

ISAAC LEVI'S POTENTIALLY SURPRISING EPISTEMOLOGICAL PICTURE

*Wolfgang Spohn
Department of Philosophy
University of Konstanz
D - 78457 Konstanz
Germany*

1. A brief look 40 years back

There can be no doubt about how much Isaac Levi has taught us in epistemological matters, I certainly have no doubt as far as I am concerned. Yet, not for the first time I wonder at how philosophers can be so close and so different at the same time. The better the comparability, I assume, the sharper the comparison. Hence, this paper is devoted to a brief comparison of Levi's epistemological picture and mine. It mainly stays on an informal level, aiming at the basic features. The intention is not a critical one. It is rather to promote mutual understanding, because the similarities and differences are not easy to grasp.

Current discussions in formal epistemology tend to be quite specialized, building on rich, but not well reflected presuppositions. So, a brief look at its history may be healthy. It is indeed necessary for understanding Levi's role and position.

The history of formal epistemology¹ is predominantly that of probability theory. Mathematical probability was the only clear structure that emerged through the centuries. Not that there would have been no alternatives at all. There are various ideas, perhaps subsumable under the heading "Baconian probability", which hung around for centuries as well. But they never took a clear and determinate shape, and thus probability theory could develop its unrivalled power, culminating, as far

¹ There can be no doubt about the lasting importance of formal epistemology. It may be the lesser part of epistemology. Still, the major part is using its central notions like "inference", "justification", "reason", etc. *without any adequate theory* in reserve. Not that formal epistemology would make unanimous offers here. But it is making offers at least – and is therefore desperately needed.

as its philosophical use is concerned, in the work of Frank Ramsey, Bruno de Finetti, Leonard Savage, Rudolf Carnap, and others.

20th century philosophy of science made renewed attempts at alternatives.² Popper developed his account of corroboration, Hempel heroically engaged in qualitative confirmation theory. However, such attempts were not well received. So, the situation at the beginning of the 60's was characterized by the dominance of (subjective) probability theory, the impotence of alternatives, and yet the remaining feeling that probability cannot be everything. This feeling had at least three sources: The search for alternatives, even if failed so far, was at the same time an expression of a sense of incompleteness. The observation that probability theory was the hobby only of formal epistemologists, but hardly used by all the other ones was at least irritating. And the rise of doxastic and epistemic logic (see in particular Hintikka 1962) had shown that at least an elementary account of belief and knowledge besides probability was rigorously possible. All three sources pointed into the same direction: what was missing was a full-blown formal account of belief or acceptance (and, possibly, knowledge).

In my retrospective, the situation was essentially dramatized by the so-called lottery paradox, introduced by Kyburg (1961, p. 197) (but see also Hempel 1962, pp. 163f.). It expressed an already sharpened view of the situation, and its centrality soon became obvious. This was a crucial time at which formal epistemology definitely raised to a new stage and quality. Since then, the developments have been breath-taking, but I locate their starting point rather there than anywhere else.

For sure, Isaac Levi was one of the central actors in this epistemological revolution. His first book (1967) was his response to this situation. In retrospective, it has become clear how strongly his book has determined his research agenda till the present day. This is why I say one can understand his work only by looking at that situation in the 60's.

The beauty of the lottery paradox lies in its perfect simplicity and in its surprising capacity to crystalize various basic options in epistemology. The immediate result of the lottery paradox was that there is no simple analysis of belief or acceptance in terms of subjective probability. The immediate question it raised was: what then is the relation between belief and probability, two obviously fun-

² Did I forget formal logic starting in late 19th century? Of course, its import was overwhelming and still is. But deduction is not even half of epistemology. So, logic is auxiliary to epistemology as to many other disciplines, rather than formal epistemology in itself.

damental epistemological notions? The range of options is surprisingly similar to that concerning the relation of the mental and the physical.

One position is *eliminativism*: If belief is not analyzable in probabilistic terms, just drop it! We need not talk of anything besides probability in epistemology. This is the position Richard Jeffrey, one of the other central actors, ably defended throughout his life (see, e.g., Jeffrey 1992). He called it radical probabilism. In a way, it is not as radical as eliminativism in the philosophy of mind, which, in effect, is an unredeemed cheque on the future. Jeffrey's eliminativism is conservative, since it is a philosophically conscientious defense of the previous status quo in formal epistemology. Without doubt, as far as theoretical unity and elegance is concerned, Jeffrey's position is superior to all others. I was always attracted by this elegance.

