

7-2001

Theories of Uncertainty: Explaining the Possible Sources of Error in Inferences

Vern R. Walker

Maurice A. Deane School of Law at Hofstra University

Follow this and additional works at: https://scholarlycommons.law.hofstra.edu/faculty_scholarship

Recommended Citation

Vern R. Walker, *Theories of Uncertainty: Explaining the Possible Sources of Error in Inferences*, 22 Cardozo L. Rev. 1523 (2001)
Available at: https://scholarlycommons.law.hofstra.edu/faculty_scholarship/166

This Article is brought to you for free and open access by Scholarly Commons at Hofstra Law. It has been accepted for inclusion in Hofstra Law Faculty Scholarship by an authorized administrator of Scholarly Commons at Hofstra Law. For more information, please contact lawcls@hofstra.edu.

THEORIES OF UNCERTAINTY: EXPLAINING THE POSSIBLE SOURCES OF ERROR IN INFERENCES

Vern R. Walker*

A central task in legal factfinding is evaluating the warrant for a finding or the soundness of an inference from the evidentiary propositions to a conclusion. This task is especially difficult when there is much at stake, but the evidence is incomplete and the soundness of the inference is uncertain. Analyses of how to improve such inferences have been made at various levels of generality, and for different types of evidence. For example, one general problem is distinguishing “scientific knowledge” from “junk science,” as required for admissibility in judicial proceedings under Federal Rule of Evidence 702, following the *Daubert v. Merrell Dow Pharmaceuticals, Inc.*¹ decision.² Another general problem is evaluating inferences about unique historical events, the kind of factfinding necessary in criminal cases.³ As opposed to such general problems, some theorists address only particular areas where inferences are difficult in law, such as the “lost chance” cases,⁴ cases involving “indeterminate plaintiffs,”⁵

* Professor of Law, Hofstra University School of Law. Some of the ideas in this Article were first presented on April 30, 2000, at the symposium, Artificial Intelligence and Judicial Proof sponsored by the Benjamin N. Cardozo School of Law, the *Cardozo Law Review*, and the Jacob Burns Institute for Advanced Legal Studies, in New York, New York. Preparation of this Article was supported in part by a research grant from Hofstra University.

¹ 509 U.S. 579 (1993).

² See, e.g., *Gen. Elec. Co. v. Joiner*, 522 U.S. 136 (1997); *Kennedy v. Collagen Corp.*, 161 F.3d 1226 (9th Cir. 1998); cf. *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 147 (1999) (holding that Federal Rule of Evidence 702 mandates a judicial gatekeeping obligation to determine evidentiary reliability for all expert “knowledge,” not merely “scientific knowledge”).

³ See, e.g., JOSEPH B. KADANE & DAVID A. SCHUM, *A PROBABILISTIC ANALYSIS OF THE SACCO AND VANZETTI EVIDENCE* (1996).

⁴ See, e.g., Joseph H. King, Jr., *Causation, Valuation, and Chance in Personal Injury Torts Involving Preexisting Conditions and Future Consequences*, 90 YALE L.J. 1353 (1981); Vern R. Walker, *Direct Inference in the Lost Chance Cases: Factfinding Constraints Under Minimal Fairness to Parties*, 23 HOFSTRA L. REV. 247 (1994).

⁵ See, e.g., Bert Black & David E. Lilienfeld, *Epidemiologic Proof in Toxic Tort Litigation*, 52 FORDHAM L. REV. 732, 767-69, 782-84 (1984); Michael Dore, *A Commentary on the Use of Epidemiological Evidence in Demonstrating Cause-in-Fact*, 7

inferences from "naked statistical evidence,"⁶ or inferences based on DNA identification.⁷ Such problems of correct inference cannot be solved purely by formal logic, nor can theorists merely duplicate the role of the factfinder by evaluating the specific evidence in a particular case. To be useful as theories of inference, accounts can be neither too general nor too specific. They must provide useful models for handling recurring types of inference in situations where findings must be warranted by incomplete evidence.⁸

In Western logic, the traditional strategy has been to distill from inference problems any steps that have no uncertainty associated with them. Such inferential atoms could then be the foundation for inference. Traditional logic therefore isolates deductively valid inferences. In a deductively valid inference, if the premises are true, then the conclusion is necessarily true.⁹ The conclusion must be true whenever the premises are true. Deductively valid inferences are also examples of what logicians and mathematicians call "monotonic" inferences. An inference is monotonic if new evidence cannot affect the validity of the

HARV. ENVTL. L. REV. 429 (1983); Khristine L. Hall & Ellen K. Silbergeld, *Reappraising Epidemiology: A Response to Mr. Dore*, 7 HARV. ENVTL. L. REV. 441 (1983); Steve Gold, Note, *Causation in Toxic Torts: Burdens of Proof, Standards of Persuasion, and Statistical Evidence*, 96 YALE L.J. 376 (1986).

⁶ See, e.g., Craig R. Callen, *Adjudication and the Appearance of Statistical Evidence*, 65 TUL. L. REV. 457 (1991); David Kaye, *The Limits of the Preponderance of the Evidence Standard: Justifiably Naked Statistical Evidence and Multiple Causation*, 1982 AM. B. FOUND. RES. J. 487; David Kaye, *The Paradox of the Gatecrasher and Other Stories*, 1979 ARIZ. ST. L.J. 101 (1979); Richard Lempert, *The New Evidence Scholarship: Analyzing the Process of Proof*, 66 B.U. L. REV. 439 (1986); David Kaye, *Naked Statistical Evidence*, 89 YALE L.J. 601 (1980) (book review).

⁷ See, e.g., 2 PAUL C. GIANNELLI & EDWARD J. IMWINKELRIED, *SCIENTIFIC EVIDENCE* 1-39 (2d ed. 1993); D.H. KAY, *SCIENCE IN EVIDENCE* 153-258 (1997); NATIONAL RESEARCH COUNCIL, *THE EVALUATION OF FORENSIC DNA EVIDENCE* (1996); Judith A. McKenna et al., *Reference Guide on Forensic DNA Evidence*, in FED. JUDICIAL CTR., *REFERENCE MANUAL ON SCIENTIFIC EVIDENCE* 273-329 (1994); Kenneth E. Melson, *Determining Individuality by DNA*, in ANDRE A. MOENSSSENS, *SCIENTIFIC EVIDENCE IN CIVIL AND CRIMINAL CASES* 870-963 (4th ed. 1995).

⁸ A number of legal theorists have discussed the general problem of how to take completeness of the evidence into account in factfinding. See, e.g., Craig R. Callen, *Kicking Rocks with Dr. Johnson: A Comment on Professor Allen's Theory*, 13 CARDOZO L. REV. 423, 431-39 (1991); Neil B. Cohen, *Conceptualizing Proof and Calculating Probabilities: A Response to Professor Kaye*, 73 CORNELL L. REV. 78, 85-86 (1987); D.H. Kaye, *Do We Need a Calculus of Weight to Understand Proof Beyond a Reasonable Doubt?*, 66 B.U. L. REV. 657, 658 (1986); Richard O. Lempert, *Modeling Relevance*, 75 MICH. L. REV. 1021, 1047-48 (1977); Lempert, *supra* note 6, at 473-74; Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1349-50 (1971); Walker, *supra* note 4, at 286-97.

⁹ See, e.g., IRVING M. COPI & CARL COHEN, *INTRODUCTION TO LOGIC* 25 (10th ed. 1998).

inference.¹⁰ A deductively valid inference is monotonic because the inference will remain valid regardless of how many new premises are added and regardless of whether those additional premises are true or false. Of course, new information might undermine the truth of the original premises, and therefore the "soundness" of the original inference, but the validity of that inference is unaffected.¹¹

In contrast to monotonic inference, if reasoning is "nonmonotonic," then the addition of new premises can cast doubt on the validity of the original inference.¹² Indeed, the new information might not undermine the truth of the original premises at all, but only the validity of the inference. For example, the original premises might be the generalization that birds can fly and the singular proposition that individual *A* is a bird. Given these premises, it is reasonable to infer that *A* can fly. If new evidence arises, however, indicating that *A* is a penguin, then this inference is no longer warranted. The problem in this example is that the original premises did not include all the information material to the inference. When probabilistic, presumptive or default reasoning is based on incomplete information, new evidence can undermine previously warranted inferences.

This Article develops the concept of a "theory of uncertainty" as a tool for warranting nonmonotonic inferences. It proposes that theories of uncertainty have an explanatory function in inference that is roughly parallel to the role played by theories of causation in the empirical sciences. An adequate theory of uncertainty provides a framework for explaining the warrant for nonmonotonic inferences by explaining the possible sources of error in such inferences. A theory of uncertainty should explain the extent to which the available evidence warrants a particular conclusion and the kinds and degrees of uncertainty associated

¹⁰ See, e.g., GERHARD BREWKA ET AL., NONMONOTONIC REASONING: AN OVERVIEW 1, 24 (1997); GILBERT HARMAN, REASONING, MEANING AND MIND 30-32 (1999).

¹¹ A deductive inference is "sound" if it is deductively valid and all of its premises are in fact true. See COPI & COHEN, *supra* note 9, at 32.

¹² For discussions of nonmonotonic reasoning, see, for example, BREWKA ET AL., *supra* note 10, at 1-3, 23-24; HARMAN, *supra* note 10, at 30-32. Such reasoning includes, but is not limited to, the use of so-called statistical syllogisms. A traditionally formulated statistical syllogism has the following form: "Most *A*'s are *B*'s. This is an *A*. Therefore, this is (probably) a *B*." JOHN L. POLLOCK, NOMIC PROBABILITY AND THE FOUNDATIONS OF INDUCTION 76 (1990). For an early recognition of the difficulty of such inferences in legal theory, see George F. James, *Relevancy, Probability, and the Law*, 29 CAL. L. REV. 689 (1941), arguing that the reasoning "[n]ine-tenths of all *As* are *X*, *B* is an *A*, therefore the chances are nine to one that *B* is *X*" is not logically valid "except upon the assumption that *As* may be treated as a uniform class with respect to the probability of their being *X*." *Id.* at 697.

with that conclusion. Such a theory may also help predict how additional evidence might affect the warrant for that inference. The proposal is that factfinders develop theories of uncertainty for particular inferences from particular bodies of evidence; and that the adequacy of those theories plays an important role in warranting the inference itself.

This analysis is presented in three parts. Part I discusses theories of uncertainty generally, and some of the characteristics such theories have in common. Part II examines a theory of uncertainty for conclusions warranted by scientific data. Such a theory is particularly useful in evaluating conclusions about generic causation and in assessing whether proffered scientific opinions are admissible under *Daubert* and *General Electric Co. v. Joiner*.¹³ Part III discusses the evidence charting theory of David Schum, which is especially promising for inferences about unique historical events.¹⁴ These two extended examples illustrate how the concept of a theory of uncertainty is useful in analyzing the nonmonotonic inferences important in legal factfinding.

I. THE ROLE OF THEORIES OF UNCERTAINTY IN WARRANTING NONMONOTONIC INFERENCES

Factfinding in a legal context has, as a principal goal, an "epistemic objective": the generation of findings that are *accurate* descriptions of events and that are *warranted* by the legally available evidence.¹⁵ Factfinding undoubtedly has other objectives as well, but the epistemic objective is a paramount one. This Article does not make an argument for adopting this epistemic objective, nor does it discuss at any length why legal factfinding has the epistemic objective as a primary goal. The Article assumes, first, that findings that are accurate have instrumental value in decisionmaking, and perhaps intrinsic value as well. Accuracy should increase the effectiveness and efficiency of legal activities, as well as their fairness to private parties. Second, findings that are warranted by the legally available evidence help provide due process and help legitimate the use of governmental

¹³ 522 U.S. 136 (1997).

¹⁴ See, e.g., DAVID A. SCHUM, *THE EVIDENTIAL FOUNDATIONS OF PROBABILISTIC REASONING* (1994); KADANE & SCHUM, *supra* note 3.

¹⁵ See VERN R. WALKER, *Preponderance, Probability and Warranted Factfinding*, 62 BROOK. L. REV. 1075, 1079 (1996) (legal factfinding has the goal of producing "knowledge" in the sense of "warranted true belief"). For philosophical discussions of warrant as being the difference between mere true belief and human knowledge, see, for example, ALVIN PLANTINGA, *WARRANT: THE CURRENT DEBATE* (1993), and JOHN L. POLLOCK & JOSEPH CRUZ, *CONTEMPORARY THEORIES OF KNOWLEDGE* 7-10 (2d ed. 1999).

power. They also increase the transparency of governmental process. The requirement of warrant also provides a check on accuracy, as well as a strategy for increasing accuracy, at least in the long run. This Article assumes the epistemic objective as a primary goal in law, and uses the epistemic objective to develop a model for analyzing nonmonotonic inferences in legal factfinding.

Legal factfinding adopts as a starting point our everyday experience and our common account of it. That is, the initial touchstone for accuracy and warrant is ordinary reasoning or "common sense." The common-sense ontology of everyday objects and events is assumed to be adequate and does not itself need to be warranted. Under most everyday circumstances, people describe the world automatically, without reflection or deliberation. Sitting in my study on a winter's day, I see a Postal Service truck through the window, I hear the voice of my neighbor as she answers her door, and I remember that the person who lives down the street has gone on vacation. Such beliefs about the world are so often accurate that there is seldom a need to investigate their warrant. But in legal proceedings, when liberty and wealth hang on the outcome, a witness's descriptions of the world are routinely challenged. Whenever this is the case, we then examine the warrant for that assertion. We evaluate the probative value of the evidence we possess, and we may look for additional evidence. If we encounter uncertainty and have a need to resolve it, we turn to warrant as a means of doing so. It is this close functional relationship between accuracy and warrant that leads to bundling them together in the epistemic objective.

This Article proposes that a theory of uncertainty is one type of construct we employ to help warrant a conclusion when we are uncertain about the truth of that conclusion. This Part of the Article discusses in general terms what I mean by a "theory of uncertainty." The first section addresses the circumstances under which we begin a search for the warrant behind our beliefs. Once we understand when and why we initiate a search for warrant, we begin to understand what can bring the search to a successful close. In theory, we might successfully conclude the search by obtaining all relevant information. But factfinding usually must proceed on the basis of incomplete information. One important question is when information is sufficient even though it is incomplete. I propose that in most circumstances we develop a theory of uncertainty in lieu of obtaining complete evidence. The second section examines three dimensions that are inherent in any adequate theory of uncertainty—namely, linguistic, logical, and causal dimensions. The third section discusses why adequate

theories of uncertainty generally integrate all three dimensions in order to explain the potential for error within nonmonotonic inference.

A. *Triggering and Ending the Search for Warrant*

We typically engage in the activity of warranting because we encounter an apparent inconsistency within a set of propositions, all of which we are otherwise inclined to believe. We go through life acquiring beliefs, holding them with more or less conviction, often modifying them, and sometimes discarding them. At any point in time during this dynamic process, we may find ourselves inclined to believe two propositions that are inconsistent with each other, in the sense that they cannot both be true. The strongest form of inconsistency is *logical*, in which the inconsistency is demonstrable by considering only the logical form and meanings of the propositions involved. The proposition "I am the only person in the house" and its negation ("It is not the case that I am the only person in the house") are logically inconsistent. To understand the meanings of the two propositions is to understand that they cannot both be true. In symbols, the propositions **P** and **Not-P** are logically inconsistent. But it is generally useful to relax the notion of "inconsistent" to include *physical* inconsistency and say that propositions are inconsistent when they cannot all be true given our most basic beliefs about physical objects. For example, alibi evidence in law can create an inconsistency in the evidence because we believe that a single human being cannot be in two very different locations at precisely the same time. The inconsistency involved is not a logical impossibility, for the physical world could in theory obey other laws than it does, but we consider it a physical impossibility given our standard worldview.¹⁶

When we find ourselves inclined to believe a set of inconsistent propositions,¹⁷ we can conclude deductively that, if the

¹⁶ Stated in terms of propositional logic, we are inclined to believe three propositions: **P**, **Q** and **Not-both-P-and-Q**. Let **P** be the proposition that a person is in place P at a certain time, while **Q** is the proposition that the same person is in place Q at precisely the same time, with locations P and Q being very different. **Not-both-P-and-Q** states our belief that the person could not be in both places at the same time. The set of these three propositions leads deductively to contradictions. For example, from **P** and **Not-both-P-and-Q**, we deduce **Not-Q**, which yields the contradiction **Q-and-Not-Q**. Discarding any one of the original three propositions eliminates the contradiction. However, in this example the proposition we are least likely to discard is **Not-both-P-and-Q**, because it is too well warranted by our physical theories and general experience.

¹⁷ It is a question for psychology and psycholinguistics whether we should say that human beings *hold* inconsistent beliefs or that they *withhold* belief when they become aware of an inconsistency. There is no need or basis in this Article for deciding which is the better description of human behavior. It would seem to make no difference to the

inconsistency is real, then one or more of the inconsistent propositions must be false. An apparent inconsistency within a set of our beliefs creates a condition of uncertainty concerning that set. Although we conclude that one or more of the relevant propositions must be false, we may not know initially which proposition or propositions are in error. For example, the inconsistency might result from a complicated series of inferences based on disparate pieces of evidence, and untangling that logical web to find the problem is not always simple.¹⁸ Multiple assumptions or inferences might be in error, as even the simplest example will illustrate. Suppose that I am in my office, and I believe that my new, red sports car is parked in parking lot B. My colleague arrives from lot B, however, and claims that he did not see it there and that he would have noticed it if it had been there. I am now uncertain whether my car is there or not. I have some reason to believe that the car is in lot B and some reason to believe that the car is not there. There are numerous ways to account for this apparent inconsistency. Perhaps I have forgotten which car I drove today or where I parked it. Perhaps my colleague was inattentive when walking through the parking lot or is playing a joke on me. Maybe someone has stolen my car. The epistemic task, if I wish to undertake it, is to determine which of these propositions is (probably) in error. One task of a theory of uncertainty is to help understand and resolve such inconsistencies.

Before examining this epistemic task further, we need to recognize and set aside the pragmatic aspects of the task. The epistemic task may be to resolve the inconsistency, but in practice I might decide not to undertake that task. It takes time and effort to resolve inconsistencies among propositions we are otherwise inclined to believe, and the decision to do so always occurs within a practical context. Sometimes resolving the inconsistency takes so little effort that we barely notice it as a separate activity. If parking lot B is visible from my office window, I may merely look to see whether my car is there. A deliberate decision, however, to investigate and resolve inconsistencies involves weighing the expected costs of investigation against the benefits expected from the investigation. We tolerate many inconsistencies because resolving them is not worth our while.

concept of a theory of uncertainty. The important point here is that inconsistency among propositions leads to the search for warrant. Whether human beings can simultaneously believe such inconsistent propositions is irrelevant.

¹⁸ For an attempt to chart all of the evidence in a single legal case, see KADANE & SCHUM, *supra* note 3. Some aspects of this theory are discussed *infra* Part III.

For example, we tolerate different degrees of precision in measurements, depending upon what is at stake and what it would take to provide more precision. Suppose you and I are looking at an object in the distance, and your estimate of the distance to that object is 300 meters, while I think the distance is about 250 meters. I have some independent reason to trust your judgment in such matters, but I also think I am right in this particular case. In many circumstances I might not care which estimate is more accurate—for example, when I am relaxed and watching the sun set. In other circumstances, however, I might care a great deal and might take additional measures to try to resolve the issue. For example, if I were a professional golfer in a golf tournament, with money and pride at stake, I might try (within reason and the rules of the game) to resolve such an inconsistency. If I am not an excellent swimmer, and you and I are estimating the distance across the river, I might attempt to swim 250 meters, but not 300 meters. If there is a desperate need to do so, I might attempt to swim either distance. But whether golfer or swimmer, I may have no incentive to resolve inconsistent estimates of 269 meters and 270 meters. Such a difference would not affect my choice of golf club or my decision to try to swim the river. The marginal benefit to be gained by attempting to resolve such an inconsequential difference in estimates would be outweighed by any effort needed to do so.

We can generalize upon these illustrations. At any point in the dynamic of human activity, a person might decide not to resolve inconsistencies among his or her beliefs and decide to “live with” a particular uncertainty or potential for error. That person might decide to spend the available time and resources differently. We should not confuse pragmatic decisions about which inconsistencies to tolerate or leave unresolved with the problem of how best to resolve such inconsistencies given the decision to do so. Decision theorists are concerned with how to analyze and improve the pragmatic decisions. This Article addresses only the epistemic problem of how to proceed with resolving the inconsistency if we wish to do so.

The epistemic strategy is to resolve the inconsistency by evaluating the evidentiary basis for each proposition in the set of inconsistent propositions, and to determine which of the inconsistent propositions is more likely to be in error. The central strategy for making progress is to focus directly on warrant and sources of error, rather than on accuracy. Theories of uncertainty explain the origin and propagation of falsehood, as contrasted with traditional logic, which explains the preservation of truth. The approach is to analyze the potential for error inherent in each of

the inconsistent propositions by understanding the ways in which each proposition might be false despite the evidence in its favor. An objective of any theory of uncertainty is to understand the circumstances under which each of the inconsistent propositions can appear to be warranted and yet turn out to be false. If we understand the possible sources of error in the supporting nonmonotonic inferences, we may be able to determine which propositions are more likely to be in error, or to devise a method for investigating the propositions further.

A comprehensive analysis of any particular inconsistency would investigate the warrant for each of the inconsistent propositions, as well as the warrant for each proposition that could be inferred from each of those propositions. That is, warrant would be investigated “up” the inference chains that lead *to* the proposition, and “down” the inference chains that lead *from* that proposition to other propositions. Inferences in both directions supply the warrant for the conclusion. Ideally, we would construct a complete account of all of the possible sources of error associated with each of the inconsistent propositions. Of course, there are theoretical reasons why such a comprehensive approach is impossible. The number of warranting relations increases exponentially with each inferential step away from any particular proposition, because each line of reasoning usually involves two or more premises for each conclusion and each conclusion can play a role in multiple lines of reasoning. For practical reasons, even an approximation to a comprehensive approach is usually too costly and too inefficient. It would also be ineffective, because we would abandon the attempt long before the analysis is completed. Human beings—and even computers—need heuristics to shorten any investigation into warrant. We need an efficient process for identifying possible sources of error and the propositions more likely to be false. A good theory of uncertainty would have such a heuristic function. It would help us decide which uncertainties are more important than others, and why trying to resolve some uncertainties would be more productive than trying to resolve others.

This Article cannot analyze theories of uncertainty theoretically or comprehensively. The purpose here is to illustrate how a theory of uncertainty can improve nonmonotonic reasoning epistemically by explaining the possible sources of error. Parts II and III of this Article discuss two different theories of uncertainty, which address two inference problems of importance in legal factfinding—inferences about generic causation and inferences about particular historical events. Before turning to these

examples, however, there are several generic features of theories of uncertainty that highlight what to look for in such a theory.

B. *The Three Dimensions of a Theory of Uncertainty*

There are at least three distinct but interactive dimensions to any account of warrant for a proposition: a linguistic dimension, a logical dimension, and a causal dimension. This section briefly discusses each of these dimensions. The relevance of each dimension to a portrayal of uncertainty is based on the notion of warrant and the task of warranting. The final section of this Part discusses the usefulness and importance of each of these dimensions in legal factfinding.

1. The Linguistic Dimension

A theory of uncertainty must have a linguistic dimension because inconsistent propositions are indeed *propositions*, and propositions are expressed in language. Propositions are statements that are capable of being true or false, capable of being believed or doubted or denied, and capable of being translated into different languages.¹⁹ In terms of English grammar, a proposition is what is typically expressed by a simple declarative sentence.²⁰ When apparent inconsistencies arise within a set of propositions, the first step should be to investigate the meanings and structures of the linguistic expressions being used. Sometimes the apparent inconsistency vanishes once the linguistic expressions are correctly understood. Clarification of meaning or reference can disclose that the inconsistency was only “apparent.” Even when true inconsistencies remain after we clarify meaning and linguistic structure, we understand the nature of the remaining inconsistency more precisely: we understand more precisely what has to be resolved. Therefore, adequate theories of uncertainty generally have a linguistic component or dimension.

One kind of problem is clarifying the meanings of the words used. The desire to eliminate ambiguity of definition is one reason that scientific and technical enterprises develop their own technical languages. In a technical language, a new word may be unambiguously assigned to every useful meaning or definition. Unambiguous assignment is assured only if a word that has no prior meaning is invented. There are various goals and tradeoffs in inventing technical words. One loss in inventing such words is

¹⁹ See, e.g., COPI & COHEN, *supra* note 9, at 5-7.

²⁰ See *id.* at 5. Linguistic expressions comprise a very broad category, including not only natural languages, but also technical or specialized languages, and the formal or artificial languages found in mathematics and symbolic logic.

that prior connotations can sometimes inspire new concepts, hypotheses, and theories. But inventors of technical words often decide to pay that price in order to achieve clarity of definition. Lawmakers, judges, and lawyers define and refine the meanings of technical legal terms in order to treat similar cases similarly, and clarifying the relevance and probative value of evidence is critical to this effort. To the extent that terms with legal significance, such as “fault” or “intent” or “cause,” are vague or ambiguous, the litigating attorneys, the judge, and the factfinder do not know what evidence should count for or against a finding. Other problems occur in trying to mesh ordinary language, technical legal terms, and scientific or other specialized terms into a single factfinding process. For example, under *Daubert*, federal courts must base admissibility rulings in part on an evaluation of the warrant for an expert opinion that is expressed in scientific terms, but in order to do so they must determine the relevance of that opinion to a finding expressed in technical legal terms.²¹ A nonexpert factfinder, on the other hand, is charged with weighing the probative value of both expert and nonexpert testimony, and with reaching conclusions expressed in legal terms, while evaluating the inferences involved in terms of the ordinary language of the factfinder.

The second kind of problem of meaning is the problem of denotation or reference. A word can be used to refer to many particular objects or events, and confusion over reference occurs when a word is intended to refer to one thing but is understood as referring to another. Pronouns (such as “he,” “hers,” “it”) are prone to use in confusing ways, and even proper names (such as “John Smith”) can have multiple referents. Expressions about relationships can create confusion of reference as well, such as “the next one” or “the other one.” With some words the fact of relationship is only implicit, and error is likely if the terms of relationship are not clarified. An example is the concept of mathematical probability, which involves a ratio between two numbers. The failure to identify the denominator or reference class can lead to both confusion and error. When the problem is ambiguity of reference, merely inventing new words is not always a solution. The problem lies in using the naming words of a language in unclear ways.

An account of the meaning of a proposition explains not only the connotations and denotations of the terms used, but also the meaning of the proposition as a whole. Understanding what the

²¹ For an example of such a difference between scientific and legal concepts, see *infra* note 55.

proposition means involves understanding, at some level of detail, what counts for or against its truth. In order to understand the meaning of "the river is 300 meters wide," I must understand something about how to test its truth. I may not understand its meaning completely, or have a means to verify its accuracy, but I must have a reasonably appropriate notion of what counts as evidence for or against the proposition. I must know the differences between "wide" and "deep," "meters" and "inches," and "300" and "30." I do not understand "the river is 300 meters wide" if I think that it means the same as "the river is 30 inches deep," or that a test of accuracy consists of inserting a yardstick into the river vertically until it touches the river bottom. One technique for learning a proposition's meaning is to observe the methods used to prove its truth or falsehood. At least with regard to a proposition that purports to describe publicly observable objects or events, knowing what that proposition means involves knowing what the world would look like if the proposition is true or false.

An adequate theory of uncertainty would take into account the meaning of the linguistic expressions used to state both the conclusions to be reached and the propositions used to warrant those conclusions. Such a theory need not provide a comprehensive account of that meaning, but only an account that clarifies any unwanted uncertainty in that meaning. A theory of uncertainty has primarily instrumental value toward achieving the epistemic objective. It should guide the investigation into warrant and improve the warrant and accuracy of the inferences, at least in the long run. Even minimally explanatory theories can improve the likelihood of epistemic success by elucidating the meaning and structure of the propositions relevant to the inconsistency to be resolved. Moreover, the linguistic dimension of a successful theory would provide at least minimal criteria for what counts as relevant evidence. A very good theory of uncertainty would explain how the missing information makes a difference to the outcome of the inference. In the case of nonmonotonic reasoning, we ultimately would like to understand the nature and inferential significance of the missing evidence. Understanding the materiality of the available and missing evidence to the inference is as important as understanding the meaning and relevance of the evidence.

2. The Logical Dimension

Theories of uncertainty should take into account the logical relationships among the propositions relevant to the apparent inconsistency. Propositional logic investigates logical relations

between propositions (principally truth-functional relationships) and provides models for an important class of deductively valid inferences.²² Predicate or quantification logic analyzes the internal structure of simple propositions in ways that explain additional deductive inferences.²³ Systems of modal logic analyze deductive inferences involving necessity, possibility, and impossibility.²⁴ All of these formal systems investigate how logical form can explain the possible deductive inferences among propositions.

Logical form is central to warrant and, therefore, to theories of uncertainty. First, the recognition of inconsistency is what triggers the search for warrant and the need for a theory of uncertainty. Recognizing a logical inconsistency requires understanding the meaning and tautologous nature of propositions of the form, "**P** and **Not-P** cannot both be true."²⁵ Understanding a physical inconsistency requires adding the basic physical laws to the logical laws, and recognizing contradictions derived from the combination of those laws. For example, given the size and physical nature of a human body, we can deduce that the same body cannot be in two different locations one hundred miles apart at precisely the same time. An adequate theory of uncertainty would explain why or how two or more statements are in fact inconsistent.

Second, an adequate theory of uncertainty would identify inferences that are warranted by the rules of logic. For an inference to be deductively valid means that if the premises are true, then the conclusion must also be true. An inference is deductively valid if simultaneously asserting the premises and denying the conclusion creates a contradiction. Deductive reasoning warrants the conclusion simply by the logical form of that reasoning. An example of a deductively valid line of reasoning is the form:

If **P** then **Q**.

P.

Therefore, **Q**.

Whenever this logical form accurately models a line of argument,²⁶ then the two premises deductively warrant this

²² See, e.g., COPI & COHEN, *supra* note 9, at chs. 10-11; Albert E. Blumberg, *Logic, Modern*, in 5 THE ENCYCLOPEDIA OF PHILOSOPHY 12 (Paul Edwards ed., 1967).

²³ See COPI & COHEN, *supra* note 9, at ch. 12.

²⁴ See, e.g., A.N. Prior, *Logic, Modal*, in 5 THE ENCYCLOPEDIA OF PHILOSOPHY, *supra* note 22, at 5.

²⁵ This is also known as the principle of contradiction, or the law of noncontradiction. See COPI & COHEN, *supra* note 9, at 389.

²⁶ Because arguments are expressed in language, there is always a question of whether

conclusion. For example, if "P" is replaced by the proposition "this is a peach," and "Q" is replaced by "this contains carbohydrates," and the demonstrative pronoun "this" in each proposition refers to a single object, then the following reasoning is valid:

If this is a peach, then this contains carbohydrates.

This is a peach.

Therefore, this contains carbohydrates.

Propositional logic provides an explanation for why the conclusion must be true whenever the premises are true. Moreover, logical theory explains why inferences of this form are monotonic, and why additional premises cannot undermine the validity of the inference. In any line of reasoning that is deductively valid, the truth of the premises, together with logical theory, warrants the conclusion.

A line of reasoning may lack deductive validity but nevertheless be warranted. That is, an inference might be acceptable epistemically, but it is not contradictory to assert the premises and yet deny the conclusion. For example, some inferences are warranted inductively, such as inferences from sample data to conclusions about the population from which the sample was drawn. It remains to be seen whether logical theories can be formalized for nondeductive but warranted inferences. In the last several decades, more dynamic logics have been developed in an effort to model the relationships involved when tentatively held beliefs are (for example) "defeated" or "undercut" by later beliefs.²⁷ Even if completely general theories of warrant are not available for nonmonotonic reasoning, we do attempt to give reasons for making certain kinds of inferences based on certain kinds of evidence. A major portion of both the sciences and common sense is devoted to developing such theories. An adequate theory of uncertainty would explain the possible sources of error in such inferences and the ways that the premises could be

a logical structure accurately models what is being argued by the linguistic expressions used. If the logical model is itself constructed in a formalist language, the translation of the natural language argument into formal logic can be considered part of the linguistic dimension of a theory of uncertainty. On this view, whether an argument form written in symbolic logic adequately models the *meaning* of a natural language argument is part of the linguistic task of a theory of uncertainty. Even on this view, however, the theory of why only certain argument forms in symbolic notation are deductively valid is part of logical theory, not linguistics.

²⁷ See, e.g., BREWKA ET. AL, *supra* note 10 (examining "default logic"); ISAAC LEVI, *THE FIXATION OF BELIEF AND ITS UNDOING* (1991) (examining the justification for revising a system of beliefs by "contraction"); POLLOCK, *supra* note 12 (examining "prima facie reasoning" and "defeaters").

true but the conclusion false. Such a theory of uncertainty has to take into account the logical structures of the arguments involved.

3. The Causal Dimension

In addition to linguistic meaning and logical structure, we often use causal accounts to help warrant propositions about events. Examples include propositions warranted by first-person sensory perceptions, propositions about the mental states of other persons warranted by their behavior, and propositions warranted by the testimony of others. First, assertions warranted by personal observations draw upon an understanding of how sensory perception works. We do not often question the evidentiary basis for our observations about the world around us. But when our assertions do come under scrutiny, we normally warrant our claims using a causal theory of perception. From the earliest days of human awareness of hallucinations and sensory illusions, people have turned to causal accounts of perception to explain what is warranted on the basis of perception and what is not. Today we turn to optics, experimental psychology, and cognitive theory to explain how vision works and how it can lead to mistaken judgments. The cognitive and psychological sciences have also enabled us to “deceive” the senses through sensory manipulation.

The role of perception in warranting propositions describing the world depends upon the calibration of perceptions with the use of certain descriptive predicates. A person learns how to use appropriately many basic descriptive predicates (such as the basic colors, sounds, and textures) in certain standard contexts (of lighting, for example). The learner’s reports are thereby calibrated to those of others under such standard conditions. The person then learns through experience how distance and ambient light can affect the accuracy of reporting. People develop causal theories of light and perspective, although sometimes they are very crude theories. For example, in the case of vision, we use very simple geometry to understand how perceived size decreases as the distance from the perceiver to the perceived object increases. Human observers must learn to use a causal account of perception to adjust their assertions about reality, to the extent that those assertions are warranted in part by personal perceptions.

Another familiar example of the causal dimension in explaining possible sources of error is describing another person’s beliefs or other mental states on the basis of his or her behavior. Once again, most of the time we are so adept at formulating sufficiently true propositions about other persons that we have no occasion to state the warrant for such propositions. But not

infrequently in everyday life, and quite often in law, we investigate the warrant for judgments about another person's veracity or motive or emotion. If pressed for the warrant for such propositions, we resort to causal theories about "how people work," and also about the personality of the particular person under discussion. The better we know a person, the more likely it is that we can describe accurately "what is going on inside of them." We are forced to develop causal theories about specific people because of the tremendous variability in how individual human beings think and feel and act. We instinctively credit the opinions of those who know them well, because the individual's abilities, values, and patterns of behavior are essential elements of the warrant.

The third example is warranting a proposition on the basis of the testimony of another person. When a witness testifies based on his or her past personal observations, the warrant for a factfinder's conclusion based on that testimony draws on causal accounts of the witness's past perceptions, current memory, and veracity, as well as of the factfinder's perception of the witness and ability to judge the witness's credibility. The types of causal theory needed are not the same for each component of the warrant. Some components draw upon optics and sensory psychology, others upon knowledge of personality and human behavior. David Schum's attempt to model this immensely complicated warrant is discussed in Part III of this Article. The sole point here is that causal explanations play an inescapable role in any adequate theory of uncertainty about conclusions based on such testimony.

These three examples merely draw attention to the warranting role of causal accounts. This Article cannot provide a theoretical account of the relationship between causal theories and theories of uncertainty. There is no reason to think, however, that this choice of examples biases the conclusion. If we were to draw the examples from scientifically warranted propositions, we would still see the critical role of causal accounts. Scientific conclusions, for example, are at bottom warranted by the perceptions of individual investigators and the reports of many investigators taken collectively. Moreover, the conclusions based on that data depend for their warrant on a causal account of how the study was designed and how the data were gathered. This type of warrant will be discussed in greater detail in Part II.

C. *Integrative Theories of Uncertainty*

The previous section discussed the linguistic, logical, and causal dimensions usually present in adequate theories of

uncertainty. This section argues that these three dimensions are normally integrated in any particular theory of uncertainty—at least for any theory of practical use in law. Part of the argument for integration is based on an analogy of function between causal theories in the empirical sciences and epistemic theories of uncertainty. First, theories of causation and theories of uncertainty both play explanatory roles. A causal theory explains the occurrence of a given event in terms of the causal relationships between that event and other events. A theory of uncertainty explains the potential for error in a conclusion, using the warranting relationships between that proposition and the evidence. While a good causal theory explains why events happen as they do, a good theory of uncertainty explains why a given proposition might be false. This explanatory function is aimed at increasing our understanding about either the events or the propositions under study.

Second, good causal theories are not only explanatory, but also predictive. A good causal theory allows us to predict a chain of events into the future. If a particular causal account is true (for example, one about lunar eclipses), then it can be used to predict other events in the future, given the requisite conditions. Some of those predictions provide tests of the accuracy of the conclusions. If the requisite test conditions occur but the predicted event does not, there is some reason to think that the causal account itself is flawed.

The logical form of a simple predictive test is a variation on “denying the consequent” or “modus tollens.” If **T** stands for the theory or set of propositions that constitutes the causal account, **C** stands for the propositions describing the test conditions, and **E** stands for the proposition describing the predicted event, then the logical form of the reasoning behind the test is:

If **T** and **C**, then **E**.

C and **Not-E**.

Therefore, **Not-T**.

The causal account **T** comes into play in warranting the conditional premise “If **T** and **C**, then **E**,” which in turn helps warrant rejecting **T** if the test fails. The argument behind the test is that if the causal theory (and conditional premise) is true and the test conditions **C** are satisfied, but the predicted event **E** does not occur, then the theory **T** is (probably) false.

The predictive function of a theory of uncertainty is evident when the theory explains the limitations in the causal account and predicts how the test itself might be flawed. Theories of uncertainty about each proposition might explain how the test

conditions **C** might be satisfied and the predicted event **E** not occur, but the causal account **T** might still be true. Different theories of uncertainty might explain how **Not-E** could be concluded inaccurately (perhaps we misread the evidence, and **E** actually occurred), or how **C** could be in error (the conditions **C** were not satisfied after all). Our reasoning could be wrong, and the conditional “If **T** and **C**, then **E**” is incorrect. Such theories of uncertainty, taken together, would explain how it is possible to conduct the test and obtain results that seem to disprove **T**, yet for **T** in fact to be true. These theories of uncertainty probably point the way to further predictions or tests (with their own epistemic limitations).

An ordinary and concrete example is predicting the mental states of another person. Suppose I infer from your behavior that you are angry with something I have done (I infer that **A**). I reason that if you are angry (**A**), and if I ask you whether you are angry (**C**), you will acknowledge your anger (**E**): that is, I reason that if **A** and **C**, then **E**. Suppose I then conduct the test and ask you whether you are angry, but you do not acknowledge being angry (**Not-E**). I therefore decide that you are not angry, but I cannot be sure that this is the correct conclusion to draw. I might be misreading your behavior as a denial, when you do not intend to deny being angry. Or you might be hiding your anger for some reason I do not understand. I might then reevaluate my knowledge of the circumstances and my theories about your personality, in light of uncertainties about both.

Causal theories and theories of uncertainty also generate useful descriptive predicates. Describing what I see as “an automobile” bundles into that description an indefinite list of predictions about possible sequences of events. For example, if I were to walk in front of an automobile while it is moving, I predict that I would certainly be struck, probably injured, and possibly even killed. I know this because automobiles are relatively hard physical objects, compared to my body, and because I have an intuitive grasp of the basic laws of motion and mass. Descriptive terms can be highly efficient means of communicating or reasoning if they bundle into their meaning implications that have predictive and explanatory value. The descriptive function of a theory of uncertainty is less obvious from the grammar of English. Descriptive uncertainty tends to be expressed through adjectives (such as “somewhat” or “vague”), adverbs (“apparently” or “seemingly”), and verbs (“appears to be”), rather than through nouns. Nevertheless, there are numerous ways in English to express the uncertainty associated with a descriptive sentence

(such as “the automobile seemed to be moving”). Words expressing uncertainty often modify entire propositions and thereby qualify the descriptive function of that proposition.

Theories of uncertainty are like causal theories in tending to integrate these explanatory, predictive, and descriptive functions into a single account. Theories of uncertainty, however, are also distinguishable from causal theories. The primary function of a theory of uncertainty is not to explain events, but to explain how conclusions could be false, despite the available evidence. In the context of legal factfinding, we use theories of uncertainty to achieve the epistemic objective. We try to understand the warrant behind apparently inconsistent propositions, taking into account all of the available evidence. Through such theories, we understand the risk of error associated with finding a proposition to be true. If we can reduce that risk of error sufficiently, we might also achieve more accurate inferences, at least in the long run. Whether our factfinding accuracy actually improves depends upon many factors, only one of which is the adequacy of our theory of uncertainty.

Theories of uncertainty try to reconcile evidence, reasoning, and conclusions into a coherent whole, by explaining how error might arise in making either a type of inference or a specific inference. Theories of uncertainty draw selectively upon causal theories, as needed to address the particular propositions at issue, with their linguistic and logical dimensions. In order to explain error, theories of uncertainty must explain why certain events *are* “evidence” relevant to certain propositions, and why the available evidence warrants certain propositions and not others. These tasks require integrating the three dimensions of linguistic meaning, logical implications, and causal theories into an adequate theory of uncertainty about each possible conclusion. Such theories can be very general, applying to a type of evidence or a type of conclusion. Other theories of uncertainty might be specific to the particular evidence and conclusion under review. Just as useful causal theories range from broad scientific theories (such as theories of mechanics for any moving body) to particularized lay theories (such as theories to explain my son’s buying decisions), so too theories of uncertainty range from the generic to the specific. We tailor our theories of uncertainty to the task at hand. A theory of uncertainty might be either simple or complex, depending upon the pragmatic context and the level of uncertainty that is tolerable.

In a trial, the competing “stories” that adversarial attorneys tell to juries in the closing arguments usually exhibit both generic

and specific theories of uncertainty.²⁸ Such stories are designed to integrate all of the evidence presented and to emphasize some uncertainties over others, depending on the party's interests. The competing accounts explore the linguistic ambiguities of testimonial evidence, the appropriate interpretations of other evidence, the logical implications of the advocated accounts of the evidence, the possible motivations of the witnesses in giving the testimony, and so forth. Expert testimony is evaluated in much the same way, but with more emphasis on theories of uncertainty about inferences to be drawn from scientific data and from any physical evidence in the particular case, such as DNA evidence. The linguistic, logical, and causal accounts of all the legally available evidence are integrated into a single theory about which inferences are warranted and which are not. Because the evidence is almost always incomplete and warranted certitude is normally out of the question, the contest is over which conclusions are better warranted by the evidence. Each party develops theories of uncertainty to explain how the *available* evidence warrants a particular conclusion, and how the *missing* evidence (were it available) would not undermine that conclusion. Tactically speaking, each party's theory of uncertainty is framed by the standard of proof established for the case and by how the burdens of persuasion are assigned.

Legal factfinding is not confined to adversarial trials in courts, where the admissibility of evidence is governed by the law of evidence. Factfinding often occurs in administrative proceedings, for example, in either an adjudicatory or a rule-making setting, through either formal or informal procedures.²⁹ The theories of uncertainty employed in such settings are as pragmatic as the stories advocated in the closing arguments of trial attorneys. But unlike the ad hoc theories of uncertainty often found in judicial proceedings, theories of uncertainty in administrative proceedings can be more generic and explicit, with administrative regulations or policies specifying the default inferences to be drawn from a type of evidence.³⁰ The rules allocating burdens of proof also play

²⁸ On the role of stories in litigation, see, for example, Ronald J. Allen, *Factual Ambiguity and a Theory of Evidence*, 87 NW. U. L. REV. 604 (1994), and Ronald J. Allen, *The Nature of Juridical Proof*, 13 CARDOZO L. REV. 373 (1991).

²⁹ See, e.g., 5 U.S.C. §§ 553-57 (2000).

³⁰ See, e.g., Vern R. Walker, *Keeping the WTO from Becoming the "World Trans-science Organization": Scientific Uncertainty, Science Policy, and Factfinding in the Growth Hormones Dispute*, 31 CORNELL INT'L L.J. 251, 258-67 (1998) (discussing the nature of "science policies" adopted by administrative agencies, and giving examples of science policies adopted by the U.S. EPA for use in assessing the carcinogenic risk to humans posed by chemical agents).

a dominant role in administrative factfinding, as they do in judicial proceedings.³¹

The next two Parts of this Article discuss two examples of theories of uncertainty. The first is generic and oriented toward types of evidence, while the second lays out a general mode of analysis and applies it to descriptions of the evidence from a particular criminal trial. The first example, discussed in Part II, is a theory useful in evaluating any inferences warranted by scientific evidence, such as inferences about generic causation. Inferences about generic causation are crucial, for example, in tort cases and in administrative proceedings dealing with health, safety, and the environment. The second example, discussed in Part III, illustrates a method for analyzing inferences about unique, historical events—inferences central in most criminal and civil adjudications. The discussions of these two examples also explore additional aspects of any theory of uncertainty.

II. A THEORY OF UNCERTAINTY FOR SCIENTIFIC EVIDENCE ABOUT GENERIC CAUSATION

This Part of the Article outlines a general theory of uncertainty, the details of which are presented elsewhere.³² This theory explains the types of uncertainty inherent in inferences from scientific evidence to findings about generic causation. Assertions about generic causation state causal relationships among types of events, but not directly about specific events. For example, such a proposition might state that a certain type or level of exposure to a particular chemical compound can cause harmful neurological effects in humans.³³ Generic causation is distinguishable from specific causation. An example of specific causation is whether a particular effect (such as the illness of a particular plaintiff) was caused by a particular event (perhaps the plaintiff's exposure to a particular toxic agent). In a toxic tort case, a plaintiff must prove both generic and specific causation,

³¹ See, e.g., *Env'tl. Def. Fund, Inc. v. EPA*, 548 F.2d 998 (D.C. Cir. 1976), *supplemental opinion* (1977), *cert. denied sub nom.*, *Velsicol Chem. Corp. v. EPA*, 431 U.S. 925 (1977) (deciding burden of proof in administrative proceedings to suspend pesticide registrations).

³² The basic theory has been presented in detail in Vern R. Walker, *The Siren Songs of Science: Toward a Taxonomy of Scientific Uncertainty for Decisionmakers*, 23 CONN. L. REV. 567 (1991) [hereinafter Walker, *Siren Songs*]. The theory has been applied to explain determinations of baseline risk in tort law. See Vern R. Walker, *The Concept of Baseline Risk in Tort Litigation*, 80 KY. L.J. 631, 647-72 (1991-92) [hereinafter Walker, *Baseline Risk*].

³³ A proposition about generic causation can be either affirmative or negative—that is, it can either assert or deny generic causation.

while generic causation alone might be a sufficient legal basis for a protective health or safety regulation.

The theory of uncertainty discussed here takes into account the linguistic, logical, and causal dimensions of propositions about generic causation. The theory classifies types of uncertainty and explains why those types of uncertainty are inherent in each proposition about generic causation, at least when that proposition is warranted by data and analysis using standard scientific method. In each category of uncertainty, the central question is how error can arise in the inference from scientific data to a conclusion about generic causation. An ancillary question is how the extent of possible error can be characterized. The goal may be to reduce the amount of uncertainty and to understand the extent of residual uncertainty. This discussion has two parts: the first section enumerates the six types of uncertainty recognized in the theory, and the second section discusses how the theory might be useful in warranting nonmonotonic inferences in legal factfinding.

A. *Types of Inherent Uncertainty*

The theory classifies uncertainties into six types or categories: concept uncertainty, measurement uncertainty, calculation uncertainty, sampling uncertainty, mathematical modeling uncertainty, and causal uncertainty. Each type of uncertainty arises at a distinct step in the scientific process of warranting a conclusion about generic causation and adds a new possibility of error. In the end, after the process of scientific warranting is complete, there is inherent in the conclusion a residual degree of uncertainty of each type. The types are therefore logically distinct, generally cumulative, and probably inherent in any such finding.

1. *Concept Uncertainty*

Propositions about generic causation use nouns, verbs, and their qualifiers to conceptualize and categorize events in the world. The proposition "inhaling air containing high concentrations of benzene can cause leukemia in people" uses numerous concepts to identify and relate two types of events: certain inhalation events and the development of leukemia. Whenever we select particular concepts to use in identifying and describing events, we place a conceptual structure on the world and limit the way in which we conceptualize those events. For example, a human life is immensely complicated, with genetic, developmental, environmental, psychological, and social factors. The proposition above refers only to certain inhalation events and a certain health outcome. But the same events could be described in an infinite

number of ways. The event described as “inhaling benzene” might involve inhaling hundreds of other chemicals as well. We can also think of it as an “exposure-to-benzene event,” not just an inhalation event, and thereby classify it together with ingestions of benzene, dermal contacts with benzene, and other inhalations of benzene. The same time-slice of a person’s life might be described not as a benzene exposure event at all, but (for example) as an event of being within one hundred meters of a high-voltage alternating current transmission line.

Scientists call descriptive concepts “variables,” which they use to classify objects or events for purposes of scientific study. A variable might be a property that varies from individual to individual, such as height or weight. The primary logical function of a variable or concept is to provide a set of categories with which to classify individual objects or events. The variable “color” might be given the classification categories of “red,” “blue,” “green,” and so forth. A scientist conducting a study gives each individual subject of the study a “score,” reflecting the category in which the individual is classified. The scores of all the subjects in the study constitute the “data” for that study. There are very few constraints on inventing variables and classification categories. The primary criterion is pragmatic: they should produce measurements (data) that will be useful for some purpose.

Suppose, for example, there is an occupational health study of benzene exposure and leukemia. One variable in the study might be benzene inhalation. In order to gather exposure data, air samples would be taken in the workplace and their benzene content analyzed. Each air sample would receive a score on the variable “benzene concentration.” Depending upon the goals of the study, this variable might consist of classification categories in three different types.³⁴ Nominal or qualitative categories are merely distinguishable from each other (air samples might be scored simply as testing positive for benzene or not). Ordinal or comparative categories have a ranking among themselves, often based on degree of the property being measured—for example, “contains less benzene than the reference sample,” “contains the same amount of benzene as the reference sample,” and “contains more benzene than the reference sample.” Scalar or completely quantitative categories are not only distinguishable and rank-

³⁴ For discussions of types of measurement variables, see EDWIN E. GHISELLI ET AL., *MEASUREMENT THEORY FOR THE BEHAVIORAL SCIENCES* (1981); HERMAN J. LOETHER & DONALD G. MCTAVISH, *DESCRIPTIVE AND INFERENTIAL STATISTICS: AN INTRODUCTION* 16-24 (4th ed. 1993).

ordered, but also related by a unit of measure, such as "contains one part benzene per million parts of air."

Investigators are free to select the variables and design the classification categories of those variables. But once investigators select the study variables and the category structures of those variables, and once they gather data using those variables, any causal conclusions drawn from the study are open to question about whether different variables or category structures would have produced different results. The uncertainty created by the selection of variables is "concept uncertainty." Concept uncertainty is the potential for error created by conceptualizing the world in one way and not others. It is the uncertainty created by using any particular concepts to describe the world. The selection of variables certainly results in a lack of knowledge about variables not studied, but can also cause inferential error about those variables that are studied.³⁵ Selection of variables can be a source of inferential error in nonmonotonic inferences.

Scientists often anticipate and try to minimize unnecessary concept uncertainty when they design the variables they will use. For example, in the benzene occupational health example, there are advantages to using scalar variables, instead of nominal or ordinal variables, to gather air data. First, scalar data provides more information and warrants more conclusions. Scalar data will warrant both nominal and ordinal conclusions, but nominal or ordinal data will not warrant the kind of quantitative conclusion that scalar data will warrant. Second, the state of knowledge about benzene is at a point where scalar data are needed. When the first data are gathered about the potential toxicity of an environmental agent, qualitative case reports about acute incidents may be the only kind of information available. This state of knowledge makes purely qualitative information valuable. As the knowledge base grows, however, only more informative data make significant contributions to it. Third, the technology for generating scalar analyses of benzene air concentrations has improved over time to the point where the cost of gathering scalar data may be outweighed by the benefits of doing so.³⁶ Even with lower costs of

³⁵ See *infra* text accompanying note 48.

³⁶ This may be true about many studies of generic causation conducted within wealthy nations, but each decision to undertake a study to resolve uncertainty occurs within a particular socioeconomic context. The same studies might not be undertaken in a developing country, where scarce economic or occupational health resources might be put to better use. Fortunately, information about generic toxicity is often applicable to many circumstances around the globe, and such studies do not need to be replicated within each culture. Behavioral studies, on the other hand, might be far more culture specific.

Another example where a study might not be economically feasible, even in a

generating data, however, those costs must be weighed against the expected benefits from a particular study design. The kinds and degrees of uncertainty expected to be inherent in the study results must be weighed in that decision. However, it is clear that once data have been gathered employing certain variables, there are limits on the kinds of conclusions that the study can warrant.

