcal{E}_\mathbf{i} = \{E_\mathbf{i}^1, E_\mathbf{i}^2, \dots, E_\mathbf{i}^n\}$. These come with probabilities, $P(E_\mathbf{i}^k|\mathbf{i}), k = 1, \dots, n$, defined on a probability space different from the space on which the first order probabilities are defined.[11] In this second probability space, the first order probability $P(H|E_0, E_\mathbf{i})$ is a random variable that maps the elements of the set $\mathcal{E}_\mathbf{i}$ (one-to-one) into numerical values in the interval [0,1], but the number of unique values may be lower than *n*, since several of the first order probabilities can be equal in value. This in turn implies that $\overline{U}_\mathbf{i}$ is also a random variable, that varies with $E_\mathbf{i}$.

As an example, assume that the threshold for conviction is 98%, and the probability of the *factum probandum* given the current evidence, $P(H|E_0)$, is 99%, thus exceeding the threshold. Now, let us suppose that there are three possible outcomes of a further investigation, $E_\mathbf{i}^1, E_\mathbf{i}^2$ and $E_\mathbf{i}^3$, with the following posterior probabilities of the *factum probandum*: $P(H|E_0, E_\mathbf{i}^1)$ = 99%, $P(H|E_0, E_\mathbf{i}^2)$ = 95% and $P(H|E_0, E_\mathbf{i}^3)$ = 98%. Conditioning on the second potential outcome of a further investigation, the probability of the *factum probandum* does not reach the threshold. If this outcome had been part of the current evidence, the defendant would have been acquitted. Hence, the probability of a switch from conviction to acquittal if a further investigation is made is the probability of obtaining the second outcome from such an investigation, $P(E_\mathbf{i}^2|\mathbf{i})$. We can now define an indicator variable.

$$Q_\mathbf{i}(p) = \begin{cases} 1 & \text{if } P(H|E_0, E_\mathbf{i}) < p \\ 0 & \text{otherwise} \end{cases}$$

This means that for each possible outcome $E_\mathbf{i}^k$ from the further investigation there is a value $Q_\mathbf{i}^k(p)$ that is either 1 or 0. The second-order probability of a switch from conviction to acquittal can then be written as (1).

[11] To differ in notation between probabilities defined on different spaces one may use different symbols like P_1 and P_2, but to avoid less transparent formulas, we use the same symbol for probability but separate between the two spaces by conditioning on the further investigation $\mathbf{i}$ when referring to probabilities defined on the set of outcomes from that space.

$$P(\mathrm{S}_{\mathrm{c}\to\mathrm{a}}|\mathbf{i}) = P(Q_{\mathbf{i}}(p) = 1|\mathbf{i}) = \sum_{k=1}^{n} Q_{\mathbf{i}}^{k}(p)P\big(E_{\mathbf{i}}^{k}\big|\mathbf{i}\big) \tag{1}$$

And the expected utility becomes (2).

$$\begin{aligned}\overline{U}_{\mathbf{i}} = \sum_{k=1}^{n} Q_{\mathbf{i}}^{k}\left(\frac{1}{1+\frac{U_{\mathrm{cc}}-U_{\mathrm{ia}}}{U_{\mathrm{ca}}-U_{\mathrm{ic}}}}\right) &\times \Big(\big(P(\sim H\big|E_0, E_{\mathbf{i}}^{k})U_{\mathrm{ca}} + P\big(H\big|E_0, E_{\mathbf{i}}^{k}\big)U_{\mathrm{ia}}\big) \\ &- \big(P(H\big|E_0, E_{\mathbf{i}}^{k})U_{\mathrm{cc}} + P\big(\sim H\big|E_0, E_{\mathbf{i}}^{k}\big)U_{\mathrm{ic}}\big)\Big) \times P(E_{\mathbf{i}}^{k}|\mathbf{i})\end{aligned} \tag{2}$$

The requirement for sufficient informativeness, $\nexists\ \{\mathbf{i}\colon \overline{U}_{\mathbf{i}} - C_{\mathbf{i}} > 0\}$, can therefore be modeled as follows.

$$\begin{aligned}\nexists\left\{\mathbf{i}:\ \sum_{k=1}^{n} Q_{\mathbf{i}}^{k}\left(\frac{1}{1+\frac{U_{\mathrm{cc}}-U_{\mathrm{ia}}}{U_{\mathrm{ca}}-U_{\mathrm{ic}}}}\right)\right. &\times \Big(\big(P(\sim H\big|E_0, E_{\mathbf{i}}^{k})U_{\mathrm{ca}} + P\big(H\big|E_0, E_{\mathbf{i}}^{k}\big)U_{\mathrm{ia}}\big) \\ &- \big(P(H\big|E_0, E_{\mathbf{i}}^{k})U_{\mathrm{cc}} + P\big(\sim H\big|E_0, E_{\mathbf{i}}^{k}\big)U_{\mathrm{ic}}\big)\Big) \times P(E_{\mathbf{i}}^{k}|\mathbf{i}) - C_{\mathbf{i}} > 0\Big\}\end{aligned} \tag{3}$$

We will show in section 9 how this formalized decision rule can be used in a real case, by applying it to *The Case of the Missing Fingers*.