Still, eliminativism is deeply incredible, also in epistemology. It cannot be simply mistaken or confused or superfluous to talk of belief. What, then are the other options?

As in the philosophy of mind, the *prima facie* most attractive option is *reductionism*, of course. But it fares differently. Simple realizations of that option are barred by the lottery paradox.³ And we have not seen any more involved realization of any plausibility. Reductionism in the epistemological case does not seem to be a live option.

Reverse eliminativism or reductionism is even more unfeasible; to try to get rid of probability in favor of belief is obviously crazy. Hence, any kind of monism concerning belief and probability seems excluded. Maybe, both can be reduced to, or replaced by, something third. There are even ideas what that third might be.⁴ However, this is not the place to expand discussion into that direction. What remains is epistemological *dualism* (or pluralism, if the need for further basic doxastic structures should be imperative).

Dualism may take different forms. One is *interactionism*, as one may call it. This consists in an integral picture, which gives both, belief and probability, their due place and describes how they interact without reducing one to the other. Of

³ The analogous riddle is perhaps presented by Frank Jackson's color-blind Mary; but whether it can bar the reduction of the mental to the physical is highly contested.

⁴ The first account that may be interpreted in this way is, as far as I know, the theory of belief functions of Shafer (1976). Probability measures are special belief functions, and consonant belief functions as Shafer (1976, ch. 10) calls them may be understood as representing belief (though I argued in Spohn (1990) that they do not do it adequately). Plausibility measures as elaborated by Halpern (2003) are to provide an overarching structure with many familiar theories as special cases. My probabilified ranking functions (see the appendix) provide yet another option for such unification.

course, this position gains substance only by proposing specific forms of interaction. One might think that various such forms are possible. However, belief and probability do not mesh easily; the first difficulty is to find at all a workable coherent interactive scheme. Not many have seriously tackled this difficulty, and it is clearly Isaac Levi who has made the most elaborate proposal in this spirit. I shall return to it.

Which other form might dualism take, if not interactionism?⁵ *Separatism*, as I tentatively call it; what comes next to it in the philosophy of mind is perhaps psychophysical parallelism. Separatism is the view that acknowledges a useful theory of subjective probability and a useful theory of belief, and keeps them as coexisting, but separate enterprises not in need of unification in an integral picture. In a way, this view draws the most negative conclusion from the lottery paradox by accepting it as unsolvable. Is this a feasible position? Yes, I hope so. At least, separatism is the label which best fits my own position. What might be its justification? I discuss this below.

This classification of views is useful as far as it goes. In particular, it emphasizes the importance of the lottery paradox which forces one to take a stance within the spectrum of possibilities opened by it. Without such a stance one's epistemology would be incomplete. On the other hand, the classification is not entirely reliable. We need to look at the epistemological positions themselves; then we shall see that the dialectical situation is more complex than the classification reveals. So, let us inspect Levi's stance a bit more closely.

2. Levi's picture

The first cornerstone of Levi's epistemology⁶ is his notion of *full belief*. Full beliefs are free of any doubt; this is Levi's Peircean heritage. They are not in need of justification; they rather form the current base on which to proceed. However, full beliefs can change, the changes need be justified, and Levi is amply occupied with this justification.

⁵ Not epiphenomenalism; I have no idea what its epistemological analogy might be.

⁶ Levi's epistemology is surprisingly consistent over the decades. Many essential elements are laid out already in Levi (1967) that are still found today. The picture has become much richer over the years, and now and then an error (in his view) had to be corrected. I mainly refer to Levi (2004) which, among other things, beautifully summarizes Levi's view and its development.

Before turning to it, we have to locate probability. Full beliefs are undoubted or, as Levi often says, maximally and equally certain or infallible (from their own point of view), though corrigible (because they can change). They exclude certain possibilities and leave open others. They thus provide a standard of possibility, *serious possibility*, as Levi calls it. Serious possibilities define the space of inquiry, which at the same time is the *space of probability*; only serious possibilities (and sets of them) carry probability. Hence, full beliefs have probability 1, but not in the sense that I would bet my life on them, but by default, as it were, since full beliefs define the frame of probabilistic judgments. The possibilities excluded by them do not have probability 0, but no probability at all. This is, basically, Levi's dualism in the above sense.

In order to see how the dualism develops into an interactionism, we have to look at Levi's very detailed ideas about how to change full beliefs. There are two basic movements, expansion and contraction of full beliefs. Levi provides detailed justificatory accounts for them; all other changes can only be indirectly justified by getting decomposed into such basic moves. Such decomposition is also required because the two basic moves are guided by entirely different principles. Let us look at them separately.

Expansion of full belief is an epistemic decision problem, according to Levi. On the one hand, one seeks valuable information; on the other hand, one wants to avoid error. Thus, one's acceptance is torn between stronger and weaker propositions. This decision problem is solved by determining the *expected epistemic utility* $EV^*(h)$ of each proposition h .⁷ The simplest expression Levi derives for EV^* (cf. Levi 2004, ch. 3) is

$$EV^*(h) = Q(h) - q \cdot M(h).$$

Here, $Q(h)$ is the subject's *credal probability* for h . M is the subject's *informational value determining function* (the *undamped* version, to be precise), which Levi argues to behave like mathematical probability, although its interpretation must be sharply distinguished from that of credal probability. h 's informational value itself is then given by $1 - M(h)$; thus, the stronger a proposition, the larger its informational value. q , finally, is an *index of boldness* between 0 and 1.

⁷ The talk of propositions is mine. Levi prefers to talk of conjectures, hypotheses, etc. Ultimately, he analyses what doxastic attitudes refer to in terms of sentences and sets of sentences. This is an aspect I cannot, and need not, lay out here.

This label finds its explanation in Levi's *decision rule*. He argues that one should reject all and only those strongest propositions (or atoms of the propositional algebra under consideration) not yet excluded by one's full beliefs which carry negative expected epistemic utility. Equivalently, one should accept precisely the negations of these propositions (and their logical consequences). Hence, the greater one's q , the bolder one is in rejecting (and accepting) propositions.

Now it is evident why I call Levi's view interactionist; on the one hand, full beliefs delimit a space for subjective (credal) probabilities; on the other hand, credal probabilities are crucial for expanding full beliefs. In this way, belief and probability depend on each other, though none gets absorbed by the other.

Another important notion finds its place in the picture sketched so far. Instead of fixing the index of boldness in advance, one may as well consider the maximum index at which the proposition h is still unrejected and define it as its *degree of unrejectability* $q(h)$ (which is 1 if h is not rejected for any q). The more easily a proposition is rejected, the more surprising it is to learn that it obtains. Hence, the *degree of potential surprise* of a proposition h may be defined as $d(h) = 1 - q(h)$. The dual notion is the *degree of belief* or *plausibility* of h which is defined as $b(h) = d(\text{non-}h)$. That is, h is the more plausible, the more surprising its negation.

These definitions entail: All and only propositions excluded by full beliefs are maximally surprising. All and only full beliefs themselves are maximally plausible. If h is surprising (or plausible) to some positive degree, $\text{non-}h$ is not; i.e., if $d(h) > 0$, then $d(\text{non-}h) = 0$, and if $b(h) > 0$, then $b(\text{non-}h) = 0$. The potential surprise value of a disjunction is the minimum of the surprise values of its disjuncts; i.e., $d(g \text{ or } h) = \min(d(g), d(h))$. And the degree of plausibility of a conjunction is the minimum of the plausibility degrees of its conjuncts; i.e., $b(g \text{ and } h) = \min(b(g), b(h))$. Thus, as Levi emphasizes, the degrees of potential surprise and the dual degrees of belief are constructed so as to satisfy precisely the axioms already proposed by Shackle (1949).

There is no saying whether credal probabilities or degrees of plausibility have more claim on being called degrees of belief. They coexist, they are different, but they are related (via the presented chain of definitions). Credal probabilities are relevant for assessing risk and for determining expected values (or utilities) in practical and epistemic decision making. By contrast, according to Shackle's degrees of belief, the set of propositions believed at least to some degree $x > 0$ is consistent and deductively closed, just like the set of full beliefs; hence, they are

sued for assessing changes of full beliefs through expansion. This must suffice as a very brief description of Levi's account of expansion.

Contraction is also an epistemic decision problem, but quite different from expansion. There is no unfailing truth guarantee for our full beliefs, they must be conceived as corrigible, and indeed sometimes we find reasons to give up some of them. The problem is: how? In contrast to expansion, this problem is one-dimensional, as it were; there is no risk of error in contraction, since one does not acquire any beliefs at all in contraction and a fortiori not any false beliefs. Therefore, the only parameter guiding contraction is informational value, and since one is losing information in contraction, this loss must be minimized.

Thus put, contraction appears simpler than expansion. There is, however, a severe complication. At first, one may think one can apply the same informational value determining function M that was used in expansion. However, if contraction minimizes loss of information in this sense, the resulting contraction behavior is heavily defective, from an intuitive as well as a theoretical point of view. This is the starting point of a quite involved discussion in which Levi arrives at the result that loss of information must not be measured by the undamped version of M mentioned above which behaves like mathematical probability, but rather by a *damped* version. Of course, everything then depends on the specific damping applied – for all this see Levi (2004, ch. 4).

I cannot go into details here. But there is a tricky point of interpretation that I have to briefly explain. Damped informational value also satisfies the above-mentioned Shackle axioms. Therefore, one must take great care not to confuse the various uses of this structure. Damped informational value is not degree of plausibility, simply because plausibility, like probability, relates only to serious possibilities, whereas damped informational value refers to the possibilities excluded by full belief in order to govern contraction. One might be tempted then to interpret damped informational values as prior degrees of plausibility, i.e., as plausibility degrees one had in a (hypothetical) state before acquiring any substantial full beliefs. But this does not fit either. Prior plausibility rules expansion of that (hypothetical) ignorant state by bringing informational value and risk of error into a balance, but this explanation does not apply to damped informational value.

AGM belief revision theory was able to explain contraction (and revision) by an underlying entrenchment ordering, as it was called (cf. Gärdenfors 1988, ch. 4). Again, though, Levi discovers subtle differences. He proposes to reconstruct entrenchment in terms damped informational value. But they are not the same,

and hence differences in the view of contraction emerge. (For all this, see Levi 2004, ch. 5.)

To summarize: According to Levi, it must be informational value that guides contraction, and it must be a specific damped version in order for contraction to work properly. Other approaches to contraction are argued to be misguided, even if they have the same or a similar formal structure.

Of course, there are other changes of full beliefs. In particular, there are revisions in which one is forced to accept a hitherto unexpected or rejected proposition h , and residual shifts, as Levi calls them, where one replaces a formerly accepted proposition f by another proposition g . Any such change can be decomposed into a contraction and an expansion. Levi, however, insists that it *must* be so decomposed, since this is the only way to rationalize it. As explained, contractions and expansions have their own, but quite different justifications which can be applied successively, but not mixed to yield direct justifications of other forms of changes.

So much for Isaac Levi's epistemological picture. The richness of details of his many books could not even faintly be displayed here, but I hope I have fairly represented his basic moves. My picture is surprisingly different, so much so, indeed, that it would be hopeless to start discussing details. Hence, I am going to describe my picture at the same basic and informal level. If you want to be fair to me, however, please simply forget Levi's picture for the next section. Any attempt to translation would be premature and misleading. The two accounts will be brought together only in the final section.

3. My view

Like Levi I accept the moral of the lottery paradox, and like Levi I cannot simply forget about the notion of belief. Thus I am a dualist, too. However, I start as a methodological separatist. If the first aim, a reduction of belief to probability, is not feasible, then my next aim is to develop an account of belief as well and as far as possible – separately, because the clarification of the relation of that account to probability can only come afterwards as a third step. Whether I end up as a genuine separatist is not so clear. I would like too see myself as at least formally returning to reductionism (see the appendix); this would be theoretically most satisfying. Soberly, though, I prefer to present myself as a separatist (see below).

In developing an account of belief, I would very much like to naively talk of belief *simpliciter*. However, people talk of full belief, strong belief, weak belief, etc., and each adjective is backed up by a whole theory. In order to distinguish myself from all this, I introduced the notion of *plain belief*, hoping that “plain” would not be conceived as a qualification at all. 20 years later I feel that the qualification is rather misleading than helpful. Therefore I return here to talk of belief *simpliciter*. The main reason is that I cannot find that belief is in any way ambiguous. Belief is vague. When asked “*p*?”, I usually say: “yes” or “no” or “I don’t know”, in which cases I clearly express my belief or my ignorance. Quite often, I also start qualifying: “yes, I am absolutely sure” or “yes, I think so” or “I guess so” or “presumably”, etc. We have here a rich vocabulary, indicating first that belief admits of degrees (which need not be conceived as probabilities) and second that these degrees open a range of vagueness within which it is hard to determinately call someone a believer or a non-believer. But this does not make belief ambiguous; ambiguity, it seems to me, is imported only by the various theories about it. Likewise, I reject the purely theory-induced distinction between “belief” and “acceptance”.

As everybody, I start analyzing belief *simpliciter* as a set of propositions (or sentences) believed true (by a certain person at a certain time). And in order to start theorizing at all, I, like most, take such a *belief set* to be consistent and deductively closed. I like Levi’s defense of these assumptions in terms of commitment (cf. Levi 2004, sect. 1.2).

So far, the analysis is purely static. In order to be complete, an account of belief, as of any phenomenon in time, must be dynamic. Clearly, the functioning and malfunctioning of memory is the most important determinant of the dynamics of belief. But presupposing perfect memory and excluding other arational influences, it is experience, perception, observation which drives the dynamics of belief and which occupies the interest of epistemologists like me. Hence, my focus is on belief revision, which may be a simple expansion or a genuine revision, depending on the compatibility of experience with expectations.

At first, I thought that AGM belief revision theory (cf. Gärdenfors 1988, ch. 3) offers a perfect account of belief revision, but soon I realized that it is unsatisfactory because it violates what was later called the principle of categorical matching (cf. Gärdenfors, Rott, 1995, p. 37) and hence offers only a restricted account of iterated belief revision.

Thus, in my (1983 and 1988) I came to introduce ranking functions (a terminology proposed by Judea Pearl and Moises Goldszmidt to which I happily converted later on). Given a set W of all possibilities under consideration, a *ranking function* κ for W is a function from the power set $\mathcal{P}(W)$, the set of propositions, into the extended set of natural numbers $\mathbf{N} \cup \{\infty\}$ such that $\kappa(W) = 0$, $\kappa(\emptyset) = \infty$ and for all propositions A, B $\kappa(A \text{ or } B) = \min \{\kappa(A), \kappa(B)\}$. My standard explanation, (thanks to Isaac Levi) is, that ranks are *degrees of disbelief*: $\kappa(A) = 0$ says that A is not disbelieved at all, $\kappa(A) = n$ ($n \geq 1$) says that A is disbelieved to degree n . Hence A is *believed* relative to κ iff $\kappa(\text{non-}A) > 0$ (but see my important qualification in the next section). There is no difficulty in introducing the dual notion of degrees of positive belief, but ranks of disbelief are preferable for various reasons.

So far, the formal structure has many predecessors and Levi is right to point out that Shackle (1949) was the first: the above functions d of potential surprise and my ranking functions κ are obviously governed by the same axioms (the difference in the range of these functions, which is not arbitrary, need not concern us here). Of course, the informal predecessors reach back much further. However, the reference to predecessors is also misleading – in particular it has misled Levi, I think –, because it imports all the old and perhaps unwanted interpretations and because it obstructs the view on variations and new developments.

Indeed, what is doing all the work in my account of belief is the definition of conditional ranks. This is a topic that received remarkably little attention by the predecessors.⁸ The only explanation I have is that the dynamic perspective so central for me became focal within philosophy only in the 70's and that it is only in this perspective that conditional ranks unfold their power and beauty.

To be a bit more precise, the *conditional rank* of B given A is defined as $\kappa(B | A) = \kappa(A \text{ and } B) - \kappa(A)$, if $\kappa(A) < \infty$; otherwise, it is undefined. In other words, the degree of disbelief in A and the degree of disbelief in B given A add up to the degree of disbelief in A -and- B . This is intuitively convincing; it suggests a Cox-like justification of the ranking structure; it opens a way to measuring ranks (on a proportional scale) via contraction behavior (still to be explained); and it provides the clearest justification of the crucial last axiom for ranking functions: given the definition of conditional ranks, this axiom is equivalent to either $\kappa(A | A \text{ or } B) = 0$

⁸ Shackle is an exception; conditional potential surprise played an important role in his theory (cf. Shackle 1969, part IV). There (p. 205) he even briefly considers the definition of conditional ranks I use, but he rejects it. He prefers the postulate $d(A \text{ and } B) = \max \{d(A), d(B | A)\}$, which in fact takes $d(B | A)$ as an undefined primitive and which he takes to be simpler and less unrealistic.

or $\kappa(B \mid A \text{ or } B) = 0$ (or both), and this is simply a conditional consistency requirement saying that A and B cannot both be disbelieved given A -or- B .

In thorough-going analogy to probability theory conditional ranks allow a substantial development of ranking theory. Qualitative confirmation theory did not get off of the ground in the 50's and 60's, but now we can make the basic idea of confirmation theory work. A *confirms* B , or A is a *reason* for B (as I prefer to say in order to make ranking theory interesting for the traditional epistemologist), iff A is positively relevant to B , i.e., iff $\kappa(\text{non-}B \mid A) > \kappa(\text{non-}B \mid \text{non-}A)$ or $\kappa(B \mid A) < \kappa(B \mid \text{non-}A)$. We can define (conditional) dependence and independence with respect to ranking functions. Indeed, we can develop a full-blown ranking analogue to the theory of Bayesian nets, which is a beautiful formal representation of our intuitive ways of thinking, as such still grossly underrated by philosophers.

On this basis, then, my picture of belief dynamics is as follows: The subject directly perceives that A and must hence revise her body of beliefs so as to contain A . But no, this is wrong from the start; the idea of direct perception and its false understandings have done a lot of damage to philosophy in the last 350 years. I better begin in this way: Experience affects the subject's beliefs; somehow, among all the propositions under consideration, it is the proposition A which is affected first and thus believed to some degree n (i.e., $\kappa'(\text{non-}A) = n$, where κ' is the posterior ranking function). This degree n is part of the experiential process, and it can vary. I read in the newspaper that A , I hear with my own ears that A , my wife tells me that A , I see that A : these are four ways to learn A with increasing firmness of belief (since I trust my wife more than my ears). So, A as well as n form the input of the doxastic change which I take as given. But how should the subject respond to this experience? With Jeffrey (1965, ch. 11) I say that the prior ranks conditional on A and conditional on $\text{non-}A$ are not changed through this experience. That is, if κ is the prior ranking, we have for any proposition B :

$$\begin{aligned} \kappa'(B) &= \min \{ \kappa'(A \text{ and } B), \kappa'(\text{non-}A \text{ and } B) \} \\ &= \min \{ \kappa'(A) + \kappa'(B \mid A), \kappa'(\text{non-}A) + \kappa'(B \mid \text{non-}A) \} \\ &= \min \{ \kappa(B \mid A), n + \kappa(B \mid \text{non-}A) \}, \end{aligned}$$

where the first steps are transformations according to the ranking calculus and the last step realizes the crucial assumptions just explained. This κ' is what I describe as the A, n -conditionalization of κ (cf. Spohn 1983, sect. 5.3, and 1988, sect. 5). And it has many special cases. In particular, it is an expansion in case $\kappa(A) =$

$\kappa(\text{non-}A) = 0$ and $n > 0$; it is a genuine revision in case $\kappa(A) > 0$ and $n > 0$; and it is a genuine contraction in case $\kappa(\text{non-}A) > 0$ and $n = 0$. This rule of belief change can be iterated indefinitely as long as the prior rank of the experiential proposition is finite (i.e., $\kappa(A), \kappa(\text{non-}A) < \infty$); this is why I usually assumed that $\kappa(A) = \infty$ only for $A = \emptyset$. To this extent, I may claim to have offered a complete dynamics of belief. This must suffice as a sketch of my account.

Can this picture be combined with subjective probability theory? Yes and no. Ranking theory is very similar to probability theory; there is indeed an algorithm for translating probabilistic into ranking theorems which is almost guaranteed to work. Therefore, it is perfectly natural, and not a gerrymander, to integrate both into one notion which I call probabilified ranking functions (or ranked probability measures) – see the appendix for the definition. All the theoretical developments I sketched above work equally well for that notion. Hence, we may speak here of a genuine theoretical unification and of a reduction of both, probability and belief, to that notion. One may also call this simply an extended probabilistic point of view, I don't mind. In any case, If one accepts this unification, one returns to reductionism.

I do not dare so. For, the lottery paradox raises its ugly head again. Unavoidably, this unification has the consequence that disbelief in A , i.e. $\kappa(A) > 0$, entails $p(A) = 0$. In other words, believing in A , even to the lowest degree, means giving probability 1 to A . This strikes me as counter-intuitive. Believing in A is a very ordinary affair; it is not being probabilistically maximally certain of A , betting everything I have on A . Hence, from the point of view of the intended interpretation, the two notions, ranks and probabilities, *do not mesh*, as far as I can see. This is why I feel the unification to be artificial despite its theoretical beauty and why I prefer to develop the two theories separately, though always with an eye on the format parallel. Thus I remain a separatist (and may perhaps even be called a parallelist). But I would be glad if someone teaches me better.

4. Comparisons

These sketches suffice to make clear two things. First, that Isaac Levi and I are dealing with the same issues in the same spirit; that is why a comparison is relatively straightforward. Second that our conceptions differ widely; indeed, we diverge on nearly every point. Let me collect here the differences, and let me try to

identify their deeper sources. We shall find a divergence in basic attitudes, where it becomes nearly impossible to tell who is right and who is wrong.

Some differences have already been fairly explicit: (1) I have explained why Levi takes expansion and contraction to be the basic moves in belief change, and I have explained why I start considering revision which led me to state the role of A, n -conditionalization which in turn embraces contraction as a special case. (2) My contraction then endorses the AGM orthodoxy on contraction; this was my intention. Hence, I also endorse the recovery postulate for contraction not yet mentioned so far (cf. Gärdenfors 1988, p. 62). By contrast, Levi rejects the recovery postulate and he takes great pains to construct contraction accordingly; indeed, the whole chapter 4 of Levi (2004) is devoted to this issue. (3) I mentioned the importance of the principle of categorical matching for my line of reasoning. Levi (2004, sect. 5.11), however, does not accept it as a regulative principle for theories of belief change. (4) This difference is explained by a different approach to iterated belief change. I have sketched how this problem drove the development of my theory. Levi does not see such a big problem here. For him, informational value is the only parameter governing contraction and an essential parameter governing expansion which he assumes to be relatively stable, at least within a given inquiry. Thus, there is no need to develop a dynamic account for it. Given this stable parameter, Levi can, of course, also explain iterated changes.

One may certainly discover more differences of this specificity. This is good; if theories do not get down to such details, they are of no worth. In principle, we should carry out such differences. However, I refrain from starting an argument here. The specific differences only reflect deeper strategic ones, and without settling the latter, agreement on the former is spurious.

(5) The first strategic difference is – indeed, this was the red line of this paper – that Levi and I go together for a while as dualists and then split as interactionist and separatist.

(6) This difference is entailed, it seems to me, by a deep difference over the notion of belief. Levi's notion of full belief and the accompanying combination of infallibility and corrigibility do not fit into my picture, and my notion of belief and my rejection of this combination do not fit into his. Here, our views are entirely at cross-purpose. In a way, this conclusion is reached by Levi (2004, sect. 1.6 and 3.5), too.

Levi there suggests that I am a Parmenidean skeptic. This is a person, according to Levi, for whom the only standard of serious possibility is logical or con-

ceptual possibility and for whom full beliefs are restricted to unrevisably or incorrigibly a priori beliefs. Such a person is called a skeptic because for Levi the full beliefs of a subject precisely set the undoubtful frame within which she is conducting her inquiries. Thus, if the Parmenidean has such a narrow notion of full belief, he has a very large notion of doubt.

According to this suggestion, my beliefs can then be understood as expansions of Parmenidean full beliefs. This is so because Levi has assigned a firm place to Shackle's functions of potential surprise in his epistemological picture as determining which beliefs to accept according to the chosen index of boldness and because my ranking functions work the same and thus apparently do the same as the functions of potential surprise. Indeed, I am a particularly bold expander according to Levi, since by calling the minimal rank 1 already disbelief I am in effect applying the maximal index of boldness.

Maybe this is the relatively best embedding of my views into Levi's. Still, instead of forcing this embedding one better acknowledges that it does not fit at all. I grant that I have the notion of unrevisably a priori beliefs. Their negations are disbelieved with maximal rank ∞ for which no dynamics is explained; this is why they are unrevisable. