

Review of Wolfgang Spohn
*The Laws of Belief: Ranking Theory
and its Philosophical Implications*
(Oxford University Press, 2012)

Jake Chandler

Weighing in at close to 600 pages of densely written material and packed with theorems and definitions, Wolfgang Spohn's Lakatos Award winning opus magnum is an impressive volume, offering a comprehensive overview of the systematic and highly distinctive body of work that he has been developing over the past 30 years.

The central aim of *The Laws of Belief* is to outline a formal theory of rational opinion change that is capable of delivering insights into questions that have been of traditional interest to epistemologists and philosophers of science. This, of course, is a theoretical role that many would take to already be comfortably occupied by the popular Bayesian model of degrees of confidence, aka 'degrees of belief'. But Spohn disagrees. His view is that, since many traditional debates in epistemology and the philosophy of science have been framed in terms of the notion of belief, rather than of degree of confidence, Bayesianism is at least *prima facie* ill-suited to the task at hand. Furthermore, he argues, the prospects of expanding the Bayesian model so as to accommodate the notion of plain belief are rather poor: best simply start from scratch with a theory of belief dynamics.

Accordingly, Spohn's point of departure is an axiomatic theory of belief change. More specifically, it is a theory of iterated belief contraction: a theory of how one ought to adjust one's overall set of beliefs in response to a succession of retractions of one or more of its members. Outlined only fairly late in the book, in Chapter 8, the proposal strengthens in novel and intriguing ways a number of principles already found in the belief revision literature, including the AGM postulates for contraction (Alchourrón *et al* 1985) and a set of further postulates recently studied by Samir Chopra and his colleagues (Chopra *et al* 2008).

With this axiomatisation in hand, Spohn offers up a key technical result that he established some years prior in collaboration with Matthias Hild (Hild & Spohn 2008): if an agent's contraction dispositions can be deemed rational by the lights of Spohn's theory of contraction, they can then be represented as being underpinned by a corresponding sequence of operations on something that Spohn calls a 'negative ranking function' κ . This function, unique up to a certain type of transformation, associates with each proposition expressible in the agent's

‘language of thought’ a rank drawn from the extended set of positive real numbers. The association is made in such a way that, for any such proposition A , either A or $\neg A$ receives a minimal rank of zero. The derivation of this result is not intended as a gratuitous technical exercise. Indeed, Spohn further claims that these ranks characterise what one would intuitively be inclined to call the rational agent’s ‘degrees of disbelief’. What we would call the agent’s ‘degrees of belief’ are then claimed to be given by a corresponding ‘two-sided’ ranking function τ , ranging over the extended set of real numbers, positive and negative, that is straightforwardly interdefinable with κ . For Spohn, then, the laws of degree of (dis)belief fall out of the laws of belief.

One of the upshots of the general picture provided by Spohn is a rather straightforward proposal regarding the relation between graded belief and belief simpliciter: a proposition A is believed just in case $\neg A$ is disbelieved to a non-zero degree (or equivalently, just in case A is believed to a strictly positive degree). Contraction by A , then, is correspondingly understood as a particular procedure that results in reducing the degree of disbelief in $\neg A$ to zero. It is a special case of a more general class of change in degree of belief that he calls ‘ $A \rightarrow n$ -conditionalisation’, which has the result of shifting the degree of disbelief in $\neg A$ to n . In cases in which n is set to zero, the procedure ensures that A is not believed: we have a case of contraction. In cases in which n is set to a greater value, the procedure ensures that A is believed: we have an instance of what belief revision theorists would call ‘revision’. A noteworthy special case of the latter is $A \rightarrow \infty$ -conditionalisation. This, in essence, is ranking theory’s answer to the Bayesian operation of ‘strict conditionalisation’. It ensures that A becomes believed with complete certainty, with the negative rank of any proposition B being then given by what Spohn calls B ’s ‘conditional negative rank’ given A , $\kappa(B | A)$, ranking theory’s answer to Bayesian conditional probability (the new two-sided rank of B is correspondingly given by $\tau(B | A)$).

Ranking theory is of course but one of many quantitative models of uncertainty on the market and, in Chapters 10 and 11, Spohn indulges our curiosity regarding the place of his model in this broader landscape. In Chapter 10, building on a helpful critical presentation of the Bayesian model in Chapter 3, he flirts with—but ultimately discounts—the possibility that ranks might correspond to logarithms of non-standard probabilities. In Chapter 11, he reminds us of the fact that ranking theory is both isomorphic with Prade and Dubois’ possibility theory (Prade & Dubois 1998), which is subsumable as a special case of a more general well known formalism due to Dempster and Shafer (Shafer 1976; possibility functions are indeed a special ‘consonant’ case of Dempster-Shafer belief functions).

As mentioned in the second paragraph of this review, Spohn’s model provides the backdrop for detailed discussions of a rather broad range of key concepts in epistemology and the philosophy of science, ranging from reasons for belief and a prioricity (Chapter 6), through

laws of nature (Chapter 12), dispositions (Chapter 13) and causal relevance (Chapter 14), to perception and justification (Chapter 16). In what follows, due to space considerations, I will restrict my exposition to the very basics of his discussions of evidential and causal relevance, which strike me as probably being among the more central contributions of the book, beyond ranking theory proper.

