

Towards a Theory of the Weight of Evidence: Rethinking Trial Decisions

Abstract It is not easy to decide when someone accused of a crime should be convicted. Since evidence is fallible, trial decisions are subject to uncertainty. The standard model of decision-making in economics and psychology recommends assessing the balance of the evidence for and against the accused and then taking the decision with the greatest expected utility. What this model overlooks is that even if the evidence, on balance, tips strongly against the accused, it may lack weight: the evidence may be incomplete or one-sided. This project argues against the standard model by using insights from philosophy, law and decision theory. By examining issues that are conceptual, descriptive and normative, this Publication Grant will demonstrate why taking into account the *weight* of the evidence, not only its balance, is an integral part of rational decision-making. This will lay the groundwork for giving guidelines to judges and jurors on how to improve the accuracy and fairness of their decisions.

1 TABLE OF CONTENTS

| | | |
|-----------|--|-----------|
| 2 | Statement of Significance and Impact | 2 |
| 3 | List of Participants | 3 |
| 4 | Narrative | 4 |
| | A. SUBSTANCE AND CONTEXT | 4 |
| | Beyond the Standard Model: Balance and Weight | 4 |
| | Weight: Completeness and Robustness | 5 |
| | Major Issues to Be Addressed | 5 |
| | Significance for the Humanities and Beyond | 6 |
| | B. HISTORY OF THE PROJECT | 7 |
| | C. COLLABORATORS | 7 |
| | Advisory Board | 9 |
| | D. METHODS AND EXECUTION | 9 |
| | Tasks and Division of Labor | 10 |
| | Output | 11 |
| | E. WORK PLAN | 12 |
| | F. FINAL PRODUCT AND DISSEMINATION | 13 |
| 5 | Project Budget | 14 |
| 6 | Budget Justification | 18 |
| 7 | Appendices | 19 |
| | Résumés | 19 |
| | Bibliography | 35 |
| 8 | Statement of Funding Received and Requested | 42 |
| 9 | Federally Negotiated Indirect-cost Rate Agreement | 43 |
| 10 | Certification Regarding Lobbying | 60 |

2 STATEMENT OF SIGNIFICANCE AND IMPACT

When someone accused of a crime is brought to trial, judges and jurors are tasked with deciding whether the accused should be convicted or acquitted. This is a difficult task. Since evidence is fallible, trial decisions are subject to uncertainty. How should judges and jurors decide in the face of uncertainty? The standard model of decision-making in economics and psychology recommends assessing the available evidence for and against the accused to estimate the likelihood of guilt, and then taking the decision with the greatest expected utility, all things considered. Many empirical studies, however, have shown that people's patterns of decision deviate from this model in significant ways. These deviations are often interpreted as an indication that humans make bad decisions or behave irrationally. But such deviations might also suggest that the model is inadequate, not that people are irrational.

By drawing from research in philosophy of law, epistemology, economics and decision theory, this Publication Grant project will defend a richer account of rational decision-making than the standard model. The central claim of the project is that decisions should not only be guided by the balance of the available evidence, but also by what we will call the *weight* of the evidence. Even if the available evidence, on balance, tips strongly against the accused, it may be deficient and lack weight for a number of reasons. For example, the available evidence may be one-sided and partial, or relevant information may be missing. Except for a few suggestions in the literature, however, there is no well-developed theory of the weight of evidence to date. 'The Weight of Evidence Project' will fill this lacuna by addressing conceptual, descriptive and normative questions. First, the project will clarify conceptually how we should understand 'weight of evidence' and why this notion should be distinguished from 'balance of evidence.' Second, through textual analyses of court cases and interviews with judges and jurors, the project will examine descriptively whether, and if so how, trial decisions are guided by considerations about the weight of evidence. Finally, the project will examine, from a normative standpoint, to what extent decisions based on weightier evidence better align with the values and goals of a well-functioning trial system, such as accuracy and fairness.

The examination of conceptual, descriptive and normative questions, combined, will lay the groundwork to formulate guidelines for judges and jurors on how to effectively assess the weight of the evidence when they make decisions. This project, then, has the potential to shed new light on the causes of wrongful convictions—an increasingly discussed topic owing to the work of the *Innocence Project*—as well as offer suggestions on how to improve the accuracy and fairness of the US judicial system. Focusing on the legal applications of decision-making, this project also attests that interpretive methods characteristic of the humanities can help to better understand decisions in the face of uncertainty and offer guidelines for reform.

'The Weight of Evidence Project' will last for three years, with the following projected outcomes: (1) essays in magazines and newspapers with a wide readership, along with summaries, infographics and animations that describe in accessible language the central ideas of the project; (2) an on-line portal and website with annotated court decisions that illustrate the legal and epistemic significance, in practice, of evidential weight; (3) an entry in the *Stanford Encyclopedia of Philosophy* titled 'Weight of Evidence' that offers the most up-to-date discussion of the relevant literature in philosophy, law and economics; (4) ten scholarly articles to appear in a collected volume or special issue of a journal.

3 LIST OF PARTICIPANTS

Academic collaborators

| | |
|--------------------------------|--|
| Chan, Lok | Duke University, Social Science Research Institute |
| Dahlman, Christian, codirector | Lund University, Sweden, Dept. of Law |
| Di Bello, Marcello, director | Lehman College, CUNY, Dept. of Philosophy |
| Feduzi, Alberto | SOAS, University of London, School of Finance and Management |
| Fitelson, Branden | Northeastern University, Dept. of Philosophy |
| Gardiner, Georgi | University of Tennessee, Dept. of Philosophy |
| Nordgaard, Anders | Swedish Police Authority, National Forensic Centre |
| Vorms, Marion | University Paris 1, Dept. of Philosophy |

Advisory Board

| | |
|------------------|--|
| Allen, Ronald J. | Northwestern University, School of Law |
| Nance, Dale | Case Western Reserve University, School of Law |
| Skyrms, Brian | UC Irvine, Department of Logic and Philosophy of Science |
| Smith, George E. | Tufts University, Dept. of Philosophy |
| Stein, Alex | Supreme Court of Israel |

4 NARRATIVE

Project title – Towards a Theory of the Weight of Evidence: Rethinking Trial Decisions

A. SUBSTANCE AND CONTEXT

Beyond the Standard Model: Balance and Weight

It is not easy to decide whether a person accused of a crime should be convicted. Suppose Marcello is accused of stealing 500 dollars. His friend Branden says that Marcello visited him yesterday for no special reason. When Marcello left, Branden realized he was missing 500 dollars from his wallet. A store owner says that later that day Marcello paid 500 dollars in cash for merchandise. Is this enough evidence to convict? Surely not. Before jumping to this conclusion, we should consult other witnesses and confirm that Branden is not lying. However, no matter how much evidence we collect, we will never achieve certainty. Trial decisions—as with virtually all decisions—are subject to a margin of error.

How should we decide in the face of uncertainty? Refusing to take a stance because we cannot be sure can hardly be an option. According to a model of decision-making that is often considered the gold standard of rationality in economics and psychology, we should rely on the available evidence and determine the probabilities of the possible scenarios, for example, how likely it is that Marcello stole the money and how likely it is that he did not; next, we should become clear about our preferences, for example, convicting an innocent is usually regarded as far worse than acquitting a guilty individual; and finally, by taking into account our preferences and the probabilities we have assigned to the possible scenarios, we should choose the course of action that would bring about the greatest expected utility from a suitably defined standpoint. We will call this the *standard model of decision-making* (Ramsey, 1931; de Finetti, 1937; Savage, 1954).

