Luc Bovens

'Interview' In: *Probability and Statistics: 5 Questions.* Edited by Vincent Hendricks and Alan Hajek. Automatic Press, forthcoming.

1. Why were you initially drawn to probability theory and/or statistics?

I first developed an interest in statistical methods in the late 70's and early 80's during my undergraduate studies in Social Sciences at the Catholic University of Leuven. During my graduate studies at the University of Minnesota, I was introduced to modelling in microeconomics by Leonid Hurwicz in a series of graduate courses. Graham Oddie and Stephen Leeds at the University of Colorado at Boulder sparked my interest in Bayesianism in philosophy of science in the mid-90s. The University of Colorado had a strong programme in probabilistic modelling—so I lined up my book case with Sheldon Ross's books and pretended to be a student in a sequence of courses in the mathematics and applied-mathematics departments.

Beyond this, much was due to chance—as is fitting considering the theme of this volume.

In the summer of 1997 James Hawthorne and I worked on the lottery paradox and the preface paradox which led to our joint *Mind* (1999) article. He steered me towards the last chapter in Judea Pearl's *Probabilistic Reasoning in Intelligent Systems* on the connection between logic and probability. But the graphs of Bayesian Networks in the beginning of the book drew my attention. My tenure file was in and I had some time to explore new horizons. I started reading on Bayesian Networks and played around with Bayesian-Networks software.

The following year I was on a Humboldt fellowship at the University of Konstanz. I attended a seminar by Wolfgang Spohn, who has done early work on conditional independence structures—one of the pillars of Bayesian Networks. I met Erik J. Olsson who had an interest in formal epistemology and Stephan Hartmann who had an interest in models in science. We started exploring how Bayesian Network methodology could be applied to problems in the coherence theory of justification and in philosophy of science. This led to two joint publications (*Mind*, 2000 and *Erkenntnis*, 2002) with Erik and a series of articles culminating in a monograph (*Bayesian Epistemology*, 2003) with Stephan.

Wlodek Rabinowicz and I met in the University of Leipzig in 1999. Wlodek was thinking about the discursive dilemma in joint work with Philip Pettit and they had formulated some interesting conjectures. At the time, there had been some changes in the tenure proceedings in the University of Colorado and this raised some questions that were in the neighbourhood of Philip and Wlodek's conjectures about the discursive dilemma. Consider the following two procedures. One could take votes by the tenured faculty on the candidate's qualifications on teaching and research separately and award tenure just in case there is a majority in favour on each score. Or one could have the faculty assess in their own minds whether the candidate is qualified on both scores and decide the matter by a simple majority vote. Which procedure is the better truth-tracker, i.e. yields the fewest false positives and false negatives? This question opened up some fruitful new lines of inquiry into probabilistic approaches to judgment aggregation. (*Synthese*, 2006)

Wlodek and I knew that our curiosity was sparked by the same kind of things. Much of my later research started with two short newspaper clippings that we sent to each other. Wlodek was amused by the French response to a proposal by the Swedish delegation to set the weights of the various countries in the EU Council of Ministers proportional to the square root of their population sizes. Chirac commented that he failed to see the political significance of the square root. I sent Wlodek a clipping about an interesting hats puzzle that was first formulated by Todd Ebert and was occupying computer scientists. Hats are passed around in the dark and you have a 50-50 chance of obtaining a white or a black hat. The lights are turned on and you see the colour of other people's hats, but not your own. You are asked to call the colour of your own hat. If at least one person calls it correctly and nobody does so incorrectly, then, allowing for passes, the group receives a prize. What strategy should the group adopt?

As to the first clipping, I told Stephan about the Swedish delegation and the square root rule when taking a break from Bayesian Epistemology over a restaurant lunch. In fact, the square root rule for the representation in a Federal Assembly goes back to a proposal made by Penrose, in work on voting power in the 40's, but the Swedish delegation or we knew nothing about this. Our ignorance sent us down a different track. We asked ourselves what the likely implications would be of different voting weights on the welfare distribution for the various countries in the European Union. After filling up the paper table cover with calculations, we moved on to *Mathematica* for simulations, which yielded some interesting results. (*Reasoning, Rationality and Probability*, 2006 and *European Union Politics*, 2005) This was Spring 2001 and at the time there was not much literature on welfarist approaches to the evaluation of voting procedures yet.