The intuition that guides Spohn's ranking-theoretic treatment of reasons for belief is one that will be familiar from the Bayesian literature: to take A to provide a reason to believe B involves being disposed to believe that B more firmly upon coming to believe that A . Mirroring a standard Bayesian way of cashing out this intuition, Spohn then suggests that, to the extent that one's contraction dispositions can be described by means of the two-sided ranking function τ , one is committed to the claim that A is a reason to believe B just in case $\tau(B | A) > \tau(B | \neg A)$. (I should probably note that Spohn actually puts things a little differently, telling us that A is a reason to believe B 'relative to τ ' just in case the relevant inequality holds. But this of course rather unhelpfully leaves open the nature of the relevant relation to τ . The formulation that I give is my best reconstruction of what he could possibly have had in mind.) This proposal is extended to the notion of A 's being a reason to believe B conditional on C in the obvious way, by taking the relevant inequality to be $\tau(B | A \& C) > \tau(B | \neg A \& C)$. Spohn argues that his proposal has a number of desirable formal properties, as does his parallel treatment of the complementary relation of evidential independence, allowing him to exploit a rich body of literature on so-called Bayes nets.

Reasons for belief then play a central role in Spohn's treatment of claims of deterministic causal relevance, which is, as Spohn notes, strongly reminiscent of the work on indeterministic causation that traces back to Reichenbach, Suppes and others. Spohn starts by defining commitment to the existence of a relation of direct causation—i.e. of causation unmediated by intermediate causal factors—between a cause C and effect E in terms of a commitment to C 's providing a reason to believe E , conditional on the history of the world up to but excluding E . The account is then extended to commitments to causal claims of a more general nature by considering the transitive closure of direct causation. In the course of the relevant chapter, Spohn notably discusses a series of classic cases, ranging from overdetermination to so-called trumping preemption, which he takes his account to handle in a satisfactory manner.

This brief and somewhat superficial sketch does not in any way do justice to the breadth and depth of what is a painstakingly detailed piece of work, a genuine treasure trove of interesting comments and technical results. It does however afford what I take to be an adequate general sense of what is on offer. It also provides a sufficient backdrop for me to air some concerns regarding some of its key claims. Let me list these in the corresponding order.

First off, I must say that I share Spohn's concerns regarding the awkwardness of trying to shoehorn a variety of traditional problems primarily framed in terms of plain belief into a model

of opinion dynamics that trades primarily in degrees of confidence. This is a view that I have already aired in print, in a paper in which I attempt to put a model of belief dynamics to work in clarifying the logic of reasons for belief (see Chandler 2012; it should be noted that my approach to this differs somewhat from Spohn's). I am also, to some extent, sympathetic to his concerns regarding the prospects of delivering an appropriately extended Bayesian model.

However, it should probably first be noted that much of Spohn's discussion of the further philosophical application of ranking theory, be it regarding reasons for belief or causes and effects, proceeds pretty much exclusively in terms of talk of graded belief. It is hard to see the precise sense in which the Bayesian model's current limitations with respect to plain belief have much of an incidence here.

Secondly, there is a lingering concern that Spohn may have a little too rapidly dismissed the possibility of accommodating a notion of plain belief in a Bayesian framework. It is true that the reduction of plain belief that A to a subjective probability of A of 1 seems perhaps implausible, as Spohn notes a number of times. Furthermore, it also appears to be the case that this route to reduction is the only viable one to take, *so long as* we make the assumption that the believability of A is a sole function of the subjective probability that we assign to it. This is the moral that one can draw from the paper by Douven and Williamson (2006) that Spohn cites. But the appropriateness of the assumption in question remains unclear and a number of reductivist proposals do away with it. These include the suggestion made by Leitgeb (2013), which Spohn briefly considers on pp. 213–214, and a number of options that I have myself floated (Chandler 2010; 2013). Of course, these kinds of worries are not the only obstacles to reductionism: further potential concerns arise in relation to the dynamic behaviour of beliefs (Lin & Kelly 2012) and indeed Spohn's dismissal of Leitgeb's proposal seems based on this broad type of consideration. Nevertheless, it seems fair to say that the situation for the Bayesian is probably not quite as bleak as depicted.

Turning to Spohn's account of iterated contraction, it must be said that the details of his axiomatisation are somewhat hard to grasp intuitively. As mentioned earlier, Spohn does establish that it entails a number of standard postulates, which is certainly helpful. But Spohn's commitments go substantially beyond these and the justification that he offers for doing so is essentially limited to a handful of rather terse paragraphs on pp. 94–96 and pp. 179–181. It is a pity that a lengthier and more careful discussion of their motivation was not provided. Pending a closer examination of such a key element of Spohn's framework, it strikes me that any endorsement of the latter ought to be tempered by a corresponding degree of caution.