As we mentioned in the introduction, scholars who oppose the Bayesian approach to legal evidence have claimed that 'weight' (informativeness) cannot be modeled with a Bayesian approach (Brilmeyer 1986, 682-685; Cohen 1986, 648). Our model shows that this claim is mistaken. Some scholars claim that a requirement of sufficient informativeness does not follow the complementation rule for mathematical probability (Stein 2011, 241-42; Haack 2014, 61-62), since the probability that the *factum probandum* is correct and the probability that it is incorrect can both be based on insufficient information. Our model shows why this claim gets it wrong. There is no problem with the complementation rule when sufficient information is modeled as a question of information economics.

$$\textstyle\sum_{k=1}^{n} Q_{\mathbf{i}}^{k}\left(\frac{1}{1+\frac{U_{\mathrm{cc}}-U_{\mathrm{ia}}}{U_{\mathrm{ca}}-U_{\mathrm{ic}}}}\right)P(E_{\mathbf{i}}^{k}|\mathbf{i}) + \sum_{k=1}^{n}\left(1 - Q_{\mathbf{i}}^{k}\left(\frac{1}{1+\frac{U_{\mathrm{cc}}-U_{\mathrm{ia}}}{U_{\mathrm{ca}}-U_{\mathrm{ic}}}}\right)\right)P(E_{\mathbf{i}}^{k}|\mathbf{i}) = 1$$

7. The Debate over Higher Order Probabilities

As we have seen, the requirement of sufficient informativeness invokes the probability that the probability of the prosecutor's hypothesis given the evidence produced by the investigation does not meet the threshold for sufficient support, $P(P(H|E_0, E_{\mathbf{i}}) < p|\mathbf{i})$. This is a second order probability from the probability space spanned by the outcomes of a further investigation.

The second order probability is necessary to formalize shift-ability. In the shift-ability formulation of the standard of proof that we provided in section 4, the requirement of sufficient informativeness says: "…and there is no gathering of evidence that has not been undertaken where the expected benefit of a potential shift in the probability of the *factum probandum* to a probability smaller than *p* is greater than the investigation cost". A formalization of this rule will inevitably involve a second order probability.

It should be noted that it is possible to model sufficient informativeness without second order probabilities if sufficient informativeness is formulated in more general terms that do not specify that the potential benefit lies in shift-ability. We discussed such a general formulation in section 4 before we arrived at the shift-ability formulation. In the general formulation the requirement of sufficient informativeness says. "…and there is no gathering of evidence that has not been undertaken where the expected benefit of an investigation is greater than the investigation cost". This rule can be formalized with first order probabilities, since it does not specify that the expected benefit comes from a potential shift. Sufficient informativeness has been modeled in this way in the literature on forensic decision making (Gittelson et al 2013, e43-44), but a more precise exploration of the requirement of sufficient informativeness in the criminal standard of proof needs to take account of shift-ability, and therefore employ second order probabilities.

Higher order probabilities are controversial in the literature on probability theory. Some scholars have said that higher order probabilities should generally be avoided (de Finetti 1972, 145; Savage 1972, 58). Others say that we need higher order probabilities to solve certain problems (Reichenbach 1949, 324; Gärdenfors 1975, 171; Skyrms 1980, 118; Baron 1987, 27; Sahlin 1993, 13-23; Logue 1995, 87-95; Hansson 2008, 530-532). We will therefore take a closer look at this debate. The main objections to higher order probabilities can be summarized as follows: 1) *it violates the concept of probability to put a probability on a probability*, 2) *all uncertainty can be integrated in the first order probability*, and 3) *higher order probabilities lead to an infinite regress*. In this section, we will discuss each of these arguments.

The claim that *it violates the concept of probability to put a probability on a probability* is based on the argument that probabilities refer to events, but a probability is not in itself an event, and, therefore, there cannot be probabilities about probabilities. This argument was first made by Bruno de Finetti in the 1970s, and has recently been restated by Franco Taroni, Silvia Bozza, Alex Biedermann and Colin Aitken, who build on de Finetti's ground-breaking work on subjective probabilities.

> Any assertion concerning probabilities of events is merely the expression of somebody's opinion and not itself an event. There is no meaning, therefore, in asking whether such an assertion is true or false or more or less probable. (de Finetti 1972, 189)

> One can, in fact, have probabilities for events, or probabilities for propositions, but not probabilities of probabilities. (Taroni et al. 2016, 8)

The immediate rebuttal to this argument is that a probability assessment actually is an event. According to the subjective understanding of probability developed by de Finetti, to assess a probability is to assign a set of coherent preferences to a scheme of hypothetical decisions (de Finetti 2017, 70-73). In other words, a probability assessment is a *betting preference*.[12] This means that a probability assessment is not a proposition about an event, but is in itself an event. The formation of a betting preference by a certain person at a certain time is an event. And de Finetti admits this inadvertently when pointing out that probability assessment cannot be true or false: "the situation is different, of course, if we are concerned not with the assertion itself but with whether 'somebody holds or expresses such an opinion or acts according to it'; for this is a real event" (de Finetti 1972, 189). In our model of sufficient informativeness, the second order probability is the probability of the event that the fact-finder after further investigation will assess a probability to the *factum probandum* that falls below the threshold. This is a future event, but nevertheless an event.

The second objection to second order probabilities claims that higher order probabilities are unnecessary, since *all uncertainty can be integrated in the first order probability*. This argument has been made by the same group of scholars in a recent article with Biedermann as the first author.

> There is no need to assign a measure of uncertainty to the measure of uncertainty.
> (Biedermann et al. 2016, 394).

As we observed at the beginning of this section, a general requirement of sufficient informativeness that does not make shift-ability explicit can be modeled with first order probabilities, so if you are satisfied with this general notion of sufficient informativeness, Beidermann et al are right that there is no need to employ higher order probabilities. That approach makes sense if you are modeling forensic decision making, since the forensic scientist does not know what evidence will be presented at a future possible trial and therefore is not in a position to assess if further investigations have a potential to switch the fact-finder's decision at that future trial, but if you are modeling the fact-finder's decision making, and you want to model how the requirement of sufficient informativeness in the legal rule that the fact-finder is applying is conditioned on shift-ability, you need second order probabilities.