Many studies in behavioral economics and cognitive psychology have shown that people's patterns of decision deviate from this model (see, for example, Tversky and Kahneman, 1974; Kahneman and Tversky, 1979; Tversky and Kahneman, 1981). These deviations are often interpreted as an indication that humans make bad decisions or behave irrationally (Kersting and Obst, 2016; Thaler, 2016). Yet some of these deviations might not be irrational after all. Consider the legal context. Under the standard model of decision-making, if the probability of guilt is high, say, at least 95%, a defendant facing trial should be convicted; if it is lower, the defendant should be acquitted (Kaplan, 1968; Kaye, 1999; Tillers and Gottfried, 2007; Weinstein and Dewsbury, 2006; Kaplow, 2012; Hamer, 2014). Curiously, even when people believe the defendant's guilt to be extremely likely, they are unwilling to convict if the supporting evidence is quantitative in nature, a phenomenon known as the Wells effect (Nesson, 1979; Wasserman, 1991; Wells, 1992; Niedermeier et al., 1999; Ebert et al., 2018). It would be hasty to call this and other deviations from the standard model irrational. They might suggest that the model is inadequate, not that people are irrational.¹

By drawing from different areas of research in philosophy of law, epistemology, economics and decision theory, this project defends a richer theory of rational decision-making than the standard model. The central claim—to be supported conceptually, descriptively and normatively—is that decisions should not only be guided by the *balance* of the available evidence, but also by what we will call the *weight* of the evidence. The balance of the evidence can be pictorially represented as a pair of scales measuring the relative strength of the arguments for and against the accused. The more strongly the balance tips against the accused, the higher the probability of guilt. The balance reflects the evidence we have, and there can be little doubt that rational decisions should be guided by the evidence we have.² But we should not rely on the balance alone. Even if the available evidence tips strongly against the accused, it might be incomplete and one-sided because of an uneven allocation of resources between the two parties at trial. For example, it might be that while the prosecutor was able to amass a large body of incriminating evidence, the defense lacked the

resources to discover counterevidence. Before making a decision, then, we should be wary of the missing evidence that, if it had been acquired, could have shifted the balance in significant ways. Some authors have discussed these issues under the name ‘weight of evidence’ and we will adopt the same terminology. But except for a few suggestions in the literature, there is no well-developed theory of weight to date and many questions remain open (Joyce, 2005; Faulkner et al., 2017). Before discussing the major issues to be addressed in the project, we will briefly survey accounts of weight in the literature.

Weight: Completeness and Robustness

At its simplest, weight reflects the amount of evidence presented (Keynes, 1921; Peirce, 1932). A body of evidence that consists of one witness’s testimony will have less weight than one consisting of two testimonies. But this characterization does not specify when a body of evidence would count as sufficiently weighty to serve as the basis for making a decision. A good theory of the weight of evidence should define a criterion for when the acquisition of more evidence can stop and a decision can finally take place. To address this difficulty, some authors have proposed to understand the idea of weight of evidence as a measure of the gap between the expected total evidence and the evidence actually available. The smaller the gap, the weightier the evidence. As some have put it, the weight of the evidence reflects *known unknowns*—that is, the missing evidence we know we do not have (Feduzi, 2010; Faulkner et al., 2017). David Kaye (1986) gives an anecdotal example of a drunk driving case in which the prosecutor called the arresting officer to the stand who testified about the smell and breath of the man arrested, as well as the results of a field sobriety test. The incriminating evidence did not, however, include the results of a breathalyzer. Kaye notes that the jurors expected to hear about that type of evidence but did not, and thus acquitted the defendant. Dale Nance (2016) in a recent book developed this line of thought. For him, the available evidence is sufficiently weighty so long as it is *reasonably complete*—that is, it is all the evidence that someone tasked with making a decision at trial would reasonably expect to see from a conscientious investigation of the facts.

Another existing proposal for characterizing weight is based on the notion of *robustness* (Cohen, 1977, 1986; Stein, 2005; Gärdenfors et al., 1983; Dahlman et al., 2015). Evidence is robust whenever an inference that is drawn from it would survive the objections that a reasonably skilled opponent would present at trial. For example, suppose two eyewitnesses testify at trial that they recognize the defendant as the perpetrator. This evidence, on balance, makes a strong case for guilt, but we cannot tell whether it is robust unless we know it can survive the challenges, objections and counterarguments put forth by a skilled opponent.³

Major Issues to Be Addressed

‘The Weight of Evidence Project’ will focus on three sets of issues—conceptual and theoretical; descriptive and empirical; normative and value-oriented—and then offer recommendations to judges and jurors.

Conceptual and theoretical issues – As a first step, the project will lay out a conceptual map of existing accounts of evidential weight that can serve as the groundwork for a theory of weight. The project will then move beyond existing accounts. The proposals in the literature—weight as completeness of the evidence and weight as robustness—leave open a number of questions. For example, what are the criteria to determine when a body of evidence contains *all* the evidence it would be reasonable to expect there to be in a case? Our working hypothesis is that *narratives* about how different forms of evidence are normally discovered or procured help to assess whether or not the evidence is complete. It is well documented that narratives function as organizing principles for making sense of the available evidence (Pennington and Hastie, 1991; Weisberg, 1996; Burns, 2004; Simon, 2004; Bex, 2011; Griffin, 2013). It is less widely recognized that narratives alert us to the power imbalances that may have affected the presentation and acquisition of the evidence. The project will examine this narrative-based interpretive dimension as integral to an adequate evaluation of the evidence. Another conceptual and theoretical issue to be addressed is whether the notion of

‘weight of evidence’ can be part of the standard model of decision-making. The standard model can capture the notion of ‘balance of the evidence’ in terms of the probability of a hypothesis, say, the probability of guilt. The more strongly the evidence tips against the accused, the higher the probability of guilt. This would seem to leave out the weight of the evidence, although some have argued that the standard model can be more inclusive by making the probability of the hypothesis sensitive to considerations of evidential weight (Kaye, 1979; Friedman, 1996; Taroni et al., 2015). Theorists in law, economics and philosophy have disagreed, arguing that balance and weight cannot be combined into one single measure (Keynes, 1921; Peirce, 1932; Cohen, 1986; Feduzi, 2010; Sjerps et al., 2015). The project will show that the weight and balance of the evidence cannot be combined into one single measure of uncertainty.

Descriptive and empirical issues – If the weight of the evidence plays a distinctive role in guiding trial decisions, this should be reflected in how judges and jurors reason and draw inferences from the evidence. This empirical claim, however, is difficult to verify because trial decisions in the United States about someone’s guilt are not motivated in writing. This makes the decision-making process a black box. (Appellate decisions are motivated in writing, but they primarily address questions of legal interpretation, not how the evidence against an accused is assessed.) To circumvent this difficulty, ‘The Weight of Evidence Project’ will pursue two strategies. First, we will examine court opinions in countries outside the United States, such as Sweden, France and Italy, where a written motivation is required by law. There is textual evidence that weight informs the reasoning of judges in Sweden (Dahlman et al., 2015; Dahlman, 2018), but no study has been done outside Sweden. Institutional differences between countries should not be overlooked, but since our study concerns evidence-based reasoning, the principles guiding a rational evaluation of the evidence should be shared across countries even with different legal traditions. Arguably, conclusions drawn from the textual analysis of court opinions in other countries will be true, under similar circumstances, in the United States. But in order to avoid unwarranted generalizations, the textual analysis of court opinions will be complemented by interviews with judges and jurors in the United States and in select European countries. Textual data together with qualitative data from the interviews will provide the needed empirical grounding.

Normative and value-oriented issues – Besides showing whether, and if so how, judges and jurors take the weight of the evidence into account, another major issue to be addressed concerns the *value* of evidential weight. Why should judges and jurors be concerned with the weight of the evidence when they make decisions? What is gained by doing so? To address these questions, ‘The Weight of Evidence Project’ will first spell out the goals and values that inform the trial, such as the accuracy of the decisions (Damaška, 1997; Taruffo, 2009), their impartiality and fairness (Stein, 2005), their justifiability to the public (Duff et al., 2007), and their finality and ability to deter unwanted behavior in the future (Kaplow, 2015). Next, the project will examine whether, in light of these values and goals, trial decisions that are based on weightier evidence fare better than decisions based on less weighty evidence.

Recommendations to judges and jurors – The project would not be complete without applying its findings to legal practice. If we are correct, it is the task of judges and jurors to assess whether the evidence presented at trial is sufficiently weighty. We will formulate guidelines that can aid judges and jurors in assessing the weight of the evidence. We will also discuss the challenges that such guidelines present.

Significance for the Humanities and Beyond

The question of how evidence should guide decisions in the face of uncertainty is a pressing one, in the legal as well as many other contexts. Yet decision-making under uncertainty is not a topic that humanities scholars often tackle, and the contributions in this area are dominated by scientists. The danger is that the interpretive dimension of decision-making, with its nuances and complexities, will be overlooked. This project attests that scholars in the humanities can make an important contribution to the theory of decision

in distinctive ways. First, the examination of evidential weight as a guide to decisions will help to see that decision-making is richer than assigning probabilities and utilities. Decisions should of course be guided by the evidence, but also be wary of the missing evidence since imbalances of power, individual interests and other exogenous factors contribute to shaping the evidence. Secondly, the textual analysis of court opinions and the interviews with judges and jurors will offer rich insights about human reasoning that can hardly be described by one simple model. Thirdly, this project makes clear that any theory of decision-making should confront value-oriented questions. Does expected utility maximization bring about what we hope to accomplish? What do we hope to accomplish in the first place? These questions are too often overlooked.

‘The Weight of Evidence Project’ has significance beyond the humanities. Over the last twenty-five years, the *Innocence Project* (www.InnocenceProject.org) has drawn attention to numerous cases of wrongful convictions. There is a growing literature on wrongful convictions, their causes and possible remedies (Risinger, 2003; Dwyer et al., 2003; Sangero and Halpert, 2007; Thompson, 2008; Crump, 2009). Perhaps unsurprisingly, these wrongful convictions have occurred when the evidence was relatively thin and lacked weight—for example, it consisted of a single eyewitness testimony or uncorroborated expert identification. The connections that this project will draw between evidential weight and the accuracy and fairness of trial decisions will shed new light on the causes of wrongful convictions and help formulate remedies.

Finally, the research on the role of the weight of evidence in trial decisions will also advance more specific debates in philosophy and its subfields. Epistemologists, ethicists and legal theorists have become increasingly interested in standards of proof at trial, especially about the meaning of ‘proof beyond a reasonable doubt’ (Stein, 2005; Laudan, 2011; Walen, 2015; Gardiner, 2017). This literature often discusses the epistemic and ethical shortcomings of beliefs about guilt that are based on probabilistic and quantitative evidence (Thomson, 1986; Colyvan et al., 2001; Ho, 2008; Redmayne, 2008; Amaya, 2008; Enoch et al., 2012; Haack, 2014; Pritchard, 2015; Smith, 2018; Moss, 2018b; Gardiner, 2018b; Bolinger, forthcoming). The examination of evidential weight will advance these discussions. It will also advance broader debates in philosophy. For example, the notion of weight can clarify the idea of ‘moral encroachment’ in epistemology. Moral encroachment holds that the epistemic justification of a belief can be affected by moral factors (Gardiner, 2018a; Moss, 2018a; Basu, forthcoming). Evidential weight suggests a plausible way to interpret this claim. That is, when moral factors place a more stringent requirement on the justification of beliefs, this translates in the requirement that the evidence be weightier—that more evidence be considered—but not in the requirement that the balance of the evidence make the belief more probable.

B. HISTORY OF THE PROJECT

The collaborators have explored the themes of the project independently. Director and codirector, Marcello Di Bello and Christian Dahlman, have argued in previous publications that evidential weight should guide trial decisions (Dahlman et al., 2015; Dahlman, 2017, 2018; Di Bello, 2013; Di Bello and Verheij, 2018; Di Bello, 2019a,b). Alberto Feduzi has written on evidential weight in economics (Feduzi, 2007, 2010; Feduzi and Runde, 2014; Faulkner et al., 2017). Georgi Gardiner has written on proof beyond a reasonable doubt and wrongful convictions (Gardiner, 2017, 2018b, 2019a,b, forthcoming). Marion Vorms has led a Marie Curie European project on legal reasoning with evidence (Vorms and Hahn, forthcoming). Branden Fitelson has written on probability and evidence (Fitelson, 1999, 2006; Crupi et al., 2008; Easwaran and Fitelson, 2015). ‘The Weight of Evidence Project’ marks the beginning of a collaboration on a topic of mutual interest. The start of the project is expected for October 2020 and its completion for September 2023. The backbone of the project website (URL: www.WeightOfEvidenceProject.org) has been created.

C. COLLABORATORS

Marcello Di Bello, the project director, obtained his Ph.D. in Philosophy in 2013 from Stanford University.

He is an Assistant Professor in Philosophy at Lehman College of the City University of New York. During 2016-17, he was a fellow at the Institute for Advanced Study in Princeton. His fields of expertise are epistemology and the philosophy of law, specifically, reasoning with evidence, decision-making at trial and probability theory. Di Bello has published on these topics in journals such as *Ethics*, *Mind*, *Philosophical Studies*, *Synthese* and *International Journal of Evidence and Proof*. He will be the director of the project and devote 28% of his time during the first year, and 21% of his time during the second and third year. **Christian Dahlman**, the project codirector, is Professor of Jurisprudence at Lund University in Sweden. His research focuses on legal evidence assessment and decision-making. At Lund University, he heads the cross-disciplinary research group ‘Law, Evidence and Cognition,’ a collaboration between the Faculty of Law, the Department of Philosophy and the Department of Psychology. His work has appeared in peer-review journals such as *The Review of Philosophy and Psychology*, *Ratio Juris*, *Argumentation and Law*, *Probability and Risk*. He has also published widely in Swedish. Dahlman will devote 28% of his time during the first year (funded by the Söderberg Foundation), 17% during the second year and 10% during the third (funded by NEH). Director and codirector will coordinate and contribute to all parts of the project, such as conceptual analysis, textual analysis of court opinions, interviews and normative analysis.