I received a Sofja Kovalevskaja award from the vonHumboldt foundation, which permitted me to set up a research group named *Philosophy, Probability and Modelling* (PPM) in the University of Konstanz from 2002-5. Stephan Hartmann and I co-directed the group. There was a wonderful sense of synergy in this group, which centred around issues of philosophy of probability, broadly construed.

Claus Beisbart (one of the *PPM*-members) and I started working on joint projects in voting theory that involve probabilistic modelling. We have been working on the US Electoral College—first on Colorado's Amendment 36 (*Public Choice*, 2007) and now on the proposal in California to move to a district-by-district procedure. In more theoretical work, we determine optimal weights for voters in a federal assembly on egalitarian and utilitarian grounds and show how these results generalise to results in the voting power literature. (*Social Choice and Welfare*, 2007) Rousseau's remarks on factions inspired a probabilistic model that provides a novel justification for less-than-full proportionality in federal assemblies. (*Analysis*, 2007) We are currently exploring a measure of *a posteriori* voting power—i.e. of the influence that voters have on the outcome taking into consideration actual voting profiles and interactions between voters—by means of Bayesian Networks methodology. (*Texts in Logic and Games*, 2008)

As to the clipping on the hats puzzle, Wlodek responded with a variant of the puzzle that seemed to indicate that it was possible to make a Dutch Book against a group of players who are making independent decisions in the interest of the group. It only dawned on us much later that this is a strategic decision-making problem. A simple game-theoretical argument shows that rational players would actually evade the Dutch Book. This seemed like a story of paradox gained, paradox lost. But en route there were a wide array of lessons to be learned. The puzzle is relevant to Ramsey's analysis of degrees of belief as fair betting rates (*Foundations of the Formal Sciences VI*, 2009), to Dutch Book arguments for the Sleeping Beauty, and to the strategic voting literature. (*Synthese*, 2009) It also brings out curious features of doubly symmetric games with applications to the tragedy of the commons, which I am currently exploring with Maurice Koster and Ines Lindner.

I have come full circle in my most recent work and returned to problems in statistics for the social sciences. I was struck by the plight of African migrants who clung to tuna nets in the Mediterranean struggling for their lives, with Malta and Libya quibbling about who should rescue them. Chechnyan asylum seekers have a high chance of acceptance in Sweden and a close to zero chance in Slovakia. It seemed that what was needed was an expert system for asylum seekers to determine where they should file a claim. To avoid asylum shopping, the EU is actually trying to bring more unity in their dealings with asylum seekers through the 1999 Tampere Agreement. What we have available is yearly UNHCR data detailing the proportion of asylum seekers from various countries of origin that have been granted refugee status by each EU host country. In joint work with Paresh Kathrani, we are trying to assess whether there is a tendency toward or away from a more unified policy, as reflected in the acceptance rates of the EU host countries, on the basis of this data. There are various techniques for describing interrelations in multivariate data in the social sciences. The design and application of these techniques yield interesting conceptual questions.

Many thanks to everyone with whom I have had the good fortune to work. It has been a joy. Many thanks also to the Choice Group at the LSE which has been a source of inspiration. What I like about probabilistic models is that they often lead to surprising results. This is also what attracts me to doing collaborative work. You never know beforehand what will happen when you hook up two or more cognitive systems, especially when they have been trained differently. And then it does make the journey less lonesome.

2. What is distinctive about your work in the foundations of probability or its applications?

I distinguish here between the following three strands in my work. First, I have an interest in epistemic paradoxes and some of this work touches on probability theory. Second, I use probability theory and probabilistic networks in addressing issues in epistemology and philosophy of science. And third, my work is of an interdisciplinary nature. I make use of probabilistic techniques in addressing normative questions in social and political theory—and in particular in voting theory.

What is distinctive about my work on epistemic paradoxes? Maybe it is this: I don't tend to rush to dissolve a paradox, but rather, my tendency is to cherish a paradox and to see to what use it can be put and how it connects to a wide range of issues. In the work on the lottery and the preface paradoxes, we try to use these paradoxes to spell out a relationship between quantitative and qualitative doxastic notions, i.e. between degrees of belief and plain belief. I tried my hand at the connection between the surprise exam paradox and backward induction arguments in game theory—but am dissatisfied with the results and would like to pick up the topic again some day. In the work that was sparked by the hats puzzle, we explore its connection to degrees of beliefs as betting rates, doubly symmetric games, self-locating beliefs in the Sleeping Beauty problem, the tragedy of the commons and strategic voting.

Let me now turn to epistemology. The Cartesian sceptic claims that we are never justified in believing any of the information from the external world that impinges on us through our senses. All these information items may be like Descartes' stick in the water that appears bent. In response, a coherence theorist claims that we are justified in believing at least some external-world information on grounds of its coherence. It would be quite implausible that every item in the story that we have gathered about the world would be false, considering how well this story fits together. There might be an occasional misleading item like the seemingly bent stick in the water, but if all items were like that, then how could the story come out to be so coherent? The standard challenge to the coherentist is: What could possibly be meant by information being more and less coherent? The methodology we appeal to in answering this question is distinctly interdisciplinary.

Coherence of information only matters to our degree of confidence that the information is true if the information comes from multiple independent and partially reliable sources. To model information updating with independent sources, we make use of conditional independence structures and the theory of Bayesian Networks in computer science.

Furthermore, there are various influences from economic modelling.

First, we borrow the notion of separability from consumer demand theory. We search for a set of *separable* variables that determine our degree of confidence that the information provided is true. In trying to identify a measure of coherence, we try to identify a variable that (i) is a function of the probability distribution over the propositions in the information set and (ii) is in line with our ordinary notion of coherence and (iii) increases our degree of confidence in the information provided, *ceteris paribus*, where this *ceteris paribus* clause is cashed out in terms of separability.

Second, the measurement of coherence is inspired by Atkinson's work in welfare economics. One could ask—look, what is equality good for? Well if it's good for increasing overall welfare, then we could compare levels of welfare in different societies with a fixed total income. Society *A* has an income distribution that is more equal than the income distribution of society *B*, if it is the case that is greater in *A* than in *B for any strictly concave utility function of income the level of welfare*. This criterion yields a quasi-ordering for the relation '... being no less unequal than ...' Similarly, we asked—what is coherence good for? What it is good for is that it makes us more likely to believe the story that transpires upon being informed of its constituent items by partially reliable and independent witnesses. So let us assess the actual joint posterior probability after the information is in. Now suppose that the information would have come to us in *fully coherent* format—i.e. each witness would not have provided us with one single item of the story, but with the whole story. We assess what the joint posterior probability would have been under these idealised conditions. Now construct the ratio of the actual joint posterior probability over the joint posterior probability under conditions of full coherence. If this ratio for one information set exceeds the ratio for another set *no matter how we specify the degree of reliability of the witnesses*, then the former set is more coherent than the latter. This only yields a coherence quasi-ordering, but this is how we like things to be. For sets that remain unordered by this procedure, we also lack a clear intuitive judgment whether one set is more or less coherent than the other. There have been clever counter examples to this proposal. (Meijs and Douven, *Mind*, 2005) I am somewhat nervous, but there is some room to wriggle and I am not convinced that our wriggling has been unsatisfactory. (Bovens and Hartmann, *Mind*, 2005)

And third, there is the influence of social choice theory. We construct an impossibility theorem, Arrow-style. There is the surprising consequence that it is impossible to construct a coherence ordering over information sets such that more coherent sets induce a greater boost in our degree-of-confidence upon being informed by independent and partially reliable sources, *ceteris paribus*. As I mentioned, one escape route is to settle for a quasi-ordering and this may be precisely what we want. Another escape route is to construct vectors of separable coherence measures but there is no straightforward interpretation of these vectors that maps onto our ordinary notion of coherence.