Uncertainty can only be integrated in the first order probability if the probability measure used is the same for all probability assessments made, and this is not the case when we are modeling shift-ability. A person assigning – at a specific time-point – probabilities to several events uses a probability measure defined on the probability space that is spanned by the total amount of relevant information available to that person (at that time-point). In the

[12] In the words of Frank Ramsey: 'a measurement of belief *qua* basis of action' (Ramsey 1926, 67).

literature, this is symbolized by using the notation $P(\cdot\,|I)$, where I stands for the relevant base of information (see note 11). However, integrating components of uncertainty measured by probabilities defined on different probability spaces is not correct, for instance it is not generally true that $P(H|I_1) + P(\sim H|I_2) = 1$ if I_1 and I_2 are not identical bases of information. In our derivation of expression (1) in section 6, using a stricter notation the first-order probabilities involved are $P(H|E_0) = P(H|E_0, I_0)$ and $P(H|E_0, E_\mathbf{i}) = P(H|E_0, I_0 \cup E_\mathbf{i})$, where I_0 is the base of information available at the time-point of assigning the conditional probability of H given only the evidence E_0. The conditioning on E_0 is made in the outcome space defined by the base of information available, while the conditioning on $E_\mathbf{i}$ is actually an *extension* of the base of information. For each possible outcome of a further investigation, this will change the assignment of $P(H|E_0)$, used as a generic notation of the first-order probability of interest, and it is this change that is in focus when we model shift-ability.

The objection that *higher order probabilities lead to an infinite regress* was also made by Bruno de Finetti in the 1970s,[13] and has recently been restated by Taroni, Biedermann, Bozza and Aitken (de Finetti 1972, 193; Taroni et al. 2016, 8; Biedermann et al. 2016, 394). It rests on the observation that if there is a second order probability regarding the possibility that additional information would alter our first order probability, there must also be a third order probability on the possibility that additional information would alter the second order probability, and so forth *ad infinitum*. If we can ask, in *The Case of the Missing Fingers*, how probable it is that additional reference data on people affiliated with ISIS would change our assessment of the random match probability that the hand of a random person affiliated with ISIS has severed fingers matching the defendant, we can also ask how probable it is that additional information would change our assessment of the probability that additional reference data would change our assessment of the random match probability, and so forth. The observation that second order probabilities open up this infinite regress is correct, but that does not mean that higher order probabilities should be avoided. The claim that second order probabilities should be avoided must demonstrate why the infinite regress is problematic.

The infinite regress of higher order uncertainty was first discovered by David Hume, who found it deeply problematic. According to Hume, the infinite regress diminishes the first order probability to zero.

> Having thus found in every probability, beside the original uncertainty inherent in the subject, a new uncertainty deriv'd from the weakness of that faculty, which judges, and having adjusted these two together, we are oblig'd by our reason to add a new doubt deriv'd from the possibility of error in the estimation we make of the truth and fidelity of our faculties. This is a doubt, which immediately occurs to us, and of which, if we wou'd closely pursue our reason, we cannot avoid giving a decision. But this decision, tho' it should be favourable to our preceding judgment, being founded only on probability, must weaken still further our first evidence, and must itself be weaken'd by a fourth doubt of the same kind, and so on in infinitum; till at last there remain nothing of the original probability, however great we may suppose it to have been, and however small the diminution by every new uncertainty. No

[13] See also Savage (1972, 257).

> finite object can subsist under a decrease repeated in infinitum; and even the vastest quantity, which can enter human imagination, must in this manner be reduc'd to nothing.
> (Hume 1739, Book I, Part IV, Section I).

Hume's logic is rarely flawed but this is an exception. Integrating higher order uncertainty does not *a priori* decrease the first order probability, but could just as well increase it. Taking account of higher order uncertainty constantly diminishes the probability that the first order probability will be unaffected by additional information, but that does not mean that it diminishes the first order probability (Atkinson & Peijnenburg 2020, 3314).

So, if the infinite regress of higher order uncertainty is problematic, the problem lies elsewhere. But where, exactly? There seems to be a notion that higher order probabilities are problematic because they make probability assessments infinitely complex and therefore make probability assessments practically impossible. This conclusion is exaggerated (Atkinson & Peijnenburg 2020, 3316), and a fundamental problem with this infinite complexity is that it hits first order probabilities just as much as higher order probabilities. The infinite complexity of higher order probabilities is exactly the same complexity that Biedermann and his co-authors must handle when they integrate all uncertainty in the first order probability. The claim that higher order uncertainty entails an infinite regress that is too complex to handle is therefore incompatible with the claim, made by the same scholars, that all uncertainty can, in theory and practice, be integrated in the first order probability. In summary, none of the arguments against higher order probabilities survive a closer scrutiny.

8. Informativeness in Forensic Reporting

As illustrated by *The Case of the Missing Fingers*, there are situations with sub-optimal informativeness with regard to sampling of reference data. In *The Case of the Missing Fingers* the expert lacked reference data on severed fingers in the relevant reference class (people affiliated with ISIS) and therefore suspected that the assessment of the random match probability and likelihood ratio could change if more reference data were sampled. The same kind of problem occurs with other kinds of forensic evidence when the reference data is poor (for example in a case where a shoeprint at the crime scene has a sole pattern that matches the suspect's shoe, but there is poor reference data on the prevalence of the sole pattern in the general population).

How should this kind of uncertainty be reflected in a forensic report to help the legal fact-finder decide if the requirement of sufficient informativeness is met? This is a controversial issue among forensic scientists, and has recently been the topic of a hot debate. Some forensic scientists recommend that uncertainty with regard to informativeness should be communicated by reporting the likelihood ratio in the form of a *credible interval* (Morrison 2011, 95; Sjerps et al. 2016, 25). Others strongly reject this proposal and argue that the expert shall assign a point estimate[14] to the likelihood of a random match, and report the

[14] We use the term 'point estimate' following established terminology in the field. We are aware that some proponents of the described method (Taroni, Biedermann, Bozza and Aitken) have claimed that the word 'estimate' for the assessment of a subjective probability is an abuse of language, since 'estimate' implies the

likelihood ratio as a single value (Taroni et al. 2016, 2). In this paper we will refer to the latter method as *single value reporting*. Advocates of single value reporting include Franco Taroni, Silvia Bozza, Alex Biedermann, Colin Aitken, Charles Berger, Duncan Taylor, Tacha Hicks and Christophe Champod. Critics of the single value reporting include Marjan Sjerps, Kristy Martire, Gary Edmond and Geoffrey Morrison.