Branden Fitelson, a Distinguished Professor of Philosophy at Northeastern University, will contribute to the conceptual and theoretical analysis of the notion of evidential weight. Fitelson has published influential articles on probabilistic analyses of evidence in peer-reviewed journals such as the *Journal of Philosophy*, *Philosophy of Science*, *Philosophical Studies*, *Analysis*, *Journal of Philosophical Logic*, *Synthese* and more. He will devote to this project 4% of his time during both the first and second year. **Alberto Feduzi**, a Senior Lecturer in Strategy at SOAS University of London and a Fellow at Cambridge Judge Business School, will also contribute to the conceptual analysis of evidential weight. He has studied extensively how the concept of evidential weight helps to explain how economic agents make decisions in situations of extreme uncertainty. His research has been published, among others, in *Journal of Economic Behavior and Organization*, *British Journal for the Philosophy of Science* and *Cambridge Journal of Economics*. Feduzi will devote 10% of his time during the first year of the project. **Marion Vorms**, an Associate Professor of Philosophy at Paris 1 University, will help to carry out the textual analysis of court opinions. Her research examines the cognitive bases for judicial reasoning and decision-making. She was awarded a five year fellowship (2019-2014) at *Institut universitaire de France* for a project entitled ‘Reasonable Doubts: Evidential Reasoning in the Courts, the Sciences, and Everyday Life.’ One of the goals of Vorms’s project is to study the inference patterns used in judicial decisions. Vorms will devote 8% of her time during the first and second year. **Lok Chan**, a postdoctoral fellow at Duke University, will coordinate the technical aspects of the project and run the website. He will also conduct interviews with judges and jurors. Chan has extensive experience in programming and data visualization. He is now working on a project funded by the National Institutes of Health on how different types of risk and uncertainty, including evidential weight, can be communicated to patients in the medical context. Lok will devote 20% of his time during the first and second year of the project, and 10% during the third year. **Georgi Gardiner**, an Assistant Professor of Philosophy at the University of Tennessee, will contribute to the analysis of evidential weight from the normative standpoint. Her research areas include the nature of understanding and explanation, social and legal epistemology, virtue epistemology. Gardiner has published in journals such as *Philosophy & Public Affairs*, *Journal of Applied Philosophy* and *American Philosophical Quarterly*. She will devote 8% of her time during the second and third year. **Anders Nordgaard**, a forensic specialist in statistics at the National Forensic Centre of the Swedish Police Authority, will help to formulate appropriate guidelines for judges and legal practitioners on how to assess the weight of the evidence. He has published in peer-reviewed journal such as *Science and Justice*, *Forensic Science International* and *Law, Probability and Risk*. (Funds for Nordgaard’s participation in the project will be provided by his employer.)

Advisory Board

Members of the advisory board are among some of the most distinguished scholars in their field. Two members are philosophers. **Brian Skyrms**, a Fellow of the American Academy of Arts & Sciences and of the National Academy of Sciences, is the Distinguished Professor of Logic and Philosophy of Science and Economics at the University of California, Irvine and a Professor of Philosophy at Stanford University. He has worked on problems in the philosophy of science, causation, decision theory, game theory, and the foundations of probability. Skyrms was one of the first epistemologists to defend the importance of weight (his term is ‘resilience’) for a theory of knowledge (Skyrms, 1967, 1977, 1980). **George Smith** is Professor of Philosophy at Tufts University in Boston. He is one of the leading scholars on Newton in the United States. Smith has also been hired as an expert witness in legal proceedings in the U.S., Canada, Australia and France. He combines experience as an expert witness at trial and also—thanks to his scholarship in the history of science—a deep knowledge of how evidence is weighed against hypotheses.

Three members of the board are legal scholars. **Ronald J. Allen** is the John Henry Wigmore Professor of Law at Northwestern University in Chicago. He is an expert in the fields of evidence, criminal procedure and constitutional law. He has worked with various groups in China to help formulate proposals for legal reform, and he was recently retained by the Tanzanian Government to assist in the reform of their evidence law. **Alex Stein** is a Justice of the Supreme Court of Israel and a former Professor of Law at Brooklyn Law School and Cardozo Law School in New York City. He is one of the leading scholars on the foundations of evidence law and has written extensively on the notion of evidential weight (Stein, 2005). Finally, **Dale Nance** is the Albert J. Weatherhead III and Richard W. Weatherhead Professor in Law and Criminal Justice at Case Western Reserve University, School of Law. He is an internationally recognized scholar of evidence law. He is the author of numerous articles in leading law reviews and an important monograph on evidential weight in decision-making at trial (Nance, 1998, 2008, 2016).

D. METHODS AND EXECUTION

‘The Weight of Evidence Project’ will adopt different types of interpretive analyses and methods. **Conceptual analysis** is a method employed in philosophy in the analytic tradition for testing the boundaries of a given concept or differentiate between two concepts. The method uses examples, counterexamples and stylized scenarios to test whether or not a proposed definition, analysis or characterization of a concept is adequate (Gardiner, 2015a,b; Ichikawa and Steup, 2017). Another form of conceptual analysis is known as ‘reflective equilibrium.’ By working back and forth with our pre-theoretic judgments or intuitions about a given issue, the goal is to arrive at an equilibrium point that takes into account all our pre-theoretical judgments (Daniels, 2018). For the present project, conceptual analysis will serve to address the cluster of conceptual and theoretical issues described in detail earlier. The goal is to arrive at a plausible characterization of evidential weight that combines existing proposals and different pre-theoretical intuitions.

Textual analysis will serve to address the descriptive question of whether judges and jurors take the weight of the evidence into account when they make decisions at trial. We will analyze court opinions in French, Italian and Swedish building on the work of Dahlman (2018). As noted earlier, the choice of these countries is due to the fact that trial decisions about the evidence are not motivated in writing in the United States. The opinions in French will concern court decisions for minor offenses in which judges are expected to decide quickly, with limited evidence. The decisions in Swedish and Italian will instead concern difficult criminal cases with more complex bodies of evidence. The textual analysis will combine qualitative and quantitative methods. First, we will do a close reading of (a subset of) the opinions. The goal will be to identify argument patterns that make reference to the weight of the evidence (Walton et al., 2008). Second, we will analyze a larger set of opinions using a technique called TF-IDF, where TF stands for Term Frequency and IDF denotes Inverse Document Frequency (Ramos, 2003). This method allows to assess the

relative importance of certain terms, such as ‘evidence’ or ‘weight’, in a document. The project will also implement what is known as Topic Modeling and rely on Latent Dirichlet Allocation (LDA), an algorithm that defines a topic in terms of a set of words and models a document as a set of topics. Topic Modeling allows to determine, within a given document, the importance of a topic, such as the weight of evidence (Newman and Block, 2006; Newman et al., 2006).

To strengthen the project’s empirical grounding, we will conduct **interviews** with judges and jurors to probe whether the weight of the evidence guides their reasoning at trial. They will be asked to decide about the criminal liability of a defendant in different hypothetical scenarios. These will be stylized versions of the real cases used for the textual analysis, and the evidence presented in each scenario will differ in terms of its weight. The decisions by judges and jurors will be analyzed to see whether there is a statistically significant difference in the responses depending on variations in the weight of the evidence in the various scenarios. This method of research, often used in cognitive psychology, is now also common in the emergent field of experimental philosophy (Knobe and Nichols, 2008). We will also conduct extensive interviews to understand the underlying reasons for the decisions by judges and jurors. The interview transcripts will undergo a textual analysis analogous to the one to be conducted for the texts of the court opinions.

Value-oriented analysis will serve to address whether decisions that are guided by the weight of the evidence, not just its balance, promote the values and goals of the trial, such as accuracy and fairness. This analysis can be carried out by examining hypothetical trial systems, each differing in the extent to which they promote evidential weight, and then comparing the outcomes they generate, in terms of accuracy, fairness and finality of decisions. This can be done as a purely theoretical investigation relying on so-called ‘thought experiments’ often used in philosophy. A more recent innovation, part of the Digital Humanities, is to rely on computer simulations that examine how different hypothetical, simulated systems behave over time given different parameters (Goldman, 1999; Zollman, 2007; Olsson, 2011; Hegselmann and Krause, 2006). We will use a mixture of thought experiments and computer simulations.