In philosophy of science, we introduce a model of unreliable instruments. The model is characterised by two parameters, viz. a parameter that measures the prior probability of the reliability of the instrument and a parameter that measures the probability of obtaining positive test results if the instrument is unreliable. We can then use this simple model to test the variety-of-evidence thesis, the Duhem-Quine thesis and the effectiveness of calibration procedures. For instance, is it true that in hypothesis testing, we may be more confident upon receiving confirming test results from multiple independent instruments rather than from a single instrument? Not quite—it depends on the values of the parameters. For certain parameter values, consistent test results may increase our confidence in the reliability of a single test instrument and this trumps the value of consistent results from multiple independent test instruments. Using a combination of Bayesian network methodology and *Mathematica* software we construct contour lines to determine the parameter values for which the variety-of-evidence thesis (under a particular interpretation) holds and for which the thesis does not hold. What is distinctive about our work is that we import this standard methodology of scientific modelling to take on philosophical questions. This permits us to provide nuanced answers—the truth of certain claims is contingent on the values of a set of relevant parameters.

In voting theory, Banzhaf voting power is the chance that a voter is pivotal under the Bernoulli model. This is the chance that had she voted differently, then the vote would have gone differently, assuming that there is an equal chance that voters vote for or against a proposal and assuming that votes are cast independently. Our work diverges from this standard approach in two respects. First, our approach is welfarist. The implementation of a proposal may affect different people's welfare in different ways. The impact on a person's welfare of a proposal is a random variable and a proposal is represented as a vector of random variables—one for each person (or subgroup) affected. This then permits us to assess different voting rules relative to various desiderata, e.g. one may wish to institute a voting rule that maximises expected welfare or that equalises expected welfare. This approach is particularly useful for assessing the representation of states in a federal assembly such as the Council of Ministers in the EU and the Electoral College in the US. Second, our approach is cautiously *a posteriori*, i.e. we explore to what extent one can take into account information about actual voting patterns and various types of welfare dependencies in the evaluation of voting rules. In my most recent work with Claus Beisbart, we consider how the presence of opinion leaders affects the actual influence of the voters by means of the Balke-Pearl theory of counterfactuals in the framework of Bayesian Networks.

3. How do you conceive of the relationship between probability theory and/or statistics and

other disciplines?

I will take the liberty to twist this question somewhat and consider the role of probability theory and in particular, probabilistic modelling, in philosophy. At least at first sight, many philosophical questions lend themselves to a kind of methodology that is common in the sciences. Why would one want to look at philosophical problems in this vein? For the same reasons that one would use this type of methodology in the sciences. There is clarity to be gained, as one considers the following questions in model construction. What are the relevant variables? What are the relations between the variables? Are there stochastic dependencies between the variables? How can the variables be measured? The determining factors pull in different directions and intuitions often give out in these matters. What factors win out for what values of the relevant parameters?

One of the attractions of probabilistic modelling is that one often has no clue beforehand what results will materialise and there is a genuine surprise element in one's work. Let me give you two examples.

Here is a result from Bayesian epistemology. Suppose that the prior joint probability of all the propositions in set S is the same as the prior joint probability of all the propositions in set S', but that the data in S is more coherent—i.e. fits together better—than the data in S'. Then it can be shown that we would be more confident that the propositions in S are true than that the propositions in S' are true after independent and equally reliable witnesses each attest to the truth of these propositions, there being a single witness for each proposition. This is unsurprising. But now consider two information sets S and S' with positive and negative dependencies between their constituent propositions within each set and equal prior joint probabilities. Suppose that we are informed by witnesses with high reliability who each independently attest to a proposition in the respective sets. Alternatively suppose that we are informed in the same manner by witnesses with low reliability, but who are still sufficiently reliable that their testimony increases our degree of confidence. Then it is possible that in the former case – i.e. with witnesses of high reliability – our degree of confidence that all propositions in S are true is greater than our degree of confidence that all propositions in S' are true, whereas in the latter case – i.e. with witnesses of low reliability – the reverse holds. To make this more tangible, suppose that you are equally confident that all the items in information set S and that all the items in S' are true. In the first case, you receive confirmation from, say, independent first-league scientists (whom you trust very much) that each element of these sets is true. Then you would be more inclined to believe, say, S than S'. Contrast this with a scenario in which you receive confirmation from, say, independent second-league scientists (whom you trust less) that each element of these sets is true. Then you would be more inclined to believe S' than S. This me as a surprising result.

Or let me turn to a different example. We represent a proposal to be voted on in a federal assembly as a vector of utilities that indicate how the proposal affects each member in the federation. If utilities of the voters within states as well as between states are fully independent and certain other default assumptions on the probability density function hold, then utilitarianism and egalitarianism agree on a particular degree of proportionality for states within a federal assembly that is located between proportional and equal representation—viz. on the square root rule. However, the more the utilities of the voters within states are aligned, the more utilitarianism tends toward proportional representation and the more egalitarianism tends toward equal representation. Again, this strikes me as a surprising result.

These kinds of results bring an element of wonder to philosophical inquiry. At the same time, this sense of wonder is somewhat of a two-edged sword. Probabilistic modelling can be hard-going in philosophy. When you have results from simulations, the mathematician objects that she wants to see analytical results. When you achieve analytical results, then the philosopher complains that she wants to understand why results come out this way. And if you succeed in giving a good intuitive account of the result—taking away some of the surprise value—then the next question is: Did you really need to do all this work for *that*? Indeed some modelling in philosophy is just bogus—results could equally well be presented by means of an intuitive story and the modelling exercise was at best a school exercise. But this is not always the case. Sometimes intuition can do no better than indicate the way, show a general direction. For the details of the case, the model remains indispensable.

Another trap is Strawson's *dictum* that formal solutions to philosophical problems can be like giving a textbook in physiology to someone who says (with a sigh) that he wished he understood the workings of the human heart. (See *The Philosophy of Rudolph Carnap*, edited by P.A. Schilpp,1963.) Clearly, just like in the sciences, there is the caveat that one should try to construct a realistic model rather than gain simple and elegant answers at the cost of misrepresenting or shifting the problem. I sometimes do have qualms of this sort about my work. For example, our model of partially reliable instruments is contrived and some of the surprising results are no more than an artefact of the model. In such cases it's not clear what is gained in our understanding about the world. This is a problem that is not restricted to philosophy, unless it is the case that one could show that philosophical problems are such that what is essential always resists modelling. I do not think that a general argument to this effect—i.e. an argument that philosophy stands to modelling as matters of the heart stand to physiology—is forthcoming. If there are concerns of this nature, they would need to be voiced case by case.

> 4. What do you consider the most neglected topics and/or contributions in probability theory and/or statistics?

Let me tailor this question towards my own expertise – viz. the intersection between philosophy on the one hand and probability theory and statistics on the other hand.

There is an extensive interest in probabilistic causation amongst philosophers and there is lots of room for interdisciplinary work. Computer scientists contribute to questions that were once strictly in the realm of metaphysics and epistemology—such as the analysis of counterfactual statements in the work of Balke and Pearl. Philosophers contribute to the development of algorithms for causal search and engage with research in statistics on causal models. This is not surprising. Philosophers have had a long-standing and continuing interest in causation and induction at least since the publication of Hume's *Treatise*. The increase in computational power has opened up many new avenues in statistics for the social sciences. Something would be amiss if no bridges had been built between both areas of research.

What is curious is that there are two areas of research that seem geographically equally close and their respective histories are similar, but there is much less activity in bridge construction. In statistics for the social sciences, the increase of computational power has not only led to the development of causal modelling techniques, but also to the development of techniques for exploring and representing interrelations between multiple variables, such as cluster analysis, correspondence analysis, multi-dimensional scaling and latent class analysis. Why might philosophers care about such issues?

The question of classification is at the heart of the ancient debates in the theory of universals with its core question—in virtue of what are two objects tokens of the same type? Resemblance theorists worry about how to construct resemblance classes. They might find it of interest that subjective resemblance judgments of, say, colour, can be comfortably (i.e. with *stress* below a threshold value) scaled in a two-dimensional plane, whereas other types of resemblance judgments require fewer or more dimensions.

There has been a genuine explosion of multivariate methods to study interrelations between variables due to computational advances. The interpretation of these methods and, in particular, the scope and legitimacy of their applications has received too little attention. Now it's not that the statistician is like Goethe's sorcerer's apprentice and that the philosopher needs to come to her aid. But there is unexplored territory here and collaborative work between statisticians, social scientists and philosophers could prove rewarding in clarifying conceptual questions and questions of interpretation.

> 5. What do you consider the most important open problems in probability theory and/or statistics and what are the prospects for progress?

I am going to seriously dodge this question. Instead, let me float a few ideas that have been puzzling me in philosophy of probability. The first one is a comment on an ongoing debate in philosophy of probability. The second one is a curious clash of intuitions concerning counterfactuals and stochastic events that deserves some attention. The third one is a simple conjecture that seems like a school exercise, but is quite recalcitrant and has repercussions in game-theory. The fourth one is an open question in strategic voting. And considering that I have been teaching philosophy of public policy for the last five years, let me conclude with some normative questions about risk analysis. Neither one of these is 'the most important problem' in philosophy of probability theory, but I hope that you will find them of some interest.

First, ever since I read Hume's *On Miracles*, I was bothered by the following issue. Consider a miracle that is in the category of highly improbable events—six bullets fired from close range all missing their target, *Pulp-Fiction* style. Or make it as improbable as you wish. Suppose that we are informed that this happened to some monk in the 13th century and contrary to Hume, suppose that we know that this is a veridical report. Then the natural thing is to shrug it off and say—strange things happen. But the curious thing is that when that same equally improbable event happens to you, you may well go the route of Jules (in *Pulp Fiction*) and embrace theism. Is this irrational? Should I count what happens to *me* as equally good evidence as what happens to others? Stephen Leeds and I (2002) tried to dissolve the puzzle, but I still think that it is more recalcitrant than our analysis suggests. Though much work has been done on probability assignments to indexical statements, e.g. in connection with the self-selection bias, the doomsday argument, the puzzle of the absent-minded driver and Sleeping Beauty, our understanding of these issues seems very unsatisfactory. I don't foresee quick solutions that will put these paradoxes to rest. Rather, there are insufficiently understood connections with other areas in philosophy, e.g. evidential *versus* causal decision theory and possibly the A- and B-theory of time.

Second, it is a mark of a good philosophical problem that you can get a lay audience to split in half and stare at each other in utter disbelief. In philosophy of probability, Newcomb's problem comes to mind. However, in the Newcomb problem, it seems to me that non-philosophers are publicly 'two-boxers' but will deviate to 'one-boxing' in a secret ballot. But let me suggest a case of split intuitions that has not received much discussion. Suppose that I have the opportunity to invite Sue on Monday to a party that will take place on Saturday. Independently of my decision, you will flip a fair coin and invite Sue on Wednesday if and only if Heads comes up. Unbeknownst to both of us, Sue will come if and only if she is not invited by either you or me. Now suppose that I invite Sue, you flip a coin, which comes up Tails, do not invite her and Sue does not come to the party. Take the counterfactual: 'If I had not invited Sue, she would have come to the party.' This seems to me either false or indeterminate on Tuesday, because your coin flip may still come up Heads. But it is true on Thursday—i.e. after your coin flip. Is this not what it means to say that the outcome of your coin flip is causally independent of what I do? I tried to model such truth-value switches or settlings over time. (*Philosophical Studies*, 1998) However, some people vehemently deny that this counterfactual is true on Thursday. Their reasoning is that if I had not invited Sue, then I would have opened up a completely different world history and there is no telling how the coin would have come up in this world history. If I take one path in Borges's *Garden of Forking Paths* of possibilities then there is no reason to suppose that there is any parallel between how the path of actuality is implemented down the road and how it would have been implemented if I had taken the other path. So on Thursday, it is still the case that Sue might or might not have come to the party, had I not invited her, because there is no telling how your coin would have come up on the world history initiated by my not inviting her. I do not share this intuition, but I genuinely do not know what to make of it. (Since writing this, I learned that this is in essence the "Morgenbesser problem", presented by Michael Slote in the next to last footnote of "Time in Counterfactuals" (2004) and discussed by Dorothy Edgington in Dowe and Noordhof's *Cause and Chance* (2004). In philosophy one often feels like the caveman who, sitting in front of his tools, says "Hard times for inventors, everything has been invented." It seems as if we can do no better than lift a small tip of the veil and it's hard to find a tip that has never been lifted before. Nonetheless, I think that it's worth flagging the need for a systematic investigation of this problem.)