The Case of the Missing Fingers is an example of single value reporting. The expert assesses the random match probability at “1 in 10 000”. If the expert had instead reported the estimate in the form of an interval, the expert could, for example, have placed the random match probability in an interval between 1 in 5 000 and 1 in 20 000, and addressed the question of additional reference data by, for example, reporting it as a 95% credible interval.

The arguments for single value reporting can be summarized as follows. The number assigned by the forensic scientist to the likelihood of the random match hypothesis is not an estimate of an actual frequency[15] (Taroni et al. 2016, 8; Biedermann et al. 2017, 81). It is a measure of the expert’s degree of belief, and must therefore be a single number, since it would be illogical for the expert to assign different values to the same belief (Taroni et al 2016, 6). The number reflects all available information and therefore encapsulates all existing uncertainty with regard to the current legal case (Berger & Slooten 2016, 389; Biedermann et al. 2016, 395; Taroni et al. 2016, 9-10; Taylor et al. 2016, 403). According to advocates of single number reporting, the question whether the reference data is optimal with regard to informativeness is relevant when the forensic expert looks ahead at future cases and contemplates if more reference data should be collected to improve the basis for future reports, but it is not relevant for the likelihood assigned in the current legal case (Berger & Slooten 2016, 389; Taroni et al. 2016, 12-13). The only way in which uncertainty with regard to informativeness should be allowed to influence the case at hand is if the expert feels that the available reference data is so deficient that prudence requires the expert to abstain from reporting an opinion on the likelihood of the random match hypothesis (Taylor et al. 2016, 407-408), but given that the forensic expert does report such an opinion, it should be in the form of a single number.

According to the advocates of single value reporting there is no need to expand the point to an interval to handle uncertainty regarding informativeness since all uncertainty is already encapsulated in the point estimate (Biedermann et al. 2016, 394). All uncertainty can and should be integrated in the first order probability (see section 7, above). Advocates of single value reporting also stress that a single number serves well as a factor in Bayesian updating, but intervals do not (Berger & Slooten 2016, 391; Taroni et al. 2016, 11). For clarification, it

measure of something objective in the external world (Taroni et al. 2016, 2). We see the point, but believe that their conclusion goes too far. The word ‘estimate’ does not necessarily relate to something that exists outside the human mind. It can relate to some belief in the speaker’s mind, and refer to the mental process in which the speaker is trying to come to terms with how strongly he or she holds the belief in question. This is what Bayesians mean when they, occasionally, use the word ‘estimate’ about assessments of subjective probability, and it is, therefore, not an ‘abuse of language’ as Taroni and his co-authors have argued. As a matter of fact, Taroni and his co-authors have themselves used the term ‘estimate’ about assessments of subjective probabilities in previous publications (see, for example, Taroni et al. 2010, 60).

[15] In *The Case of the Missing Fingers* the frequency of severed fingers among people affiliated with ISIS.

should be stressed that the single value reporting does not mean that a forensic report should only consist of the single number assigned to the likelihood of the random match hypothesis and/or the single number assigned to the likelihood ratio. The report should also explain how the forensic expert arrived at the number, describing the methodology and the reference data used (Biedermann et al. 2016, 394-395; Slooten & Berger 2017, 468).

Critics of single value reporting oppose the idea that a degree of belief must be a point estimate. They claim that reporting a point estimate is not a conceptual necessity but a choice (Sjerps et al. 2016, 25), and propose the use of intervals instead, since an interval communicates more information relevant to the current legal case than a single number (Morrison & Enzinger 2016, 375; Sjerps et al. 2016, 26; Martire et al. 2017, 78).

We are not siding with either camp in this debate. We partially agree and partially disagree with both sides. We agree with the advocates of single value reporting that the first order probability of a random match and the resulting likelihood ratio should be reported as a single number. A degree of belief is susceptible to a Dutch Book if it is represented as an interval.[16] But we agree with critics of single value reporting that the forensic expert should report more than this. The report should also help the legal fact-finder to assess the informativeness.[17]

Advocates of single value reporting seem to think that information deficiency plays no role in the legal decision, since the law forbids a *non liquet*-verdict and thereby constrains the legal fact-finder to 'bet' on the evidence presented in court.

> … sensitivity analysis is meant to characterize the system that generates the LRs, and does not characterize the *evidence* in a particular case. Whether or not to gather more data to narrow down the input parameters further is a different and separate issue from the issue defined by the proposition in the case. (Berger & Slooten 2016, 389).

This is a misunderstanding of procedural law. As we have seen, the criminal standard of proof can include a requirement of sufficient informativeness, and the defendant can therefore be acquitted on the grounds of information deficiency. In such acquittals, the prosecution is being punished for bringing an insufficiently investigated case to court, and the fact-finder never makes a 'bet' on the first order probability of the *factum probandum*.

We agree with critics of single value reporting that a forensic report should address the question of informativeness, but we do not agree that a credible interval (or a posterior distribution) does this in a satisfying way. As we have seen, sufficient informativeness is not a purely epistemic matter, but a question of information economics, and is therefore not

[16] Suppose, for example, that a punter's degree of belief (p) in some future event is represented by the interval 1/200 ≤ p ≤ 1/50 and the punter is presented with a Dutch Book consisting of a €2 bet on the event happening at 75-to-1 odds and a €149 bet that the event will not happen at 150-to-149 odds. Given the punter's degree of belief, the expected monetary value (EMV) of the first bet is positive in the upper end of the interval (1/50 × €150 - €2 = €1) and the EMV of the second bet is positive in the lower end of the interval (199/200 × €150 - €149 = €0.25). Accepting both bets is, however, a sure loss for the punter (paying €2 + €149 = €151 to win €150 whether the event happens or not), turning the Dutch Book into a money pump for the bookie.

[17] *ENFSI Guidelines for Evaluative Reporting in Forensic Science* 2016, 16. https://enfsi.eu/wp-content/uploads/2016/09/m1_guideline.pdf

captured by credible intervals that do not take the utilities of further investigations (costs and benefits) into account. Sufficient informativeness is not an ‘acceptable’ credible interval.

We propose that the forensic report should help the fact-finder decide if the requirement of sufficient informativeness has been met. As we have seen (section 6), this involves an assessment of various further investigations with regard to their cost and their potential of producing new evidence that could decrease the probability of the *factum probandum* below the threshold and thereby lead to a switch from conviction to acquittal. The forensic expert can help the fact-finder with two of these factors. When the investigation in question is forensic (searching for more traces, running additional lab tests, gathering more reference data) the expertise of the forensic scientist typically includes knowledge about the cost of the investigation and the kind of result it has the potential to produce. This information should be communicated to the fact-finder. The forensic expert is not in a position to assess the probability that a certain further investigation will push the probability of the *factum probandum* below the threshold, since the expert does not know all the evidence in the case, but the expert can nevertheless help the fact-finder with the cost-benefit-analysis involved in this assessment. Experts can use their experience in forensic science to say what results a certain investigation could produce, and also, to some extent, how likely certain results are. An expert on DNA could, for example, help the fact-finder to assess how likely it is that an object left in the woods will still have usable DNA-traces after some time has passed.

We therefore propose that a forensic report should help the fact-finder assess sufficient probability of the *factum probandum* as well as sufficient informativeness. With regard to sufficient probability of the *factum probandum*, it should report the random match probability and likelihood ratio for the forensic evidence in question in the form of a single number. With regard to sufficient informativeness, the report should survey further investigations and report 1) their cost, and 2) what results they have the potential to produce. When assessing how strong a hypothetical result could be as counter-evidence, the expert could, as suggested by Ronald Meester and Klaas Slooten, report the assessment as a posterior likelihood ratio distribution (Meester & Slooten 2021, 73-75).

9. Estimating the Economics of Additional Investigations

In *The Case of the Missing Fingers* the court found that the evidence was insufficiently robust, due to poor reference data about people affiliated with ISIS, and acquitted the defendant. Was this decision correct according to the model of sufficient informativeness that we have developed? As we have seen, this depends on the costs of gathering more reference data and the potential benefit of the additional information. The decision to acquit is only correct if the expected benefits indeed outweigh the costs.

To make an estimate on the economic efficiency of gathering reference data about severed fingers among people affiliated with ISIS we need to make some assumptions about the utilities involved. Following the example in section 5, let us assume that the expected social utility for correct conviction, incorrect conviction, correct acquittal and incorrect acquittal is

U_{cc} = 5, U_{ic} = –500, U_{ca} = 1 and U_{ia} = –5. As we observed in section 5, this sets the threshold for sufficient probability of the *factum probandum* at 98%. Since the court In *the Case of the Missing Fingers* explains that a likelihood ratio of 10 000 would have been sufficient to reach the threshold, the prior probability is ≥ 1/200 = 0.5% given the assumed threshold of 98%.[18] With regard to information costs, let us assume that it is possible to gather a sample of 1000 images of hands belonging to people affiliated with ISIS at a cost of $C_{\mathbf{i}}$ = 80. This information can be provided by a forensic expert, and the expert can also calculate the prospective likelihood ratios, given different outcomes of such a sample. Since we have no information about people affiliated with ISIS before we receive the sample, some assumptions must be made about the proportion of severed hands in this population. For a further investigation to be cost-efficient, hands that match the hand in the video must be more frequent in the ISIS-population than the general Swedish population (1/10 000). Our estimate of the potential benefit of further reference data should therefore be focused on proportions >1/10 000. It is reasonable to think that the proportion of hands with missing fingers that match the hand in the video could be 10 times higher in the ISIS-population than the general Swedish population, but it seems implausible that it is more than 100 times higher. It is, therefore, reasonable to assume that the relative frequency of matching hands among men affiliated with ISIS could be 1/1000, but not higher than 1/100. On the basis of this assumption, we make a Bayesian inference from a sample, using a uniform prior distribution for the proportion ranging from 1/10 000 to 1/100, updated with the sample outcome to a posterior distribution. The sample outcome is modeled using a binomial distribution. Since there is no other information relevant for the assignment of the probability of the observed match, the best single-value assignment of that probability is a measure of the central tendency in the posterior distribution of a proportion.[19] For a skewed distribution, the posterior median is preferable. [20] Letting $M(k)$ denote the posterior median when there are k observed matches in a sample, the likelihood ratio assigned is $1/M(k)$.

In Figure 2, ranges of values covering 99% of the prospective likelihood ratio distribution[21] are plotted for true proportions in the range from 1/10 000 to 1/100. This information, accompanied by an estimate of the costs of collecting the sample and conducting the study, can be provided by a forensic expert to assist the fact-finder in the assessment of sufficient informativeness.

[18] 0.005/0.995 × 10 000 ≈ 0.98/0.02 (Bayes' theorem).

[19] This way of assigning a probability integrates the uncertainty, thus taking into account limited amount of reference data (see e.g. Taroni et al, 2016. 9-10).

[20] Denoting the true proportion π, with a lower limit *c* (here *c* = 1/10 000) and an upper limit d (here $d = 1/100$) and sample size n, the chosen uniform prior density is $p(\pi|c,d) = \frac{1}{d-c}, c < \pi \leq d$ and zero otherwise, the binomial likelihood from k 'successes' out of n trials is $f(k|\pi,c,d) = \binom{n}{k}\pi^k(1-\pi)^{n-k}$, and the posterior density $q(\pi|k,n,c,d)$ is proportional to $\pi^k(1-\pi)^{n-k}, c < \pi \leq d$, and zero otherwise. The posterior median is then $F^{-1}\left(0.5 \cdot \left(F(d) - F(c)\right)\right)$, where F is the cumulative distribution function of a Beta distribution with shape parameters $x+1$ and $n-x+1$ respectively, and F^{-1} is the inverse of F.

[21] A prospective likelihood ratio distribution is the set of likelihood ratios that can be obtained given a future sample from the population of interest, together with the probabilities of the outcomes of such a sample. It is related to the *posterior likelihood ratio distributions* described by Meester & Slooten (2021, 73-75).

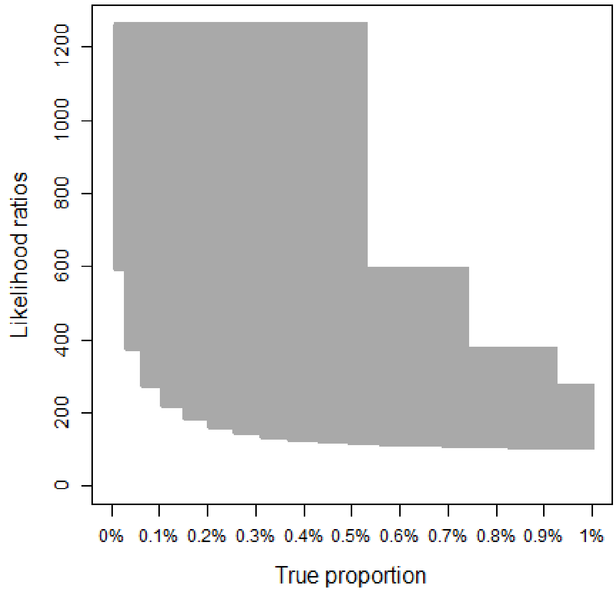

Figure 2. Ranges covering 99% the prospective likelihood ratio distribution than can be obtained for different values of the true proportion

Since we confine the analysis to true proportions in the interval [1/10 000, 1/100], we can investigate how the expected prospective utility ($\overline{U}_{\mathbf{i}}$) depends on the true proportion. $\overline{U}_{\mathbf{i}}$ will of course increase with the proportion, but the crucial question is when it exceeds the cost of making the investigation. Using the posterior median as the assigned probability of the match, the probability (ex post) of H given that there are k persons in the sample with a hand that matches the defendant's hand is $P(H|k) = 1 \cdot P(H)/\big(1 \cdot P(H) + M(k) \cdot P(\sim H)\big)$, where $M(k)$ is the posterior median. Using expression (2) (and (3)) above, $E_{\mathbf{i}}^{k}$ is simply replaced by k, $P\big(H\big|E_0, E_{\mathbf{i}}^{k}\big)$ is replaced by $P(H|k)$ and $P\big(\sim H\big|E_0, E_{\mathbf{i}}^{k}\big)$ by $1 - P(H|k)$. In Figure 3, $\overline{U}_{\mathbf{i}}$ is graphed against the true proportion, with $P(H)$ = 0.5%.

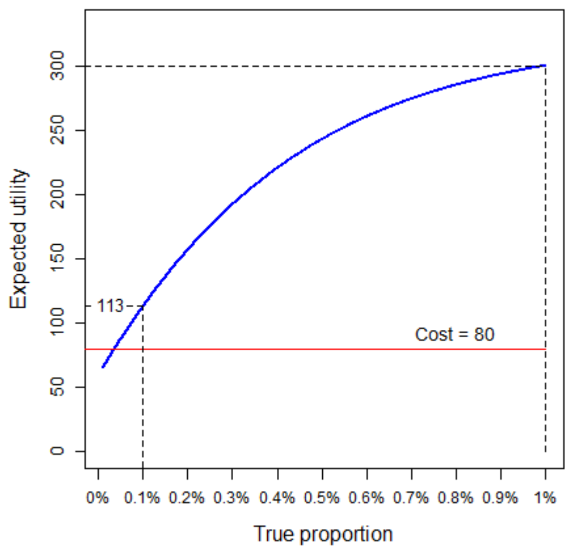

Figure 3. Expected utility of additional reference data graphed against a range of potential values for the true proportion of matching hands. The cost is shown with a straight horizontal line and dashed lines show the expected utilities when the proportion is 0.1% and 1 % respectively.

As Figure 3 shows, the expected utility of additional information ($\overline{U}_{\mathrm{i}}$ = 113) is higher than the information cost (C_{i} = 80) if the proportion of hands that match the hand in the video is ten times higher in the ISIS-population (1/1000) than the general population (1/10 000), and the net benefit increases if the true proportion is more than ten times higher. The cost-benefit relation if the true proportion is less than ten times higher is highlighted in Figure 4.

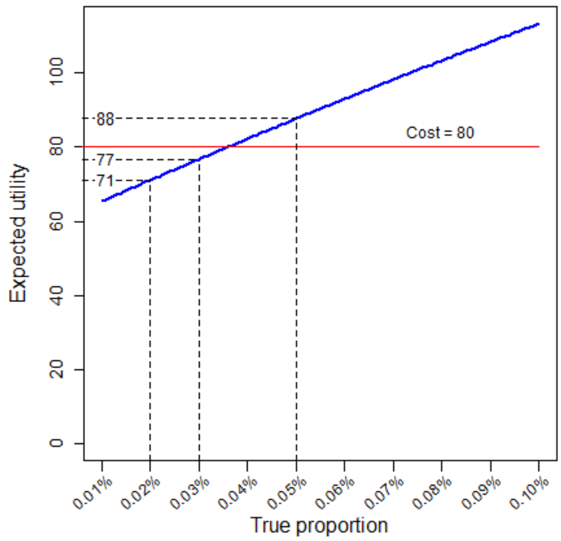

Figure 4. Expected utility of additional reference data graphed against potential proportions of matching hands ranging from 0.01% to 0.1%. The anticipated cost is shown with a straight horizontal line and dashed lines show the expected utilities for the proportions 0.02%, 0.03% and 0.05%.