The different methods used in the project are complementary to one another. The conceptual analysis should ensure that ‘weight of evidence’ is a well-defined concept with a clear scope and range of applications. But the conceptual analysis will also be informed by the descriptive findings. The investigation of descriptive issues through textual analyses and interviews may reveal that judges and jurors are guided by a conception of weight of the evidence that is neither reasonable completeness nor robustness (see Section A). If so, the conceptual analysis would be revised accordingly. But the descriptive findings will not count as the last word either. The conception of weight that judges and jurors actually deploy might turn out to be, following the normative and value-oriented analysis, less conducive to accurate and fair decisions. If that were so, there would be good reasons to recommend its revision. By combining the results of the conceptual, descriptive and value-oriented analysis, we will be able to offer recommendations on how judges and jurors can (and should) take the weight of the evidence into account when they make decisions.

Tasks and Division of Labor

The project can be broken down into six tasks. The exact work plan is mapped out in Section E. The first task consists in conceptual groundclearing—that is, to examine arguments for and against weight in the legal and philosophical literature; to draw a conceptual map of accounts of evidential weight; and to offer a more adequate theory of the weight of evidence. This conceptual groundclearing will benefit from different kinds of expertise. Fitelson will share his expertise in probability theory and Feduzi his knowledge of how weight has been deployed in economics. Dahlman and Di Bello will examine evidential weight in theories of legal reasoning. The second task of the project is to read and analyze court opinions. Dahlman will examine decisions in Swedish. Vorms will be responsible for examining decisions written in French. Di Bello will examine court decisions in Italian. Chan will conduct the analysis of court opinions using TF-IDF and

Topic Modeling. The third task of the project is to conduct and analyze the interviews with judges and jurors. This will be Chan's responsibility. The fourth task of the project is to carry out the normative and value-oriented analysis. Dahlman, Di Bello and Gardiner will be responsible for examining the relationship between weight and the values and goals of the trial. Chan will help run computer simulations in this phase of the project. The fifth task is to formulate guidelines that can help judges and other legal practitioners to assess the weight of the evidence when they make decisions. Dahlman and Nordgaard will carry out this task. Finally, the sixth task consists in synthesizing the findings in a way that is accessible to the general public. All the participants will give their contribution to this synthesis.

Meetings among the participants will take place on a regular basis through video conferences. The project director and codirector will supervise the progress of the work. In addition, the project director will travel to Lund for a week to visit the project codirector, and in turn, the project codirector will travel to New York to visit the project director. (Travel costs will be covered by the Söderberg Foundation.)

Output

We will write ten scholarly papers, each 25-30 pages (about 8,000-12,000 words), for a total of 250-300 pages. Five articles will address the theoretical and descriptive issues. **Article #1** ('Do We Need a Theory of Weight?'), co-authored by Dahlman and Di Bello, will examine arguments for and against adopting the weight of the evidence as a distinctive ingredient in a theory of decision-making. **Article #2** ('The Power of Narratives, Weight of Evidence and Known Unknowns'), co-authored by Feduzi and Di Bello, will discuss how narratives about what normally happens guide the assessment of the weight of the evidence and so-called known unknowns. **Article #3** ('How Weight Guides Decisions at Trial: The Case of Sweden, France and Italy'), co-authored by Vorms, Dahlman and Di Bello, will use the textual findings to examine the extent to which weight plays a role in decision-making at trial. **Article #4** ('How Weight Guides Decisions: An Empirical Investigation'), co-authored by Chan and Di Bello, will describe the findings from the interviews with judges and jurors. **Article #5** ('What Theory of Weight Best Fit Decision-making at Trial?'), co-authored by Dahlman and Di Bello, will delineate a theory of evidential weight that is informed by the descriptive findings. Five more articles will present the results of the value-oriented analysis and offer practical recommendations. **Article #6** ('Does Presenting More Evidence at Trial Lead to Fewer Erroneous Decisions?'), co-authored by Di Bello and Chan, will examine the connection between evidential weight and the accuracy of trial decisions. **Article #7** ('Robustness at the End of the Trial'), co-authored by Dahlman and Di Bello, will examine the connection between evidential weight and the final resolution of disputes at trial. **Article #8** ('She Said/He Said: Weight of Evidence and Preponderance in Title IX Proceedings'), written by Gardiner, will argue that the principles of a fair trial require to interpret the standard of proof for civil cases as a threshold on the weight of the evidence. **Article #9** ('Weight of Evidence and the Value of Knowledge'), again written by Gardiner, will examine the notion of evidential weight in light of contemporary theories of knowledge in analytic epistemology, especially developing the relevant alternatives framework for legal proof in Gardiner (2019b). Finally, **Article #10** ('How to Talk About Weight'), co-authored by Dahlman and Nordgaard, will outline guidelines for judges and jurors on how to assess the weight of the evidence at trial.

Another output of the project will be a piece that discusses, in a comprehensive and up-to-date manner, evidential weight. The piece will be an **encyclopedia article**, divided in three sections: weight in philosophy; weight in law; and weight in economics. Its estimated length is about 40 pages (circa 15,000 words). Given their different expertise, Fitelson, Feduzi and Di Bello will write this piece. This will spark new research by drawing connections between research in philosophy, law and economics.

We will also make the project's outputs available on a **website** (URL: www.WeightOfEvidenceProject.org). The website will include an on-line portal showcasing the textual findings based on TF-IDF and Topic Mod-

eling. A digital copy of the court opinions in their original languages will also be made available, along with translations of relevant passages. In addition, pre-print versions of the articles described above will be available for download, along with a description, in accessible language, of the main ideas, as well as infographics and animations when appropriate. The frontend of the website will be programmed in the open-source javascript framework React.js. The web portal will utilize D3.js, a data visualization javascript library. The frontend will be connected to a backend server built using Python and integrated with a PostgreSQL database. The source component will be hosted at Github, the frontend at Netlify and the backend at Heroku. These are professional platforms that offer generous on-line data storage at no cost. This will ensure the long-term availability of the project's outputs. All the collaborators will curate the web content, and Chan will coordinate the technical aspects.

Finally, we will publish three **essays** describing the central ideas of the project in a language accessible to the general public. The working titles of these pieces (7-10 pages each; 3,000-5,000 words) are 'How Known Unknowns Guide Decisions,' 'How Wrongful Convictions Can Be Avoided' and 'What Is Evidence Good For?' These will be co-authored by all the collaborators.

E. WORK PLAN

First six months: October 2020–March 2021 The project will begin with researching existing accounts of weight in the literature in philosophy, law and economics. Fitelson will focus on the literature in philosophy and probability theory, Feduzi on the literature in economics, and Dahlman and Di Bello on the literature in law. Concurrently, Chan will finalize the website and Vorms will start her analysis of court decisions in French. Her findings will start to populate the website with textual data.

Second six months: April 2021–September 2021 Di Bello, Fitelson and Feduzi will work on a first draft of the encyclopedia article on the weight of evidence. Di Bello and Dahlman will draft a piece discussing arguments for and against evidential weight (article #1). Di Bello and Feduzi will work on a piece on the weight of evidence and the power of narratives (article #2). Di Bello, with the help of Dahlman, will also start analyzing court decisions in Italian and Swedish and populate the website with more textual findings. Vorms will continue the analysis of opinions in French. Chan will conduct interviews with judges and jurors.

Third six months: October 2021–March 2022 The draft of the encyclopedia article by Di Bello, Fitelson and Feduzi will be revised on the basis of comments from the advisory board. The same applies to articles #1 and #2 by Di Bello, Dahlman and Feduzi. On the basis of their findings from the textual analysis of court opinions, Vorms, Dahlman and Di Bello will draft a piece on evidential weight in legal reasoning (article #3), as well as a piece developing a theory of weight informed by the textual findings (article #5). Chan and Di Bello will draft a piece describing the interviews with judges and jurors (article #4).