Third, consider the following problem. Suppose that you can flip a coin n times. If Heads comes up i times, you receive a certain sum of money x_i. The sequence of payoffs is single-peaked, i.e. it has the following form $x_0 < x_1 < \ldots x_{i-1} < x_i > x_{i+1} > \ldots > x_n$ for $i = 0, \ldots, n$. You know the values of each x_i and are allowed to pick the probability p that the coin will come up Heads. Show that there exists exactly one value of p which maximises your expected payoff. Proving this conjecture is challenging and opens up some interesting venues for the analysis of doubly symmetric games which are relevant to the tragedy of the commons and to strategic voting in juries. But too much needs to be said to fill in all the details.

Fourth, this brings me to an open question in strategic voting in juries. One tends to think that unanimity voting (rather than majority voting) in juries makes it less likely that the innocent will be convicted. But this is not so obvious. A jury member may reason as follows. The only time that my vote matters is when all other voters vote guilty. Now if they are voting strictly in accordance with their private assessments of the guilt or innocence of the suspect, then this provides me with additional information. I may privately assess that the suspect is innocent, but this information could be outweighed by the fact that all others privately assess her to be guilty. And should I not take this into account in my decision? But of course, if we all reason like this, then we are bound to convict the innocent. The rational strategy is a Nash Equilibrium of mixed strategies in which each player votes guilty with some probability, even if their private assessment is that the suspect is innocent. And given some plausible values of the relevant parameters, the chance of convicting the innocent under unanimity vote may exceed this chance under majority vote. This is the moral that is at the core of the literature on strategic voting in juries. Now if our private assessments of guilt are equally reliable as of innocence, the prior probability of guilt equals the prior probability on innocence, and it is equally bad to convict the innocent as it is to acquit the guilty, then majority vote will induce truthful voting – i.e. voting one's private assessment – and this voting rule maximises utility. But what happens if we vary the reliability of our private assessment of guilt *versus* innocence? For some crimes, it may be quite difficult to find the incriminating evidence for a guilty suspect. For some crimes it may be all too easy to stumble upon misleading evidence against an innocent suspect. In such cases, majority vote may not maximise truthfulness and may not maximise expected utility. The voting rule that maximises truthfulness has a quota that diverges from simple majority. Furthermore, the voting rule that maximises truthfulness sometimes maximises expected utility, but this is not always the case. And what happens if we vary the prior probability of guilt? Or if we stipulate that it is much worse to convict the innocent than to acquit the guilty? There are some interesting results addressing these issues in the literature, but there are still many open questions.

Finally, let me close with some issues in philosophy of public policy. Defendants of risk analysis often take it to be a mark of a rational policy that for a given benefit the expected harm should be decisive between policies. Clearly, the public is prone to misread evidence and one does not want to dictate policy on grounds of a misguided hysteria or carelessness. Science should have the authority to overrule certain irrational biases. But there are interesting cases in which more than the expected harm seems to matter. Here are two examples. First, Monsanto was under attack that Bovine Growth Hormone (BGH) increases the incidence of mastitis and since mastitis is treated by antibiotics, there is a certain expected harm to the public. Monsanto defended itself by saying that BGH increases the milk production and it is the increased milk production that increases the incidence of mastitis. So if there is a Markov chain from BGH over milk production to mastitis, then the increase in expected harm is permissible, but if there is a causal fork from BGH to milk production and to mastitis, then it is not. It is interesting that at least for some people causal structure seems to matter to responsibility in the face of the same expected harm. Second, suppose that we are comparing two projected public works. Let us assume that the expected death rate is precisely one. Now it matters a great deal how this expectation is generated. For each person there is a sequence of tasks and for each task there is an associated exponential distribution that characterises the risk. Now we can imagine that the risk is spread evenly across persons and across times. This is less objectionable than when the risk is concentrated in one person or in one time slot. We may prefer public works with a higher expected death rate to public works with a lower expected death rate because the spread of the risk across persons and across times is greater in the former than in the latter. There are many cases in which policy-making is not informed by a scientific assessment of the expected harm. But the challenge is to determine what considerations other than expected harm ought to enter the calculus and how this should be done in a principled fashion and what considerations ought to be dismissed as expressions of irrational biases and fears.

I am afraid that this was not quite akin to Hilbert's 23 open problems in mathematics, but I hope that these more modest questions may spark someone's interest.