The expected utility of additional information falls below the information cost if the true proportion of matching hands is only two or three times higher in the ISIS-population, but the expected utility exceeds the information costs if the true proportion is four times higher, and is clearly above the cost if it is five times higher. In other words, if we suspect that the true proportion could be at least 4-5 times higher, and therefore estimate the potential benefit of gathering additional reference data to exceed its cost, the court was right to acquit the defendant in *The Case of Missing Fingers.*

10. Conclusions

In this paper, we have modeled the criminal standard of proof as a twofold standard requiring sufficient probability of the *factum probandum* and sufficient informativeness. The focus of the paper is on the latter requirement, and we have used decision theory to develop a model for sufficient informativeness.

We have demonstrated that sufficient informativeness is fundamentally a question of information economics and switch-ability. In our model, sufficient informativeness is a cost-benefit-analysis of further investigations that involves a prediction of the possibility that such investigations will produce evidence that switches the decision from conviction to acquittal.

Critics of the Bayesian approach to legal evidence have claimed that 'weight' and sufficient informativeness cannot be captured in a Bayesian model. Our model shows that this is wrong. As we have demonstrated, sufficient informativeness can be modelled as a second order probability.

Advocates of the Bayesian approach have so far been reluctant to use second order probabilities, claiming that higher order probabilities are problematic in several respects. We have shown that this general aversion towards higher order probabilities is unfounded. None of the arguments against higher order probabilities survive a closer scrutiny.

We have also proposed that forensic reports help the legal fact-finder to assess sufficient informativeness. Forensic reports should survey further investigations, with regard to their cost and the results they have the potential to produce.

References

Atkinson, D. & Peijnenburg, J. "Till at Last there Remain Nothing – Hume's Treatise 1.4.1 in Contemporary Perspective", *Synthese*, Vol. 197, 2020.

Baron, J. "Second-Order Probabilities and Belief Functions", *Theory and Decision*, Vol. 23, 1987.

Berger, C. & Slooten, K. "The LR does not Exist", *Science and Justice*, 2016.

Biedermann, A., Bozza, S., Taroni, F. & Aitken, C. "Reframing the Debate – a Question of Probability, not of Likelihood Ratio", *Science and Justice*, 2016.

Biedermann, A., Bozza, S., Taroni, F. & Aitken, C., "The Consequences of Understanding Expert Probability Reporting as a Decision", *Science and Justice*, 2017.

Biedermann, A., Bozza, S. & Taroni, "Normative Decision Analysis in Forensic Science", *Artificial Intelligence and Law*, 2020.

Brilmeyer, L. "Second-Order Evidence and Bayesian Logic", *Boston University Law Review*, vol. 66, 1986.

Buchak, L. "Instrumental Rationality, Epistemic Rationality and Evidence-Gathering", *Philosophical Perspectives*, 2010.

Cohen, L.J, *The Probable and the Provable*, Oxford University Press, Oxford, 1977.

Cohen, L.J. "The Role of Evidential Weight in Criminal Proof", *Boston University Law Review*, Vol. 66, 1986.

Cohen, N. "Confidence in Probability – Burdens of Persuasion in a World of Imperfect Knowledge", *New York University Law Review*, Vol. 60, 1985.

Cullison, A. “Probability Analysis of Judicial Fact-Finding – A Preliminary Outline of the Subjective Approach”, *University of Toledo Law Review*, vol. 1, 1969.

Dahlman, C., Sarwar, F. & Wahlberg, L. “Robust Trust in Expert Testimony”, Humana Mente, 2015.

Davidson, B. & Pargetter, R. “Weight”, *Philosophical Studies*, Vol. 49, 1986.

Davidson, B. & Pargetter, R. “Guilt Beyond Reasonable Doubt”, *Australasian Journal of Philosophy*, Vol. 65, 1987.

Ferrer Beltrán, J. “Prolegómenos para una teoría sobre los estándares de prueba – El test case de la responsabilidad del Estado por prisión preventiva errónea”, *Filosofía del derecho privado*, 2018.

de Finetti, B. *Probability, Induction and Statistics*, Wiley, New York, 1972.

de Finetti, B. *Theory of Probability*, Wiley, New York, 2017.

Finkelstein, M. & Fairley, W. “A Bayesian Approach to Identification Evidence”, *Harvard Law Review*, Vol. 83, 1970.

Friedman, R. “Answering the Bayesioskeptical Challenge”, International Journal of Evidence and Proof, vol. 1, 1997.

Gärdenfors, P. “Qualitative Probability as an Intensional Logic”, *Journal of Philosophical Logic*, Vol. 4, 1975.

Gärdenfors, P. & Sahlin, N. ”Unreliable Probabilities, Risk Taking and Decision Making”, *Synthese*, Vol. 53, 1982.

Good, I.J. “On the Principle of Total Evidence”, *British Journal for the Philosophy of Science*, vol. 17, 1967.

Haack, S. *Evidence and Inquiry*, Blackwell, London, 1993.

Haack, S. *Evidence Matters – Science, Proof and Truth in the Law*, Cambridge University Press, 2014.

Hamer, D. “Probability, Anti-Resilience and the Weight of Expectation”, *Law, Probability and Risk*, vol. 11, 2012.

Hansson, S.O. “Do we need Second-Order Probabilities?”, *Dialectica*, Vol. 62, 2008.

Hedden, B. & Colyvan, M. “Legal Probabilism – A Qualified Defense”, *Journal of Political Philosophy*, vol. 27, 2019.