Fourth six months: April 2022–September 2022 The project will focus on the normative and value-oriented analysis. Dahlman, Di Bello and Chan will draft two papers on evidential weight in relation to the values of accuracy and finality, one on accuracy and weight (article # 6), and the other on finality and weight (article # 7). Gardiner will draft a paper on weight of evidence in civil cases, specifically on Title IX litigations (article # 8), as well as a piece on evidential weight and the value of knowledge (article #9).

Fifth six months: October 2022–March 2023 Drafts of the different papers written so far will be circulated for feedback by the advisory board. Dahlman and Nordgaard will draft a piece discussing how to instruct judges and jurors about the weight of the evidence (article #10). All the collaborators in the project will discuss and synthesize the main findings and draft the three essays for the general public.

Sixth six months: April 2023–September 2023 Dahlman and Nordgaard will finalize their paper with recommendations for legal practitioners. Di Bello and Dahlman will curate an edited volume or special issue containing the ten scholarly articles. Chan will create visualizations and infographics that convey the main ideas of the project. The essays for the general public will be finalized and published.

F. FINAL PRODUCT AND DISSEMINATION

The research articles we plan to write will be published either as a **collected volume** edited by Di Bello and Dahlman or as a **special issue** for a journal. At the beginning of the second year, we will contact an academic publisher such as Oxford University Press to gauge interest. We will also contact peer-reviewed journals in philosophy who publish special issues, for example, *Synthese* and *Ratio*. In deciding where to publish our research, we will ensure that our articles are available to the public at no cost. One option is to make available on the project website pre-prints of the articles.

The other output of the project is an **encyclopedia article** on the weight of evidence, co-authored by Feduzi, Fitelson and Di Bello. Our plan is to publish it as part of SEP - Stanford Encyclopedia of Philosophy (plato.stanford.edu). This is a free, on-line repository of high quality articles written by the foremost scholars in the field on virtually any topic in philosophy. It is one of the most well-respected and widely read on-line resources for anyone interested in philosophy. We have already contacted the chief editor, Edward Zalta at Stanford University, who has expressed interest. An alternative venue of publication is *Philosophy Compass*, a well-respected journal in philosophy that offers pointers for researchers, teachers and students, to help them find and interpret the best research in the field.

A **website** (URL: www.WeightOfEvidenceProject.org) will help to disseminate the findings of the project. As described earlier, the website will contain an on-line portal with the findings from the textual analysis of court cases. Links to the full text of the scholarly articles will be available, along with summaries of the main findings and infographics when appropriate. As mentioned earlier, we will use GitHub, Netlify and Heroku to freely store content on-line, ensuring the long-term availability of the project outputs. Since the website will showcase scholarly papers and more accessible materials, its intended audience includes researchers in philosophy, law and the digital humanities, but also educators, legal practitioners and the general public.⁴

Finally, we plan to publish three **essays** in magazines presenting the main ideas of the project. One possible venue is *The Stone* in *The New York Times*, a ‘forum for contemporary philosophers and other thinkers on issues both timely and timeless’ (www.nytimes.com/column/the-stone). Another venue is *Aeon Magazine*, publishing essays about ‘the big questions’ in science, philosophy, society and the arts (aeon.co).

NOTES

¹Psychologists claim to have identified several errors in human reasoning, but when these are scrutinized, some turn out not to be errors (see, e.g., Stenning and van Lambalgen, 2008; Hedden, Forthcoming). Besides, people often deviate from the standard model because of the ‘framing effect’ (Tversky and Kahneman, 1981; Martino et al., 2006), but not all such deviations need to be irrational (Mandel, 2014). For a review of the literature on biases, see Hahn and Harris (2014).

²Some believe that probability is ill-suited to model legal reasoning (Tribe, 1971b,a; Pardo and Allen, 2008; Pardo, 2013; Haack, 2014). There is also a debate about the right probabilistic theory (Lempert, 1977; Dawid, 1987; Cheng, 2013; Biedermann et al., 2014; Cheng and Pardo, 2015; Di Bello, 2019b).

³The notion of robustness also underpins the literature on sensitivity analysis (Berger et al., 1994), value of information (Raiffa and Schlaifer, 1961; Good, 1967, 1985), resilience and stability (Skyrms, 1977, 1980; Davidson and Pargetter, 1987; Leitgeb, 2014).

⁴An example of this dissemination strategy is the Stanford Open Policing Project (openpolicing.stanford.edu).

Bibliography

- Amaya, A. (2008). Justification, coherence, and legal responsibility in legal fact-finding, *Journal of Social Epistemology* **5**: 206–319.
- Basu, R. (forthcoming). The wrongs of racist beliefs, *Philosophical Studies* .
- Berger, J. O., Moreno, E., Pericchi, L. R., Bayarri, M. J., Bernardo, J. M., Cano, J. A., la Horra, J. D., acinto Martín, Ríos-Insúa, D., Betrò, B., Dasgupta, A., Gustafson, P., Wasserman, L., Kadane, J. B., Srinivasan, C., Lavine, M., O'Hagan, A., Polasek, W., Robert, C. P., Goutis, C., Ruggeri, F., Salinetti, G. and Sivaganesan, S. (1994). An overview of robust bayesian analysis, *Test* **3**(1): 5–125.
- Bex, F. J. (2011). *Arguments, Stories and Criminal Evidence: A Formal Hybrid Theory*, Springer.
- Biedermann, A., Hicks, T., Taroni, F., Champod, C. and Aitken, C. (2014). On the use of the likelihood ratio for forensic evaluation: Response to Fenton et al., *Science and Justice* **54**(4): 316–318.
- Bolinger, R. J. (forthcoming). The rational impermissibility of accepting (some) racial generalizations, *Synthese* .
- Burns, R. P. (2004). The distinctiveness of trial narratives, in A. Duff, L. Farmer, S. Marshall and V. Tadros (eds), *The Trial on Trial (VOL 1): Truth and Due Process*, Hart Publishing.
- Cheng, E. (2013). Reconceptualizing the burden of proof, *Yale Law Journal* **122**(5): 1104–1371.
- Cheng, E. and Pardo, M. S. (2015). Accuracy, optimality and the preponderance standard, *Law, Probability and Risk* **14**(3): 193–212.
- Cohen, J. L. (1977). *The Probable and the Provable*, Oxford University Press.
- Cohen, J. L. (1986). Twelve questions about Keynes's concept of weight, *British Journal for the Philosophy of Science* **37**(3): 263–278.
- Colyvan, M., Regan, H. M. and Ferson, S. (2001). Is it a crime to belong to a reference class?, *Journal of Political Philosophy* **9**(2): 168–181.
- Crump, D. (2009). Eyewitness corroboration requirements as protections against wrongful conviction: The hidden questions, *Ohio State Journal of Criminal Law* **7**(1): 361–376.
- Crupi, V., Fitelson, B. and Tentori, K. (2008). Probability, confirmation, and the conjunction fallacy, *Thinking and Reasoning* **14**(2): 189–199.
- Dahlman, C. (2017). Determining the base rate for guilt, *Law, Probability and Risk* **17**(1): 15–28.
- Dahlman, C. (2018). *Beviskraft: Metod för bevisvärdering i brottmål*, Norstedts.
- Dahlman, C., Wahlberg, L. and Sarwar, F. (2015). Robust trust in expert testimony, *Humana.Mente - Journal of Philosophical Studies* **18**: 17–37.
- Damaška, M. R. (1997). *Evidence Law Adrift*, Yale University Press.