Spohn's identification of ranks with degrees of disbelief also strikes me as being somewhat under-argued. It seems that there is a rather substantial inferential leap from (i) the observation that rational iterated contractions are representable by a particular type of operation (*viz* $A \rightarrow 0$ -conditionalisation) on a ranking function to (ii) the claim that what we would intuitively

identify as degrees of disbelief (a) have a ranking-theoretic structure, (b) bear the suggested relation to plain belief and (c) ought to be updated in the manner prescribed.

This general kind of concern is not new: similar worries have been voiced in relation to a somewhat analogous representation result in the Bayesian literature, which has been taken by some to prop up the view that what we would intuitively identify as degrees of belief have a probabilistic structure. There, it is noted that, if an agent's *preferences* obey a number of putative rationality constraints, then these preferences can be represented as being underpinned by a *subjective probability function* and utility function, also unique up to a certain type of transformation. It is then further argued that that the subjective probability associated with a proposition characterises the rational agent's degree of belief in the proposition. But as Zynda (2000) has pointed out, it turns out that the very same preferences can equally well be represented as being somewhat differently underpinned by a *non-probabilistic* doxastic function and utility function. Thus, even if one grants the aforementioned rationality constraints on preferences and grants some type of connection between rational degrees of belief and preferences, further considerations need to be adduced in order to secure the Bayesian conclusion.

Although I am not aware that it *can* be shown that rational iterated contractions are representable by an entirely different type of operation on a structure that is not a ranking function, Spohn would be well-advised to show that it *cannot*. Short of demonstrating this, it would seem that an independent line of argumentation is required. But while Spohn does admittedly consider two such potential lines, respectively due to Hild and Huber, he ultimately finds both to be inconclusive. (I should probably add that the provision of an independent justification is all the more important in view of some of the peculiarities of the constraints that Spohn would have us impose. Indeed, unlike what is the case with the Bayesian model, or indeed the more general Dempster-Shafer model alluded to above, Spohn's framework requires of agents that they do not assign a positive degree of disbelief to both a proposition and its negation. This, I take to constitute a significant departure from our commonsense understanding of the constraints that are in place. As such, it seems to naturally require some form of further justification.)

Does any of this matter further down the line? It seems that it might well do. Indeed, if ranks do not correspond to our pre-theoretic degrees of disbelief, it is no longer clear what justifies the claim that commitments to evidential claims (and hence causal claims) can be cashed out in terms of the inequality of conditional ranks that Spohn considers. The claim may of course be correct, but it is no longer obvious why it should be thought to be so.

All in all, it therefore strikes me that there remain at this stage too many key uncertainties to be able to accurately judge the success of Spohn's enterprise. I should also probably note that the book is not a breezy read, to say the least, as Spohn warns us in the introduction. It will

likely only be accessible to a somewhat specialist (and enduring!) audience working at the more technical end of the philosophical spectrum. And indeed, it is somewhat of a shame that Spohn did not seize the occasion to write something a little more accessible than what his previous publications have accustomed us to (for a more approachable survey of his work, I would recommend his excellent 2009 article). This notwithstanding, I would like to stress that *The Laws of Belief* is a very fine book indeed and there is much to be learned from it even for those who, like myself, fall short of buying into the whole package on offer. Spohn has delivered a bold and iconoclastic contribution that amply deserves its place among the great works of contemporary formal philosophy.

REFERENCES

- Alchourrón, C.E., P. Gärdenfors and D. Makinson (1985). On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50(02), pp. 510–530.
- Chandler, J. (2010). The lottery paradox generalized? *British Journal for the Philosophy of Science*, 61(3), pp. 667–679.
- Chandler, J. (2012). Transmission failure, AGM style. *Erkenntnis*, 78(2), pp. 383–398.
- Chandler, J. (2013). Acceptance, aggregation and scoring rules. *Erkenntnis*, 78(1), pp. 201–217.
- Chopra, S., A. Ghose, T. Meyer, and K. Wong (2008). Iterated belief change and the recovery axiom. *Journal of Philosophical Logic*, 37, pp. 501–520.
- Douven, I. and T. Williamson, (2006). Generalizing the lottery paradox. *British Journal for the Philosophy of Science*, 57(4), 755–779.
- Dubois, D. and H. Prade (2008). *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. New York: Plenum Press.
- Leitgeb, H. (2013). Reducing belief simpliciter to degrees of belief. *Annals of Pure and Applied Logic*, 164(12), pp. 1338–1389.
- Lin, H., and K.T. Kelly (2012). Propositional reasoning that tracks probabilistic reasoning. *Journal of Philosophical Logic*, 41(6), pp. 957–981.
- Shafer, G. (1976). *A Mathematical of Evidence*. Princeton: Princeton University Press.
- Spohn, W. (2009). A survey of ranking theory. In C. Schmidt-Petri and F. Huber (eds.) *Degrees of Belief*, Synthese Library 342, pp. 185–228.

Zynda, L. (2000). Representation theorems and realism about degrees of belief. *Philosophy of Science*, 67(1), pp. 45–69.