Ho, H.L. A Philosophy of Evidence Law, Oxford University Press, Oxford, 2008.

Howard, R. “Uncertainty about Probability – A Decision Analysis Perspective”, *Risk Analysis*, Vol. 8, 1988.

Hume, D. *A Treatise of Human Nature*, John Noon, London, 1739.

Kaplan, D. “Decision Theory and the Fact-Finding Process”, *Stanford Law Review*, Vol. 20, 1968.

Kaplow, L. “The Value of Accuracy in Adjudication – An Economic Analysis”, *Journal of Legal Studies*, 1994.

Kaye, D. “Limits of the Preponderance of the Evidence Standard”, *American Bar Foundation Research Journal*, Vol. 7, 1982.

Kaye, D. “Do we need a Calculus of Weight to Understand Proof beyond a Reasonable Doubt?”, *Boston University Law Review*, Vol. 66, 1986.

Keynes, J.M. *A Treatise on Probability*, Macmillan, London, 1921.

van Koppen, P. & Mackor, A.R. ”A Scenario Approach to the Simonshaven Case”, *Topics in Cognitive Science*, vol. 12, 2020.

Laudan, L. & Saunders, H. “Re-Thinking the Criminal Standard of Proof – Seeking Consensus about Utilities of Trial Outcomes”, *International Commentary on Evidence*, Vol. 7, 2009.

Lawrence, D. *The Economic Value of Information*, Springer, New York, 1999.

Lillquist, E. “Recasting Reasonable Doubt – Decision Theory and the Virtues of Variability”, *University of California Davis Law Review*, Vol. 36, 2002.

Logue, J. *Projective Probability*, Clarendon, Oxford, 1995.

Martire, K., Edmonde, G., Navarro, D. & Newell, B. “On the Likelihood of ‘Encapsulating all Uncertainty’”, *Science and Justice*, 2017.

Meester, R. & Slooten, K. *Probability and Forensic Evidence*, Cambridge University Press, 2021.

Morrison, G.S. “Measuring the Validity and Reliability of Forensic Likelihood-Ratio Systems”, *Science and Justice*, vol. 51, 2011.

Morrison, G.S. & Enzinger, E. “What should a Forensic Practitioner’s Likelihood Ratio be?”, *Science and Justice*, 2016.

Nance, D. “Evidential Completeness and the Burden of Proof”, *Hastings Law Journal*, vol. 49, 1998.

Nance, D. “The Weights of Evidence”, *Episteme*, 2008.

Nance, D. *The Burdens of Proof – Discriminatory Power, Weight of Evidence and Tenacity of Belief*, Cambridge University Press, Cambridge, 2016.

Nance, D. “Weight of Evidence”, in Dahlman, C. et al (ed.) *The Philosophical Foundations of Evidence Law*, Oxford University Press, Oxford, 2021.

Ramsey, F.P. “Truth and Probability”, 1926 (published in Ramsey, F.P. *Philosophical Papers*, edited by D.H. Mellor, Cambridge University Press, Cambridge, 1990).

Reichenbach, H. *The Theory of Probability*, 2nd ed., University of California Press, Berkeley, 1949.

Sahlin, N. "On Second-Order Probabilities and the Notion of Epistemic Risk", in Stigum, B.P. et al (ed.) *Foundations of Utility and Risk Theory with Applications*, Springer, Dordrecht, 1983.

Sahlin, N. "On Higher-Order Beliefs", in Dubucs, J. (ed.) *Philosophy of Probability*, Springer, Dordrecht, 1993.

Savage, L. *The Foundations of Statistics*, 2nd ed., Dover, New York, 1972. Schauer, F. & Zeckhauser, R. "On the Degree of Confidence for Adverse Decisions", *Journal of Legal Studies*, Vol. 25, 1996.

Sjerps, M., Alberink, A., Bolck, A., Stoel, R., Vergeer, P. & van Zanten, H. "Uncertainty and LR – to Integrate or not to Integrate, that is the Question", *Law, Probability and Risk*, 2016.

Skyrms, B. "Higher Order Degrees of Belief", in Mellor, D.H. (ed.) *Prospects for Pragmatism – Essays in Memory of F.P. Ramsey*, Cambridge University Press, Cambridge, 1980.

Slooten, K. & Berger, C. "Response Paper to 'The Likelihood of Encapsulating all Uncertainty' – the Relevance of Additional Information for the LR", *Science and Justice*, 2017.

Stein, A. *Foundations of Evidence Law*, Oxford University Press, Oxford, 2005.

Stein, A. "The Flawed Probabilistic Foundations of Law and Economics", *Northwestern University Law Review*, 2011.

Taroni, F., Bozza, S., Biedermann, A., Garbolino, P., & Aitken, C., *Data analysis in forensic science – A Bayesian decision perspective*, Wiley, New York, 2010.

Taroni, F., Bozza, S., Biedermann, A. & Aitken, C. "Dismissal of the Illusion of Uncertainty in the Assessment of a Likelihood Ratio", *Law, Probability and Risk*, 2016.

Taroni, F., Biedermann, A. & Bozza, S. "The Logic of Inference and Decision for Forensic Evidence", in Dahlman, C., Stein, A. and Tuzet, G. (ed.) *Philosophical Foundations of Evidence Law*, Oxford University Press, Oxford, 2021.

Taylor, D., Hicks, T. & Champod, C., "Using Sensitivity Analyses in Bayesian Networks to Highlight the Impact of Data Paucity and Direct Future Analyses – a Contribution to the Debate on Measuring and Reporting the Precision of Likelihood Ratios", *Science and Justice*, 2016.

Tribe, L. "Trial by Mathematics – Precision and Ritual in the Legal Process", *Harvard Law Review*, Vol. 84, 1971.

Wendt, D. "Value of Information for Decisions", *Journal of Mathematical Psychology*, Vol. 6, 1969.