- Daniels, N. (2018). Reflective equilibrium, *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University.
- Davidson, B. and Pargetter, R. (1987). Guilt beyond reasonable doubt, *Australasian Journal of Philosophy* **65**(2): 182–187.
- Dawid, P. (1987). The difficulty about conjunction, *Journal of the Royal Statistical Society. Series D (The Statistician)*, **36**(2/3): 91–92.
- de Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives, *Annales de l'Institut Henri Poincaré* **7**: 1–68.
- Di Bello, M. (2013). *Statistics and Probability in Criminal Trials: The Good, the Bad and the Ugly*, PhD thesis, Stanford University.
- Di Bello, M. (2019a). Probability and plausibility in juridical proof, *International Journal of Evidence and Proof*.
- Di Bello, M. (2019b). Trial by statistics: Is a high probability of guilt enough to convict?, *Mind* **128**(512): 1045–1084.
- Di Bello, M. and Verheij, B. (2018). Evidential reasoning, *Handbook of Legal Reasoning and Argumentation*, Springer.
- Duff, A., Farmer, L., Marshall, S. and Tadros, V. (2007). *The Trial on Trial (vol 3): Towards a Normative Theory of the Criminal Trial*, Hart Publishing.
- Dwyer, J., Neufeld, P. and Scheck, B. (2003). *Actual Innocence*, Signet.
- Easwaran, K. and Fitelson, B. (2015). Accuracy, coherence, and evidence, in T. S. Gendler and J. Hawthorne (eds), *Oxford Studies in Epistemology*, Vol. 5, Oxford University Press.
- Ebert, P. A., Smith, M. and Durbach, I. (2018). Lottery judgments: A philosophical and experimental study, *Philosophical Psychology* **31**(1): 110–138.
- Enoch, D., Fisher, T. and Spectre, L. (2012). Statistical evidence, sensitivity, and the legal value of knowledge, *Philosophy and Public Affairs* **40**(3): 197–224.
- Faulkner, P., Feduzi, A. and Runde, J. (2017). Unknowns, Black Swans and the risk/uncertainty distinction, *Cambridge Journal of Economics* **41**(5): 1279–1302.
- Feduzi, A. (2007). On the relationship between Keynes's conception of evidential weight and the ellberg paradox, *Journal of Economic Psychology* **28**(5): 545–565.
- Feduzi, A. (2010). On Keynes's conception of the weight of evidence, *Journal of Economic Behavior and Organization* **76**(2): 338–351.
- Feduzi, A. and Runde, J. (2014). Uncovering unknown unknowns: Towards a Baconian approach to management decision-making, *Organizational Behavior and Human Decision Processes* **124**: 268–283.
- Fitelson, B. (1999). The plurality of Bayesian measures of confirmation and the problem of measure sensitivity, *Philosophy of Science* **66**(Proceedings): 362–378.

- Fitelson, B. (2006). Likelihoodism, Bayesianism, and relational confirmation, *Synthese* **16**(47): 1–22.
- Friedman, R. D. (1996). Assessing evidence, *Michigan Law Review* **94**: 1810–1838.
- Gärdenfors, P., Hansson, B., Sahlin, N.-E. and Halldén, S. (1983). *Evidentiary Value: Philosophical, Judicial, and Psychological Aspects of a Theory. Essays Dedicated to Sören Halldén on his Sixtieth Birthday*, C.W.K. Gleerup.
- Gardiner, G. (2015a). Normalcy and the contents of philosophical judgements, *Inquiry* **58**(7): 700–740.
- Gardiner, G. (2015b). Teleologies and the methodology of epistemology, in D. K. Henderson and J. Greco (eds), *Epistemic Evaluation: Point and Purpose in Epistemology*, Oxford University Press.
- Gardiner, G. (2017). In defence of reasonable doubt, *Journal of Applied Philosophy* **34**(2): 221–241.
- Gardiner, G. (2018a). Evidentialism and moral encroachment, in K. McCain (ed.), *Believing in Accordance with the Evidence: New Essays on Evidentialism*, Springer.
- Gardiner, G. (2018b). Legal burdens of proof and statistical evidence, in D. Coady and J. Chase (eds), *Routledge Handbook of Applied Epistemology*, Routledge.
- Gardiner, G. (2019a). Legal epistemology, in D. Pritchard (ed.), *Oxford Bibliographies: Philosophy*, Oxford University Press.
- Gardiner, G. (2019b). The reasonable and the relevant: Legal standards of proof, *Philosophy and Public Affairs* **47**(3): 288–318.
- Gardiner, G. (forthcoming). Legal evidence and knowledge, in M. Lasonen-Aarnio and C. Littlejohn (eds), *The Routledge Handbook of the Philosophy of Evidence*, Routledge.
- Goldman, A. (1999). *Knowledge in a Social World*, Oxford University Press.
- Good, I. (1967). On the principle of total evidence, *British Journal for the Philosophy of Science* **17**(4): 319–321.
- Good, I. (1985). Weight of evidence: A brief survey, in J. Bernardo, M. DeGroot, D. Lindley and A. Smith (eds), *Bayesian Statistics*, Elsevier Science.
- Griffin, L. K. (2013). Narrative, truth, trial, *Georgetown Law Journal* **101**: 281–335.
- Haack, S. (2014). *Evidence Matters: Science, Proof, and Truth in the Law*, Cambridge University Press.
- Hahn, U. and Harris, A. J. (2014). What does it mean to be biased: Motivated reasoning and rationality, *Psychology of Learning and Motivation* **61**: 41–102.
- Hamer, D. (2014). Presumptions, standards and burdens: managing the cost of error, *Law, Probability and Risk* **13**: 221–242.
- Hedden, B. (Forthcoming). Hindsight bias is not a bias, *Analysis*.
- Hegselmann, R. and Krause, U. (2006). Truth and cognitive division of labour: First steps towards a computer-aided social epistemology, *Journal of Artificial Societies and Social Simulation* **9**(3): URL:<http://jasss.soc.surrey.ac.uk/9/3/1.html>.

- Ho, H. L. (2008). *Philosophy of Evidence Law*, Oxford University Press.
- Ichikawa, J. J. and Steup, M. (2017). The analysis of knowledge, in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University.
- Joyce, J. M. (2005). How probabilities reflect evidence, *Philosophical Perspectives* **19**: 154–178.
- Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk, *Econometrica* **47**(2): 263–292.
- Kaplan, J. (1968). Decision theory and the fact-finding process, *Stanford Law Review* **20**(6): 1065–1092.
- Kaplow, L. (2012). Burden of proof, *Yale Law Journal* **121**(4): 738–1013.
- Kaplow, L. (2015). Information and the aim of adjudication: Truth or consequences?, *Stanford Law Review* **67**(6): 1303–1371.
- Kaye, D. H. (1979). The paradox of the Gatecrasher and other stories, *The Arizona State Law Journal* pp. 101–110.
- Kaye, D. H. (1986). Do we need a calculus of weight to understand proof beyond a reasonable doubt?, *Boston University Law Review* **66**: 657–672.
- Kaye, D. H. (1999). Clarifying the burden of persuasion: What Bayesian rules do and not do, *International Commentary on Evidence* **3**: 1–28.
- Kersting, F. and Obst, D. (2016). Behavioral economics, *Exploring Economics* (URL: www.exploring-economics.org/en/orientation/behavioral-economics/).
- Keynes, J. M. (1921). A treatise on probability.
- Knobe, J. and Nichols, S. (eds) (2008). *Experimental Philosophy*, Oxford University Press.
- Laudan, L. (2011). The rules of trial, political morality and the costs of error: or, is proof beyond a reasonable doubt doing more harm than good?, in L. Green and B. Leiter (eds), *Oxford Studies in Philosophy of Law*, Vol. 1, Oxford University Press.
- Leitgeb, H. (2014). The stability theory of belief, *Philosophical Review* **123**(2): 131–171.
- Lempert, R. O. (1977). Modeling relevance, *Michigan Law Review* **75**(5/6): 1021–1057.
- Mandel, D. R. (2014). Do framing effects reveal irrational choice?, *Journal of Experimental Psychology* **142**(3): 1185–1198.
- Martino, B. D., Kumaran, D., Seymour, B. and Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain, *Science* **313**(5787): 684–687.
- Moss, S. (2018a). Moral encroachment, *Proceedings of the Aristotelian Society* **118**(2): 177–205.
- Moss, S. (2018b). *Probabilistic Knowledge*, Oxford University Press.
- Nance, D. A. (1998). Evidential completeness and the burden of proof, *Hastings Law Journal* **49**(621-662).

- Nance, D. A. (2008). The weights of evidence, *Journal of Social Epistemology* **5**: 267–281.
- Nance, D. A. (2016). *The Burdens of Proof: Discriminatory Power, Weight of Evidence, and Tenacity of Belief*, Cambridge University Press.
- Nesson, C. R. (1979). Reasonable doubt and permissive inferences: The value of complexity, *Harvard Law Review* **92**(6): 1187–1225.
- Newman, D. J. and Block, S. (2006). Probabilistic topic decomposition of an eighteenth century newspaper, *Journal of the American Society for Information Science and Technology* **57**(6): 753–767.
- Newman, D., Smyth, P. and Steyvers, M. (2006). Scalable parallel topic models, *Journal of Intelligence Community Research and Development* .
- Niedermeier, K. E., Kerr, N. L. and Messeé, L. A. (1999). Jurors’ use of naked statistical evidence: Exploring bases and implications of the Wells effect, *Journal of Personality and Social Psychology* **76**(4): 533–542.
- Olsson, E. J. (2011). A simulation approach to veritistic social epistemology, *Episteme* **8**(2): 127–143.
- Pardo, M. S. (2013). The nature and purpose of evidence theory, *Vanderbilt Law Review* **66**(2): 547–613.
- Pardo, M. S. and Allen, R. J. (2008). Judicial proof and the best explanation, *Law and Philosophy* **27**(3): 223–268.
- Peirce, C. S. (1932). *Collected Papers (vol 2)*, Harvard University Press.
- Pennington, N. and Hastie, R. (1991). A cognitive theory of juror decision making: the story model, *Cardozo Law Review* **13**: 519–557.
- Pritchard, D. (2015). Risk, *Metaphilosophy* **46**(3): 436–461.
- Raiffa, H. and Schlaifer, R. (1961). *Applied Statistical Decision Theory*, Harvard University.
- Ramos, J. (2003). Using TF-IDF to determine word relevance in document queries, *Proceedings of the First Instructional Conference on Machine Learning* **241**: 133–42.
- Ramsey, F. P. (1931). Truth and probability, in R. Braithwaite (ed.), *The Foundations of Mathematics and other Logical Essays*, Kegan, Paul, Trench, Trubner and Co. Ltd, chapter 7, pp. 156–198.
- Redmayne, M. (2008). Exploring the proof paradoxes, *Legal Theory* **14**(4): 281–309.
- Risinger, M. (2003). Innocents convicted: An empirically justified factual wrongful conviction rate, *Journal of Criminal Law and Criminology* **97**(3): 761–806.
- Sangero, B. and Halpert, M. (2007). Why a conviction should not be based on a single piece of evidence: A proposal for reform, *Jurimetrics Journal* **48**(1): 43–94.
- Savage, L. (1954). *The Foundations of Statistics*, Wiley.
- Simon, D. (2004). A third view of the black box: Cognitive coherence in legal decision making, *The University of Chicago Law Review* **71**(2): 511–586.

- Sjerps, M. J., Alberink, I., Bolck, A., Stoel, R. D., Vergeer, P. and van Zanten, J. H. (2015). Uncertainty and Ir: to integrate or not to integrate, that's the question, *Law, Probability and Risk* **15**(1): 23–29.
- Skyrms, B. (1967). The explication of 'X knows that p', *Journal of Philosophy* **64**(12): 373–389.
- Skyrms, B. (1977). Resiliency, propensity, and causal necessity, *Journal of Philosophy* **74**(11): 704–713.
- Skyrms, B. (1980). *Causal Necessity*, Yale University Press.
- Smith, M. (2018). When does evidence suffice for conviction?, *Mind* **127**(508): 1193–1218.
- Stein, A. (2005). *Foundations of Evidence Law*, Oxford University Press.
- Stenning, K. and van Lambalgen, M. (2008). *Human Reasoning and Cognitive Science*, MIT Press.
- Taroni, F., Bozza, S., Biederman, A. and Aitken, C. G. (2015). Dismissal of the illusion of uncertainty in the assessment of a likelihood ratio, *Law, Probability and Risk* **15**(1): 1–16.
- Taruffo, M. (2009). *La semplice verità*, Laterza.
- Thaler, R. H. (2016). Behavioral economics: Past, present, and future, *American Economic Review* **196**(7): 1577–1600.
- Thompson, S. G. (2008). Beyond a reasonable doubt? Reconsidering uncorroborated eyewitness identification testimony, *UC Davis Law Review* **41**(4): 1487–1545.
- Thomson, J. J. (1986). Liability and individualized evidence, *Law and Contemporary Problems* **49**(3): 199–219.
- Tillers, P. and Gottfried, J. (2007). Case comment—United States v. Copeland, 369 F. Supp. 2d 275 (E.D.N.Y. 2005): A collateral attack on the legal maxim that proof beyond a reasonable doubt is unquantifiable?, *Law, Probability and Risk* **5**(2): 135–157.
- Tribe, L. H. (1971a). A further critique of mathematical proof, *Harvard Law Review* **84**(8): 1810–1820.
- Tribe, L. H. (1971b). Trial by mathematics: Precision and ritual in the legal process, *Harvard Law Review* **84**(6): 1329–1393.
- Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases, *Science* **185**(4157): 1124–1131.
- Tversky, A. and Kahneman, D. (1981). The framing of decisions and the psychology of choice, *Science* **211**(4481): 453–458.
- Vorms, M. and Hahn, U. (forthcoming). In the space of reasonable doubt, *Synthese*.
- Walen, A. (2015). Proof beyond a reasonable doubt: A balanced retributive account, *Louisiana Law Review* **76**(2): 355–446.
- Walton, D., Reed, C. and Macagno, F. (2008). *Argumentation Schemes*, Cambridge University Press.
- Wasserman, D. T. (1991). The morality of statistical proof and the risk of mistaken liability, *Cardozo Law Review* **13**: 935–976.

- Weinstein, J. B. and Dewsbury, I. (2006). Comment on the meaning of ‘proof beyond a reasonable doubt’, *Law, Probability and Risk* **5**(2): 167–173.
- Weisberg, R. (1996). Proclaiming trials as narratives: Premises and pretenses, in P. Brooks and P. Gewirtz (eds), *Law’s Stories: Narrative and Rhetoric in the Law*, Yale University Press, pp. 61–83.
- Wells, G. L. (1992). Naked statistical evidence of liability: Is subjective probability enough?, *Journal of Personality and Social Psychology* **62**(5): 793–752.
- Zollman, K. J. S. (2007). The communication structure of epistemic communities, *Philosophy of Science* **74**(5): 574–587.