2. Measurement Uncertainty

Measurement is the process of classifying individual objects or events into the categories of a variable, and it generates the data for a scientific study. Measuring creates the potential for misclassification. Measurement error can lead to inaccurate conclusions, not only about the particular individual that is misclassified, but also about the groups of individuals under study. Misclassifying an event as “inhalation of benzene” can be viewed either as a mistake in measurement or a mistake in description. From a linguistic standpoint, measurement error is simply an error in predication or description. From a scientific perspective, mistakes in describing a particular object or event are often misclassification or measurement errors.³⁷

When the goal is to reach warranted generic conclusions about the groups of individuals under study, it is often epistemically useful to classify measurement uncertainty into reliability problems and validity problems. Such problems manifest themselves in either random or systemic error. A measurement process (as well as the resulting data) is said to be “unreliable” to the extent that repeated measurements of the same object or event by the same measurement process would yield inconsistent results in a random fashion.³⁸ If we were to measure or classify the same individual repeatedly, using an unreliable process, we would classify the individual into different categories of the variable, but in a random pattern. For example, analyses of

wealthy nation, is in the context of a private lawsuit. If plaintiffs would have to finance a study that has variables tailored to the particular lawsuit, the private parties might not have the resources to do so, even assuming that such a study would be methodologically feasible. Hence, it is frequently a question in toxic tort or products liability litigation whether the studies available in the public domain are sufficient to warrant factfinding in a particular case. *See, e.g., Gen. Elec. Co. v. Joiner*, 522 U.S. 136 (1997) (upholding the district court’s exclusion from evidence of expert testimony based on animal and epidemiologic studies, where issue of fact was causation in a human being).

³⁷ Not all descriptive inaccuracies are measurement errors, however. Propositions that inaccurately describe a generic causal relation may have sources of error other than measurement error. Sufficient measurement accuracy is a necessary, but not sufficient, condition for warranting descriptive accuracy.

³⁸ *See* EDWARD G. CARMINES & RICHARD A. ZELLER, *RELIABILITY AND VALIDITY ASSESSMENT* 13 (1979); GHISELLI ET AL., *supra* note 34, at 184, 191.

the same air sample by a certain laboratory process might yield different results for the concentration of benzene in parts per million, but all those results might fall in a random pattern and largely within one percent of the mean value. For many measurement processes, especially in the physical sciences, we can conduct reliability studies to determine distributions of error under different sets of circumstances. In the behavioral sciences, however, the feasibility of retesting the same individual is especially difficult, since the measurement process itself might change the behavior of the subject individuals. Retaking standardized exams, for example, might produce higher scores simply because taking the exam educates the individual in exam-taking skills. Despite such methodological problems with certain types of subjects, scientists still employ a fairly clear notion of measurement unreliability, in which the source of error is due to random variations in the measurement process itself.

It may be epistemically useful to identify the extent of random variability within measurement data. If it is possible to take a large number of repeated measurements of the same, unchanged individual, we might find that the data vary in a random pattern that suggests many minor causes at work producing the results. It may be possible to reduce this "noise" through changing the technology or measurement methods. If we can reduce the range of random variation in measurements, we increase the "precision" of the measurements.³⁹ Another method for reducing the random error in inferences based on measurements is to take many measurements and use statistical techniques to derive a mean value from the data. Under certain conditions, statistical theory warrants ascribing less uncertainty to the mean than to a single measurement.⁴⁰ Therefore, this combination of techniques for increasing measurement precision and reducing measurement uncertainty may warrant a conclusion about the "true measurement value" that has less uncertainty than a conclusion based simply on a single act of measurement. This approach to reducing measurement uncertainty is warranted if the measurement error varies randomly.

Even if we develop a perfectly reliable measurement process that would yield the same classification for the same individual every time the process is repeated, however, we could still have uncertainty about measurement validity. A measurement process is valid to the extent that it measures exactly what we think it

³⁹ See, e.g., Theodore Peters & James O. Westgard, *Evaluation of Methods*, in TEXTBOOK OF CLINICAL CHEMISTRY 410, 412 (Norbert W. Tietz ed., 1986).

⁴⁰ See, e.g., DAVID FREEDMAN ET AL., STATISTICS 90-97 (2d ed. 1991).

measures.⁴¹ Measurement invalidity occurs when an individual is placed in the wrong category of the variable, even if this is done consistently. Validity problems are external questions about the adequacy of our measurement process relative to alternative measures of the same variable. If, for example, there is a "criterion method" or "reference method" that can serve as a standard for evaluating certain measurements (such as measuring benzene in air), then a new technology or process for measuring the same variable should be tested for validity against the criterion method.⁴² To the extent that we obtain inconsistent results from two different but reliable methods when they are applied to the same individuals, we have identified a validity problem with one method or the other. While reliability is a matter of internal consistency (using the same method, same subjects), validity is a matter of external consistency (different methods, same subjects). Both kinds of measurement uncertainty are about the possible errors introduced by the measurement process, not about true variations in the individuals being measured.

Two different measurement methods that are used to classify the same individuals on the same variable may yield systematically different results. For example, one method may overestimate the benzene concentrations, compared to the other method. Known systematic error allows a different kind of remedy than random error does. If the data have a known bias in some direction, one remedy might be to adjust for that bias in reaching conclusions based on those results. If, for example, we know that one measurement process tends to overestimate by ten percent, we might reduce the uncertainty in our inferences by discounting for that systematic error. Watches that gain a minute every hour can be useful in telling time, provided the user employs a suitable calculation to adjust for this bias. Of course, there may also be technological remedies that are feasible and economically worthwhile.

For a theory of uncertainty about conclusions that are based on a measurement process, it would be useful to distinguish reliability problems from validity problems, and random inconsistencies from systematic inconsistencies. The epistemic implications and remedies probably differ depending upon the type of measurement uncertainty involved. Scientists generally take great care to identify the kinds and sources of measurement error, to reduce the potential for measurement error to tolerable

⁴¹ See CARMINES & ZELLER, *supra* note 38, at 12; GHISELLI ET AL., *supra* note 34, at 266.

⁴² See, e.g., Peters & Westgard, *supra* note 39, at 412.

levels, and to characterize the extent of residual measurement uncertainty within a set of data. Without any appreciation of the extent of measurement uncertainty in the data, we cannot appreciate the epistemic constraints on the inferences warranted by that data.

3. Calculation Uncertainty

When scientists record data, manipulate symbols within data bases, calculate statistics from data, and so forth, there is the potential for performing the procedure incorrectly. Such mistakes can produce errors in the conclusions that are drawn from the data. This source of uncertainty is all too familiar from elementary school drills in arithmetic and spelling, and is so well appreciated that it might be overlooked as a distinct type of uncertainty. But conducting even basic arithmetic operations can provide an illustration of how we use theories of uncertainty to help warrant conclusions. We devise theories of uncertainty, for example, when we try to balance a checkbook and resolve an inconsistency between our calculated balance and the bank's calculated balance. Even if we and the bank have entered all check amounts correctly, with no measurement errors, we or the bank might have made a calculation error in adding or subtracting. The complexity of the theory of uncertainty we use to guide our investigation into the source of the inconsistency depends upon the amounts of money at stake, the number of withdrawal or deposit items involved, and the amount of time we wish to devote to the task. We might decide that it is highly unlikely that the bank's computer made a purely computational error and check only that withdrawal and deposit items are in agreement. Or we could, if we wish, check every aspect of our and the bank's calculations. Of course, we might also decide at the outset that minor inconsistencies of a few cents are not worth investigating at all. If we accept the bank's balance as our own, we eliminate the need for theories of uncertainty altogether. But if the inconsistency is not acceptable, then the hunt for its causes will be more efficient if it is guided by a theory of uncertainty.

4. Sampling Uncertainty

Scientific data record actual measurements taken on particular objects or events. We often want to generalize beyond the past measurements, however, and warrant generic conclusions about objects or events as yet unmeasured. Scientists distinguish between the sample (the individuals measured or the data gathered from measuring them) and the population (the group to

be the subject of the conclusion). Making an inference from sample data, or from the descriptive statistics summarizing that data, to conclusions about populations creates the possibility that error will be introduced because the sample does not adequately represent the population. The epistemic challenge is to warrant a conclusion about a population on the basis of evidence that is descriptive merely of a sample.

Scientists help establish the needed warrant by the methods they use to draw the sample and analyze the data. First, the sample must be drawn in such a way that we can assign a probability to drawing a sample with the same or similar statistics merely by chance. Ideally, the study would be designed so that a probability distribution could be determined for all important statistical results in samples of a given size. For example, if the population contains 60% men and 40% women, there is a certain probability of randomly drawing a sample of 200 people in which there would be 102 men and 98 women. If the population contains 58% men and 42% women, there is a probability of randomly drawing a sample of 103 people consisting of 40 men and 63 women. A sample for which such probability determinations are warranted is a "probability sample."⁴³ The warrant for assigning a probability to a type of sample is based upon the procedure used to draw the sample, the nature of the relevant variable, and the nature of the subject individuals. For example, a simple randomizing procedure for drawing the sample might alone provide sufficient warrant for assigning probabilities to sample results, but an occupational health study might also rely on the typicality of the biological processes being studied and on the representativeness of the human subjects involved. Regardless of the precise basis and reasoning, what must be warranted is the probability of obtaining samples with the relevant statistical compositions.

Second, once a sample has been drawn in such a way that probabilities can be assigned to drawing the relevant results, those probabilities and results can help warrant conclusions about the population. For example, suppose a simple random sample contains 168 people, consisting of 100 men and 68 women. Reasoning hypothetically, if the population in fact contains 10% men and 90% women, then the probability of drawing this sample is extremely low—so low that this hypothesis about the makeup of the population is probably false. The warrant for rejecting this hypothesis about the population as improbable rests on the way

⁴³ See, e.g., WILLIAM L. HAYS, *STATISTICS* 224-25 (5th ed. 1994).

the sample was drawn, on the sample results, and on probability theory itself. Scientists have invented various techniques for warranting inferences about the population on the basis of a probability sample. These techniques include hypothesis testing, significance testing using P-values, confidence intervals, and statistical power.⁴⁴ These are all techniques for characterizing the extent of the sampling uncertainty inherent in any inferences from sample to population. They help characterize the potential for error that is created by the fact that the empirical evidence is limited to sample data. Such techniques cannot eliminate the possibility of sampling error, but they can inform our judgments about populations.

From a linguistic and logical perspective, sampling uncertainty is inherent not only in propositions expressly about statistical characteristics of groups, but also in many propositions that are not explicitly quantitative at all. In English, there are many ways to express what logicians call "quantification." Examples are: "almost all birds can fly," "most parents love their children," "children need nurturing," and "some cases of leukemia in workers are caused by inhaling air containing benzene." From the standpoint of logical analysis, such propositions combine a reference to some quantity of a subject group identified by a variable, with a predication of a second variable. Words such as "some," "many," "almost all," and "all" identify some portion of a group ("birds," "parents," "children," or "cases of leukemia"), and the proposition predicates something as true of that portion, such as "being able to fly." In other words, the subject-predicate structure of the proposition classifies some quantity of individuals into the categories of two different variables. The first example asserts that almost everything classified as a bird is also able to fly, and it may also assert that there are in fact such things that are both birds and able to fly. Unless propositions with quantified classes are warranted by complete enumeration, they are susceptible to sampling uncertainty. The concept of sampling uncertainty developed in science is also useful in analyzing epistemic problems in ordinary generalizations about the world.

The scientific approach to sampling uncertainty provides an excellent example of the causal dimension of a theory of uncertainty. Sampling uncertainty poses an epistemic problem for nonmonotonic inferences from sample results to generalizations about a population. There are data available on the individuals in the sample, but there is incomplete information about all the

⁴⁴ See, e.g., DAVID H. KAYE & DAVID A. FREEDMAN, *Reference Guide on Statistics*, in REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 373-93 (1994).

individuals in the population. What helps warrant an inference about the population is a causal account of how the sample was drawn from that population. If the causal process of drawing a particular subset of individuals was sufficiently random and unbiased, it can help warrant the assignment of probabilities to types of sample results. Without such a sampling process, the remainder of the scientific reasoning is itself unwarranted. The reasoning is warranted in critical part by the causal process by which the sample was selected.

5. Mathematical Modeling Uncertainty

In ordinary language there are a variety of ways to combine two or more concepts into a single thought. Nouns name groups of individuals that are identifiable by common characteristics (“workers”), and nouns can be linked by prepositions (“workers in the company”), verbs (“workers quit the company”), or mere juxtaposition (“company workers”). Adjectives can narrow the group named by the noun to the subgroup identified by the adjective (“full-time workers”). Verbs similarly designate groups of events, such as “inhaling air containing benzene.” There are a variety of grammatical structures in English for expressing complex concepts and propositions.

Scientists express their conclusions in ordinary language. But they also want to provide scientific warrant for those conclusions, based on data generated and analyzed according to acceptable scientific methods. In order to warrant conclusions using multiple concepts, scientists must combine data that are gathered on multiple variables. A common scientific paradigm for doing so is the use of mathematical models, such as the quantitative methods of statistics for multivariate analysis. I have elsewhere discussed two important types of mathematical models: relative risk models for nominal data and linear regression models for scalar data.⁴⁵ Mathematical models characterize quantitative relationships between sets of data on two or more different variables. They also can be used to predict values for one variable based on values for other variables.

Even a very simple mathematical model will illustrate the nature of the uncertainty that use of such models introduces. If a

⁴⁵ See Walker, *Siren Songs*, *supra* note 32, at 598-608 (discussing linear regression models); Walker, *Baseline Risk*, *supra* note 32, at 651-62 (discussing relative risk model); see also Linda A. Bailey et al., *Reference Guide on Epidemiology*, in REFERENCE MANUAL ON SCIENTIFIC EVIDENCE, *supra* note 44, at 147-56 (discussing relative risk and odds ratio in epidemiologic studies); Daniel L. Rubinfeld, *Reference Guide on Multiple Regression*, in REFERENCE MANUAL ON SCIENTIFIC EVIDENCE, *supra* note 44, at 415-69.

particular watch “runs fast” but is the only available means for determining the time of day, we can still use it to determine the time if we use the right mathematical model. If, for example, the watch gains approximately two minutes per hour, then the algorithm would be to multiply two minutes by the number of hours since the watch was last set, and subtract that product from the time shown on the watch. This watch and this mathematical model, used together, might warrant predictions about time of day, as validated by other watches. This mathematical formula, however, might be the wrong model for two reasons—either because “two minutes” is not the correct amount to use in the multiplication or because the form of the model is incorrect. As an example of the first error, the correct constant might not be two minutes, but rather 105 seconds.⁴⁶ As an example of the second error, the correct formula might not be adding a constant number of seconds for each passing hour, but rather multiplying the duration of the previous hour cycle by 1.02. In the latter case, the incremental error in minutes per hour is not constant, but increases over time. When scientists use mathematical models to predict values or statistics on one variable from the value or statistics on other variables, they create a potential for error due to the mathematical models they choose to employ.

6. Causal Uncertainty

Even when sound mathematical analysis shows that there are quantitative associations among variables or data, there is a potential for error in asserting the existence of any underlying causation for this phenomenon. Two types of events can occur together without one causing the other.⁴⁷ Even when one type of event is a reliable predictor of the other, that event may have dubious explanatory value. If, for example, a standardized test score is a good predictor of success in academic courses, the standardized test does not cause that success. Any causal account of the statistical association itself, or of either variable in isolation,

⁴⁶ The formula used in the mathematical model is correct, but what is incorrect is the value assigned to a constant in that formula.

⁴⁷ In addition, even when no regular concurrence has been observed, there may be underlying causation involved, as when some third type of event counteracts or masks the otherwise causal action. See JAMES A. DAVIS, *THE LOGIC OF CAUSAL ORDER* 24-27 (1985); Bailey et al., *supra* note 45, at 157-70. For example, the complexities of human metabolism often make it very difficult to determine what is causing what within the human body. The design of a controlled experiment is intended to create a situation in which a statistically significant difference between the test group and the control group would warrant an inference of causation. Controlled laboratory experiments are not always feasible, however. The world of events is so complicated that without controlled experiments, conclusions about causal action are often difficult to warrant.

can be quite complicated and may be largely unavailable. Causal action is a means of explaining why events occur, as well as explaining the similarities and differences between those events. Causal accounts explain the data, not merely predict it.

Even if measurement, calculation, sampling, and mathematical modeling uncertainties could be eliminated, any causal interpretation would still incur causal uncertainty. A causal explanation can introduce causal error in various ways. First, there may be a real statistical association between events of type A and events of type B, but A might cause B, B might cause A, there may be interactive causality between A and B, events of type C might cause both A and B, and so on. One result of concept uncertainty—gathering data on selected variables and not others—is that variables are ignored and causal accounts may therefore become inaccurate.⁴⁸ Perhaps causal concepts themselves would be more useful if they were defined probabilistically and quantitatively, not merely qualitatively.⁴⁹ Statistical associations among sampling data may warrant a conclusion or finding about generic causation, but that conclusion has an inherent risk of causal error.

B. *Evaluation of the Theory as a Theory of Uncertainty*

This theory of uncertainty states that each proposition about generic causation that is warranted by scientific data is susceptible to six possible types of error. Each of these six types of uncertainty is logically distinct from the others, and the potential for error is cumulative from each type of uncertainty to the next. In the end, every causal conclusion probably has inherent in it a residual amount of each kind of uncertainty. In any particular instance of legal factfinding, the decision to be made by the factfinder is whether those kinds and residual degrees of uncertainty are tolerable for the purposes of the legal proceeding. The theory of uncertainty discussed here can help a factfinder to analyze which kinds of uncertainty are present in what degree, to perform a cost/benefit analysis for actions to reduce those uncertainties, and to decide which uncertainties are tolerable under the circumstances. For example, if a scientific study appears to warrant the proposition “inhalation of benzene can cause cancer in a human being,” then this theory of uncertainty explains the ways in which this proposition could still be false, despite the

⁴⁸ See *supra* text accompanying note 35.

⁴⁹ See, e.g., GLENN SHAFER, *THE ART OF CAUSAL CONJECTURE* 91-111, 299-357 (1996).

study. The theory explains the types of uncertainty that would be associated with such a finding.

The theory could also help explain which kinds of evidence are relevant, and help characterize the cumulative uncertainty in an inference based on the available evidence. The theory could explain what evidence is missing and why that type of evidence is important. It would also suggest what a complete set of scientific evidence would contain. A factfinder could use this theory of uncertainty to guide an epistemic evaluation of the available and missing evidence.

A theory of uncertainty should be useful in evaluating the warrant for inferences in a particular pragmatic context. No theory of uncertainty is likely to incorporate complete linguistic, logical, and causal accounts of any particular inference. Nevertheless, the theory discussed above draws attention to certain linguistic, logical, and causal dimensions of warrant that are generally applicable. For example, the theory provides a way of thinking about the semantics of English sentences—by analyzing predication as a process of classifying individual objects or events into the categories of a variable and by analyzing variables as having nominal, ordinal, or scalar classification categories.⁵⁰ Such a theory would prompt questions about the criteria for describing something (concept and measurement uncertainty) or for generalizing from observed instances to larger groups (as often occurs in propositions or predications). By anchoring the analysis of uncertainty in a semantics of linguistic expressions, the theory becomes useful in analyzing many findings, whether they are explicitly scientific or not. The theory ties the causal language of a finding to the kinds of uncertainty likely to underlie such a finding.

This theory of uncertainty also claims that the different types of uncertainty are logically distinct and that the potential for error is cumulative. This hypothesis about logical independence could be a fruitful heuristic, even if we were to find in a particular case that some types of uncertainty are not entirely independent of the others. Moreover, even if the types of uncertainty are independent and cumulative, the theory undermines the notion that uncertainties are simply additive, or even that they can be characterized collectively by a single quantitative measure. That is, the incremental uncertainties cannot be simply added into a single sum. In a particular case, measurement errors could have complicated effects on sampling errors or mathematical modeling errors. “Sensitivity analyses” are usually needed to determine and

⁵⁰ See *supra* Part II.A.1.

characterize the complicated interactive effects among various types of possible error within a model, and the effects of those errors on the conclusion to be warranted.⁵¹ Indeed, the theory suggests that while certain types of uncertainty may be characterized quantitatively (such as measurement reliability, sampling uncertainty, and mathematical modeling uncertainty), other types are unlikely candidates for quantitative treatment (such as concept uncertainty, some measurement validity, and causal uncertainties). The theory provides a good reason to reject the notion, popular among theorists, that all of the uncertainties inherent in a finding could be characterized by a single probability.⁵²

The theory also integrates a causal dimension by showing why causal theories are needed to help warrant inferences of generic causation. For example, part of the scientific warrant for a conclusion about causation is that the data resulted from a study in which the sampling and measurement processes were well designed and well executed. If in such a study the incidence of disease in the test group (exposed group) is higher than the incidence in the control or comparison group, and if the difference in rates is unlikely to be due to chance, this may warrant the conclusion that the exposure probably caused the increase in disease in the test group. The warrant for this conclusion, however, is strong only if the sampling design and measurement methods make any alternative causal account less probable. Therefore, the causal account of how the study was conducted is an essential part of the warrant for the conclusion. This theory of uncertainty integrates the causal, linguistic, and logical dimensions of scientific inference to produce an explanation of the possible sources of error in a finding.

This theory of uncertainty should also have explanatory uses outside the narrow class of propositions about generic causation warranted by scientific data. Many of the generalizations used in everyday inferences are not warranted by scientific studies, but are warranted, if at all, by personal experiences. Such everyday generalizations are prone to error in the same ways that scientific inferences are. Uncertainties in everyday generalizations can arise from concept definition and classification error, inaccurate summarization and nonrepresentative samples, or flawed predictive models and unwarranted causal hypotheses. In Part III

⁵¹ See, e.g., M. GRANGER MORGAN & MAX HENRION, UNCERTAINTY 39 (1990) ("*Sensitivity analysis* is the computation of the effect of changes in input values or assumptions (including boundaries and model functional form) on the outputs.").

⁵² See discussion of Schum's theory *infra* Part III.

of this Article, I discuss some examples, such as "police officers do not usually lie when testifying under oath." We often employ such generalizations to warrant inferences, and in the face of inconsistent conclusions we develop theories of uncertainty about how such a generalization might be inaccurate. The theory of uncertainty considered above should provide a useful analysis for such nonscientific generalizations as well.

This theory of uncertainty can also help resolve the problems surrounding admissibility standards for scientific opinions in judicial proceedings—the issue addressed by the U.S. Supreme Court in *Daubert*⁵³ and *Joiner*.⁵⁴ First, the theory shows a method for analyzing the uncertainties underlying scientific warrant. It summarizes part of what is meant by the adjective "scientific," and does so in a way that explains why the use of scientific method should provide a sufficient condition for admitting an expert opinion. That is, it can guide the epistemic analysis required by *Daubert*. What the theory does not answer is the pragmatic question of what level of uncertainty is tolerable for judicial purposes.⁵⁵ How to manage the risk of error in judicial factfinding is ultimately a decision of judicial management. However, a useful theory of uncertainty can help characterize the uncertainty inherent in such findings.

One critical question of judicial management is how to allocate factfinding roles between judge and jury. Rulings on the admissibility and legal sufficiency of evidence are traditional devices by which judges can take the factfinding function away from juries, in light of the proffered or admissible evidence in the particular case.⁵⁶ The theory of uncertainty above may help to

⁵³ *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579 (1993).

⁵⁴ *Gen. Elec. Co. v. Joiner*, 522 U.S. 136 (1997).

⁵⁵ In *Daubert*, the Supreme Court used the phrase "evidentiary reliability" to indicate a legal concept. *Daubert*, 509 U.S. at 589-92. That concept probably should include the pragmatic assessment that the evidence under scrutiny is "reliable enough" or "trustworthy enough" for judicial purposes, when the legal problem being addressed is admissibility. This pragmatic assessment goes much further than the scientific concept of "measurement reliability," or even "measurement validity," discussed *supra* Part II.A.2. See *id.* at 590 n.9. These scientific concepts are useful to scientists because they allow assessments of measurement techniques divorced from the pragmatic question of whether those techniques are reliable enough or valid enough for a particular purpose. By contrast, the legal concept of evidentiary reliability probably includes a judgment about whether the residual inconsistencies (if any) between inferences drawn by experts based on the proffered evidence are tolerable for purposes of the admissibility decision. On this reading, what experts are expected to provide federal judges in a *Daubert* hearing are theories of uncertainty tailored to the proffered evidence and the findings at issue in the particular legal case. The point of contention between proponents and opponents of the evidence is what the residual uncertainties are and whether they should be tolerated within the body of evidence admitted into the case.

⁵⁶ See, e.g., FLEMING JAMES, JR. ET AL., *CIVIL PROCEDURE* 357-409 (4th ed. 1992).

draw a more principled line between judge and jury, at least when it comes to the role of scientific expert opinion in factfinding. The *Daubert* and *Joiner* decisions underscore the “gatekeeping” role of the judge in requiring proffered expert opinions to pass scientific muster. To the extent that scientific warrant and the warrant for ordinary generalizations are epistemically congruent, there is no conceptual gap between the gatekeeping task and the broader role of the judge in policing the boundaries of reasonable jury deliberation. Despite the difficulties and inefficiencies in requiring an expert to explain the scientific warrant for his or her opinion first to a generalist judge and then to a jury, there may be clear epistemic benefits to requiring some expert witnesses to undergo that dual exercise. Moreover, if the theory of uncertainty discussed above can be used to draw reasonable boundaries around the scope of a *Daubert* inquiry, it can help control the trial court’s discretion in ruling on admissibility. Similar things can be said for rulings on legal sufficiency of evidence and the need for expert testimony to support a finding of fact. When assessing the epistemic merits of such legal rules, a useful theory of uncertainty should be of considerable value.

III. A THEORY OF UNCERTAINTY FOR EYEWITNESS TESTIMONY ABOUT PARTICULAR EVENTS

In his work on Wigmorean inference networks, David Schum has developed the notion of “decomposing” an inference from initial evidence to ultimate conclusion. Schum’s method of decomposition is not only general in form, but also applicable to inferences about the occurrence of specific historical events. For example, eyewitness *W* might testify at trial that some particular event *E* occurred at time *t* in the past (such as “Sacco was present at the scene at the time the crime occurred”). Witness *W* might claim to know that *E* occurred at *t* because *W* was present at the time and witnessed the event. A factfinder listening to this testimony may try to determine whether it is true that *E* occurred at *t*. The available evidence is, in part, *W*’s testimony at the trial. The inference from such evidence to a finding about *E*’s occurrence is undoubtedly nonmonotonic: the evidence is incomplete and the conclusion is not certain. The question in this Article is how an adequate theory of uncertainty could help warrant such a nonmonotonic inference. In this section, I set out Schum’s theory of decomposition, with particular attention to its features as a theory of uncertainty. In the second section, I briefly discuss several limitations of Schum’s theory, again solely from the standpoint of its treatment of uncertainty.

A. *Schum's Account of Inference Decomposition*

Schum suggests that a factfinder faced with determining whether event E actually occurred might “decompose” the ultimate inference from W’s testimony to a conclusion about E into a chain of intermediate inferences about, for example:⁵⁷

- (α) W’s *veracity* (if W is in fact remembering, at the time of trial, that E occurred at t, *whether* W is testifying in accord with this present memory *or* whether W is lying); and
- (β) W’s *memory* (if W believed at t, contemporaneously with the alleged event, that E was occurring, *whether* W’s memory at trial is in accord with that earlier belief *or* whether W’s memory is faulty); and
- (γ) W’s *observational sensitivity* and *objectivity* (*whether* W’s belief at t, contemporaneous with event E, was in accord with what was actually occurring at that time in W’s perceptual environment, *or* whether W was mistaken in what he thought he saw).

This chain of inferences is charted according to Schum’s method in Figure 1. In Schum’s terminology, the “probative force” of W’s testimony at trial with respect to the conclusion that event E occurred at t should be, at least in part, a function of W’s veracity, memory, objectivity, and observational sensitivity. At a minimum, W’s testimony is strong evidence that E occurred (in the way that W describes it) *only if* (γ) W perceived E correctly, (β) W accurately remembers his past perception, and (α) W is telling the truth about what he remembers. To the extent that any of these three conditions is not met, the probative value of W’s testimony diminishes accordingly.

Schum characterizes inference chain decomposition as making explicit the sources of uncertainty within the ultimate inference.⁵⁸ For each step identified in the chain, Schum thinks that “we have to make an assertion about what we believe provides the ground for or gives us license to take each one of these steps.”⁵⁹ Schum calls such an assertion a “generalization,” because it is an assertion “about what happens *in general*.”⁶⁰ He gives a number of

⁵⁷ See SCHUM, *supra* note 14, at 100-09; KADANE & SCHUM, *supra* note 3, at 53-60.

⁵⁸ See SCHUM, *supra* note 14, at 2, 77, 109, 138; KADANE & SCHUM, *supra* note 3, at 71, 76.

⁵⁹ SCHUM, *supra* note 14, at 81-82, 472 (stating “generalizations are required”); KADANE & SCHUM, *supra* note 3, at 88-89 (“[B]ehind every arc or link in a chain of reasoning, there must reside an appropriate generalization that licenses the probabilistic inferential step at this arc.”).

⁶⁰ SCHUM, *supra* note 14, at 82. At one point, Schum suggests that the term “generalization” is synonymous in this context with “warrant.” *Id.* at 81. This is not the meaning given to “warrant” in this Article. In my use, “warrant” is the account or

examples, such as “If a person’s senses provide evidence of an event, then this event (usually, often, frequently, etc.) has occurred.”⁶¹ A generalization contains some qualifier or “hedge” about the proposition being asserted—hedges such as “sometimes,” “frequently,” “often,” “very often,” or “usually.”⁶²

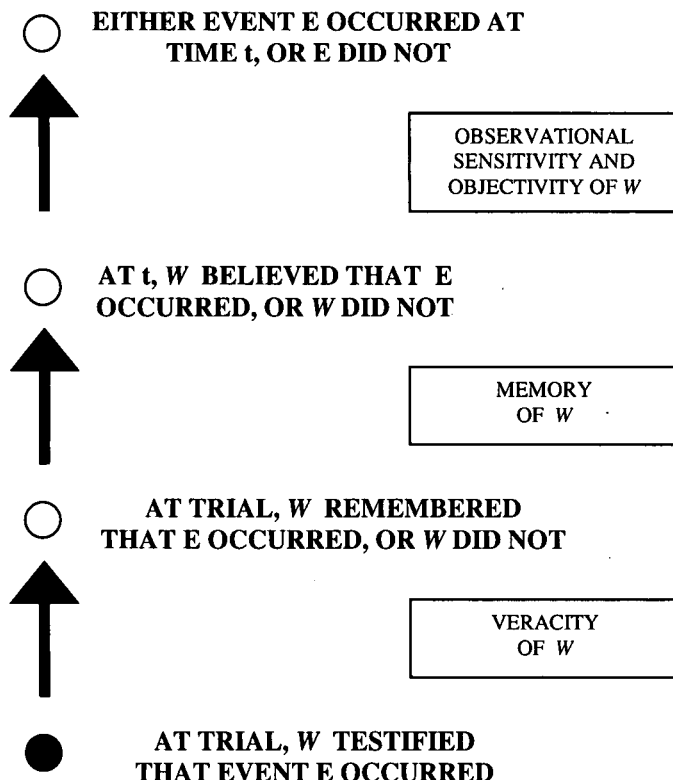


Figure 1

argument, including the evidence, that justifies finding that the conclusion is true or probably true. “Warrant” as “generalization” would merely refer to a major premise in a particular kind of warranting argument. When the word “warrant” is used throughout this Article, however, it has my broader sense, not Schum’s narrower meaning.

⁶¹ *Id.* at 102. For additional examples, see *id.* at 86-92, 101-02, 110-11.

⁶² *Id.* at 81-82, 101-02, 110-11. It is not clear whether the hedge term is about frequency (how often the asserted sequence of events occurs), warrant (how good the evidential support is), or subjective confidence (how convinced the speaker is). On the one hand, Schum states that “[w]hich hedge I choose depends upon the strength of my own belief based upon the experiences I have had in evaluating this kind of evidence.” *Id.* at 81-82. On the other hand, Schum thinks that the nature of the evidence is important. He states:

In situations in which our inferences involve replicable processes, we may have statistical or frequentist backings for these generalizations. In nonreplicable situations involving singular or unique events, we may either support or weaken a generalization on the basis of ancillary evidence resulting from a *variety* of different tests of these generalizations.

Id. at 210.

Some generalizations are derived from common sense, while others require some expertise or background knowledge of a specific situation.⁶³ Schum states, however, that “[t]here is no general standard according to which we may grade the plausibility of generalizations that we may assert.”⁶⁴

According to Schum, generalizations used in a legal context are often “asserted in fuzzy terms,” using the concept of “fuzzy” developed by Lotfi Zadeh.⁶⁵ A descriptive term is “fuzzy” in this sense if its meaning is “not crisply defined,” if the set of instances identified by the term is not determined with precision.⁶⁶ Membership in the set identified by the term is a matter of degree.⁶⁷ Examples of fuzzy sets are “elderly person,” “good eyesight,” and “concealed weapon.” Even hedge terms employed in generalizations are often fuzzy. For example, in trying to make the meaning of “usually” more precise, one person might include in that concept any relative frequency between 0.5 and 1.0, but think that the strongest examples are proportions greater than 0.7.⁶⁸ The meaning given by each user to a fuzzy term can be modeled as a “membership function” (with values between zero and 1.0). The membership function expresses the degree to which the user believes that a particular individual belongs to the “fuzzy set” identified by the term.⁶⁹ For example, I might regard Jones (age 50) as “old,” with a degree of membership of only 0.3 (that is, I am not inclined to call Jones “old”), while my teenage son clearly considers Jones to be “old” (membership = 0.95). Such “membership functions” assigned to descriptive terms are said to be “naturally subjective.”⁷⁰ Schum’s theory recognizes, therefore, that some uncertainty in a generalization is due to the vagueness and ambiguity of the terms used.

The “glue” holding arguments together consists of not only generalizations, but also “ancillary evidence.”⁷¹ The relevance of

⁶³ See *id.* at 91, 472.

⁶⁴ *Id.* at 472.

⁶⁵ *Id.* at 261-69.

⁶⁶ *Id.* at 261-63. Concerning the linguistic and logical dimensions of inferences, Schum’s analysis assumes that linguistic expressions can be found to describe the evidence accurately and to formulate the conclusions adequately. While there is some discussion of incorporating Zadeh’s fuzzy sets to capture the vagueness in set membership, Schum’s theory provides little insight into the possible sources of error below the propositional level. In order to use Schum’s method in an actual case, it would have to be supplemented by an account of linguistic meaning and logical structure that clarifies the evidence and conclusions in that case.

⁶⁷ See *id.* at 262.

⁶⁸ See *id.* at 263-64 (providing a possible membership function for “usually”).

⁶⁹ See *id.* at 262.

⁷⁰ *Id.* at 264.

⁷¹ *Id.* at 82, 263. Supporting generalizations with ancillary evidence is especially

ancillary evidence is to the applicability of a generalization to the particular instance under consideration.⁷² Suppose, for example, that a police officer named Smith testified to seeing Sacco at the crime scene, being therefore an instance of witness *W* above. Even if I believe, as a generalization, that “police officers testifying under oath usually tell the truth,” I can question whether the generalization applies in the case of Officer Smith testifying about Sacco.⁷³ Ancillary evidence relevant to Officer Smith’s veracity might include his demeanor on the witness stand, information about his motives, his actions in other contexts, and so forth.⁷⁴ Such ancillary evidence can weigh either for or against the applicability of the generalization to Officer Smith’s testimony.

Schum classifies warranting evidentiary propositions into three categories: generalizations, ancillary evidence, and directly relevant evidence. Some evidentiary propositions in a case are “directly relevant” to the ultimate inference. Evidence is “directly relevant” if there is a defensible chain of reasoning directly from the evidence to the conclusion.⁷⁵ An example is the testimony of Officer Smith, which is directly relevant to event *E*. Smith’s testimony, if credible, makes the occurrence of event *E* more probable than it would have been without the testimony.⁷⁶ By contrast, ancillary evidence is “indirectly relevant,” because it addresses “the strength or weakness of links in chains of reasoning set up by directly relevant evidence.”⁷⁷ Ancillary evidence becomes relevant if the probative value of the directly relevant evidence is questionable.⁷⁸ When the applicability of a generalization to a particular instance is questionable, the ancillary

important in factfinding about unique events. As Kadane and Schum state:

The basis for any defensible probability assessments for singular or unique events . . . is given by the strength of the *generalization* we can assert in defense of a link together with *ancillary evidence* that either supports or weakens the generalization’s being applicable in the inference at hand. Absent any ancillary evidence, our inference at a link would be based only upon an unsupported generalization. A chain of reasoning based only on unsupported generalizations cannot be very strong . . .

KADANE & SCHUM, *supra* note 3, at 87.

⁷² See SCHUM, *supra* note 14, at 71, 82, 112-14, 187-92; KADANE & SCHUM, *supra* note 3, at 85-88.

⁷³ See SCHUM, *supra* note 14, at 82.

⁷⁴ See *id.* at 249-50.

⁷⁵ KADANE & SCHUM, *supra* note 3, at 51.

⁷⁶ See FED. R. EVID. 401; KADANE & SCHUM, *supra* note 3, at 50-52.

⁷⁷ SCHUM, *supra* note 14, at 117; see also KADANE & SCHUM, *supra* note 3, at 52-53.

⁷⁸ See SCHUM, *supra* note 14, at 112, 157-60, 187-91, 207; KADANE & SCHUM, *supra* note 3, at 52-53, 85-88.

evidence might bring in a variety of indirectly relevant factors, such as Officer Smith's demeanor and motive.⁷⁹

An important aspect of Schum's theory of uncertainty is his attempt to quantify the probative value of directly relevant evidence for an inference, and thereby characterize the degree of uncertainty associated with that inference. Although there are various theories for analyzing probative "force,"⁸⁰ Schum devotes his major attention to the likelihood ratio derived from Bayes' Theorem.⁸¹ I have examined the details of this use of likelihood ratios elsewhere,⁸² and my purpose here is merely to assess Schum's account as a theory of uncertainty. If proposition **E*** describes the evidence **E*** or its content (in the case of testimony) and the proposition **E** describes event **E**, then the likelihood that **E*** is true given that **E** is true can be symbolized as $(\text{Prob}(\mathbf{E^*}|\mathbf{E}))$.⁸³ The likelihood ratio has as its numerator the likelihood (conditional probability) that the evidence **E*** would be true if the event at issue **E** occurred. The denominator of this ratio is the likelihood that **E*** would be true even if the event **E** did not occur $(\text{Prob}(\mathbf{E^*}|\mathbf{Not-E}))$. The ratio of these two likelihoods expresses the comparative likelihood of the evidence under these inconsistent descriptions of the world (**E** and **Not-E**). Schum's examples describe specific historical evidence and events—such as, **E*** = "witness Pelser testified at the trial that he saw Sacco at the scene of the crime," and **E** = "Sacco was at the scene of the crime."⁸⁴ According to Schum's theory, the probative force of the evidence **E*** relative to **E** can be captured by the likelihood ratio $L_{\mathbf{E^*}} = \text{Prob}(\mathbf{E^*}|\mathbf{E}) / \text{Prob}(\mathbf{E^*}|\mathbf{Not-E})$.⁸⁵ If the likelihood ratio $L_{\mathbf{E^*}}$ is greater than one, then the direct evidence increases the odds on **E** over **Not-E** (i.e., the odds that **E** is true); if $L_{\mathbf{E^*}}$ is less than one, then evidence **E*** decreases those odds; if $L_{\mathbf{E^*}}$ equals one, then **E*** leaves the odds unchanged.⁸⁶ While the likelihood ratio quantifies

⁷⁹ See SCHUM, *supra* note 14, at 210, 249-51; see also KADANE & SCHUM, *supra* note 3, at 85-88. As Schum states: "A generalization is supported to the extent that this generalization survives our best attempts to show that it is invalid in the particular instance of concern." SCHUM, *supra* note 14, at 251.

⁸⁰ See SCHUM, *supra* note 14, at 200-69; KADANE & SCHUM, *supra* note 3, at 150-55.

⁸¹ See SCHUM, *supra* note 14, at 213-22, 290-449; KADANE & SCHUM, *supra* note 3, at 121-57, 119-21.

⁸² See Vern R. Walker, *Language, Meaning, and Warrant: An Essay on the Use of Bayesian Probability Systems in Legal Factfinding*, 39 JURIMETRICS 391, 397-404 (1999).

⁸³ Formally, the likelihood ratio quantifies the change in odds on a proposition (such as the ultimate conclusion **E**) as a new proposition is taken into account (here, the evidence **E***). See SCHUM, *supra* note 14, at 215-22; KADANE & SCHUM, *supra* note 3, at 124-26.

⁸⁴ See KADANE & SCHUM, *supra* note 3, at 92, 290 chart 4.

⁸⁵ See SCHUM, *supra* note 14, at 215-22; KADANE & SCHUM, *supra* note 3, at 116-50.

⁸⁶ See KADANE & SCHUM, *supra* note 3, at 125.

the probative force of E^* for drawing an inference to E , it can be seen at the same time as reflecting the uncertainty associated with making such an inference.

B. *Assessment of Schum's Theory of Uncertainty*

Schum's evidence charting appears to be a useful means of analyzing chains of nonmonotonic inferences and organizing all of the relevant evidence. When inferences are decomposed into chains, the warrant for each step of the chain can be determined and the ways in which uncertainty enters the inference can be made more explicit. Schum's method forces us to determine the function of each item of evidence in warranting the findings. There are difficulties, however, with the adequacy of likelihood ratios to quantify the probative value of direct evidence. In particular, it is doubtful that a likelihood ratio provides any explanation for the possible sources of error in an inference or adequately characterizes the kinds of uncertainty associated with the inference.

Schum's method distinguishes three roles for propositions in warranting an inferential step: they can function as generalizations, descriptions of ancillary evidence, or descriptions or content of the direct evidence.⁸⁷ The use of any particular generalization to help warrant an inference could introduce error in at least two ways. The first way is if the warrant for the generalization itself is weak. To the extent that the generalization is unwarranted or weakly warranted, the generalization cannot perform a warranting role. Schum provides no theory of warrant for generalizations.⁸⁸ Without an analysis of the warrant for generalizations, however, Schum's method leaves a major source of uncertainty unexplained.⁸⁹ Moreover, the degree of warrant for the generalization used surely influences the probative value of the evidence for the inference. If a generalization has very little basis, any inference made in reliance on that generalization should have diminished warrant as well. Part of Schum's theory is that the likelihood ratio L_{E^*} is an adequate measure of the probative force

⁸⁷ See SCHUM, *supra* note 14, at 263; cf. KADANE & SCHUM, *supra* note 3, at 82 (stating that even tangible evidence, such as bullets or weapons, are charted in his method as propositions, since "[n]o item of tangible evidence 'speaks for itself'").

⁸⁸ See *supra* note 64 and accompanying text.

⁸⁹ The requirement of a "hedge" within a generalization is not a sufficient answer. A hedge term can indicate the intended quantification for the proposition—for example, that the generalization might be true in "all," "many," "forty-two percent," or "at least one" of the cases. But there can still be uncertainty about whether the generalization (with its intended quantification) is true or false. Uncertainty is about the potential for error, whatever the generalization's quantification.

of E^* for E . But it is unclear how the likelihood ratio $\text{Prob}(E^*|E) / \text{Prob}(E^*|\text{Not-}E)$ should be influenced by the degree of warrant for an applicable generalization. That is, it is unclear what method the factfinder should use to “adjust” one or both likelihoods to take into account the degree of warrant for the generalization.

Perhaps my theory of uncertainty for scientific warrant, which was discussed in Part II, could provide a theory of uncertainty for Schum’s generalizations. The rationale for my theory is sufficiently general, so that it can be applied to noncausal generalizations (such as statistical generalizations about a population) and to generalizations with nonscientific warrant.⁹⁰ But adding my theory to Schum’s would cause a serious problem for Schum’s use of the likelihood ratio as an integrated measure of probative value. The theory of scientific warrant suggests that the uncertainties inherent in a generalization cannot be characterized by a single probability.⁹¹ The likelihood ratio L_E contains far too little information about the nature of the underlying uncertainties, and sheds little or no light on the reasons why the conclusion might be inaccurate. To save the sufficiency of likelihood ratios, Schum needs an account of how to quantify the warrant for generalizations and how to adjust likelihood ratios appropriately.

The second source of inferential error introduced through generalizations is the risk of applying an inappropriate generalization in a particular case. This is a possible source of error even if the generalization is itself warranted. For example, it may be a true generalization that police officers usually do not lie under oath, but should this generalization be applied to Officer Smith’s testimony?⁹² Schum confronts this problem by defining ancillary evidence as evidence that helps to warrant the application of a generalization in a particular case. By definition, however, the probative force of the ancillary evidence is not a function of event E , nor can it be captured by a likelihood ratio relative to E . Schum is candid that a major reason for creating a separate category of “ancillary” evidence is that if ancillary evidence A were directly linked to E , the Bayesian portion of his theory would require assigning values to $\text{Prob}(A|E)$ and $\text{Prob}(A|\text{Not-}E)$.⁹³ But assigning such values does not seem promising, especially since there may be very little causal relationship between event E and

⁹⁰ See *supra* Part II.B.

⁹¹ See *supra* note 52 and accompanying text.

⁹² Perhaps Schum merely assumes what logicians call a “direct inference” from the statistical generalization to the probability about a specific case. This is, however, a complicated inference when it comes to providing warrant. See Walker, *supra* note 4, at 279-307.

⁹³ See SCHUM, *supra* note 14, at 188-90.

the evidence **A**.⁹⁴ Schum concludes that the probative force of ancillary evidence must be graded into *ordinal* categories.⁹⁵ This makes it difficult to mathematically integrate the probative value of ancillary evidence into the likelihood ratio for direct evidence. To capture the probative value of ancillary evidence in the likelihood ratio for the direct evidence, one would have to “adjust” the constituent likelihoods based on the ordinal probative value of the ancillary evidence.

Making the likelihoods subjective, as Schum does, acknowledges the problem without solving it.⁹⁶ As measures merely of the factfinder’s degree of confidence, subjective likelihoods can be freely adjusted to reflect the factfinder’s assessment of the ancillary evidence and the factfinder’s degree of confidence in the generalizations employed. But there is nothing to warrant the subjective likelihood assignments, except the intuition and authority of the person making those assignments. Theories of uncertainty are useful if they help explain the warrant for a nonmonotonic inference by explaining the possible sources of error in such inferences. The warrant for an inference, or for an assessment of the probative value of evidence, consists of the reasons why the conclusion is probably true. Relying on the mere fact that someone thinks it is probably true threatens to short-circuit the hunt for warrant.

Schum’s theory retreats too quickly to the shelter of subjective likelihood assignments, instead of probing further for the warrant for those assignments. This means that the theory not only falls short, but also undermines the likelihood ratio as a promising source of warrant. The likelihood ratio does seem to play an important epistemic role in assessing the probative value of direct evidence. The value of the likelihood ratio is that it can be used to compare two causal relationships: (1) the causal link from event **E** to the occurrence of the evidence **E***, and (2) the causal link from the world without event **E** to the occurrence of **E***. The question whether **E*** is more likely to occur if **E** occurred than if **E** did not occur seems highly relevant to the probative value of **E*** for **E**.⁹⁷ A promising direction for Schum’s theory would be to require any value assigned to L_{E^*} to be warranted by a

⁹⁴ See *id.* at 150-55.

⁹⁵ See *id.* at 187-92, 245-51, 272-74, 301-06.

⁹⁶ Schum considers probabilities, including likelihoods (conditional probabilities), to be subjective in nature. See *id.* at 40, 52-54, 209-10, 263-65; KADANE & SCHUM, *supra* note 3, at 24, 117-18, 267-68.

⁹⁷ On the use of reasonably well-specified and reasonably stable causal systems to warrant assigning probabilities to sequences of events, see Walker, *supra* note 4, at 279-81, 292-97.

causal account. The likelihood ratio for direct evidence is potentially useful if it is warranted causally, but a precipitous retreat to subjective probabilities hobbles that potential.

A final problem for Schum's theory is that every inferential step could have multiple warranting generalizations. For example, consider the possible generalizations for concluding that Officer Smith is testifying truthfully. One generalization might be that "police officers usually do not lie when testifying under oath." But a factfinder might also be tempted to use the generalization that "men older than forty do not lie under oath," or even that "blue-eyed people never lie." Yet all three generalizations, and an infinite number of others, could be invoked to warrant an inference about Officer Smith's veracity. By what method should a factfinder prioritize the possible generalizations by their degrees of warrant? When multiple generalizations are possible and seem to have different degrees of warrant, what is the method for rank ordering them with respect to their warrant? And how are different degrees of warrant for multiple generalizations to be combined into a single value? This problem is critical for any inference, because even a seemingly simple inference can be supported by multiple generalizations.⁹⁸ The question is whether there is any epistemically useful method of doing so. It seems likely that any successful attempt at solving this problem would be accompanied by a theory of uncertainty to explain the possible sources of error.

CONCLUSION

Legal factfinding routinely relies on nonmonotonic inferences and on the incomplete evidence underlying them. Theories of uncertainty can help warrant findings of fact in such situations. A theory of uncertainty explains the extent to which the available evidence warrants a particular conclusion, what kinds and degrees of uncertainty are associated with that conclusion, and how additional evidence might affect the warrant for that inference. It explains the possible sources of error in nonmonotonic inferences. Unlike a theory of validity in traditional logic, which explains why truth is necessarily preserved from certain premises to certain

⁹⁸ This is a different problem than the regress that is possible because any inference can be decomposed further, *see* KADANE & SCHUM, *supra* note 3, at 52, 246, and because more generalizations are needed to support the inferential use of ancillary evidence, *see* SCHUM, *supra* note 14, at 187-92. The problem of adjusting probative force to reflect multiple generalizations must be faced in order to apply the method to any inference at all. In this respect, it resembles the problem of combining multiple items of evidence, to which Schum devotes considerable attention. *See id.* at 366-449.

conclusions, a theory of uncertainty explains the possibility of falsehood and the risk of error. If we understand how the evidentiary premises could be true but the proposed conclusion false, then we understand the potential for error associated with making that finding on the basis of the legally available evidence.

A theory of uncertainty has an explanatory function roughly parallel to the role played by a theory of causation in an empirical science. Just as a good causal theory should increase the accuracy of predictions about events, a good theory of uncertainty should increase the accuracy of findings of fact, at least in the long run. It does this by clarifying linguistic meaning, by identifying logical relationships among apparently warranted but inconsistent propositions, and by providing causal explanations linking the events at issue to the evidentiary events. Linguistic, logical, and causal accounts of the legally available evidence and the possible findings are integrated into a single theory about which inferences are more warranted than others.

Understanding the uncertainty or potential for error in an inference might lead to effective measures to reduce that uncertainty and increase the likelihood of more accurate factfinding. A good theory might help us decide which uncertainties are more critical than others and how productive it might be to reduce some uncertainties rather than others. In some situations, the risk of factfinding error might be reduced through better reasoning or obtaining new information. Whether factfinding accuracy actually increases even in the long run, however, depends upon many factors, only one of which is the adequacy of the adopted theories of uncertainty. In other situations, a party or the factfinder might decide that the residual uncertainties for a finding are tolerable and that the cost of further analysis or evidence gathering outweighs the expected benefits. Instrumental to such pragmatic decisions is an understanding of the nature of the risk of error—an understanding based on a theory of uncertainty.

This Article also examined two particular theories of uncertainty. The first theory of uncertainty, presented in Part II, is a theory of uncertainty for scientific warrant. It is especially useful in explaining the possible sources of error in conclusions about generic causation, particularly when those conclusions are warranted by scientific evidence. This same theory of uncertainty should be applicable to statistical (noncausal) generalizations and to generalizations warranted by ordinary experience. The reasoning behind the theory is general in applicability. According to this theory, for example, any proposition about generic

causation has inherent in it uncertainties introduced by the concepts employed, the classifications or measurements performed, the calculations made, the samples selected and analyzed, the mathematical models used, and the causal interpretations imposed. If any of these types of uncertainty is unacceptably high, there may be techniques for reducing the risk of error to acceptable levels. The appropriate technique depends upon the type of uncertainty to be reduced. Moreover, there are various techniques for characterizing the degree of residual uncertainty of many of these types.

The second theory discussed in this Article is Schum's method of evidence charting and use of likelihood ratios to measure the probative force of an inference. Schum's theory, discussed in Part III, is especially valuable in analyzing the sources of error in inferences about unique historical events. Schum classifies warranting evidentiary propositions into three categories (generalizations, ancillary evidence, and directly relevant evidence), and uses likelihood ratios to quantify the probative force of an inference based on directly relevant evidence. This theory has advantages, but also contains several weaknesses. A number of these weaknesses can be remedied, however, and Schum's methods for analyzing such inferences might be improved. For example, if the theory of uncertainty for scientific warrant discussed in Part II were used to supplement Schum's theory of generalizations, then the factfinder could analyze the warrant for the generalizations identified through Schum's evidence charting. The theory of uncertainty for scientific warrant, however, undermines Schum's use of the likelihood ratio as a single measure of probative value.

These general remarks about the nature of a theory of uncertainty, together with the discussions of two leading examples of such theories, map out a promising approach to warranting nonmonotonic inferences. What is needed to warrant such inferences is an appreciation of the epistemic limits of what we know, an understanding of the nature of what we do not know, and a theory about how the information that is missing is likely to affect the accuracy of our findings.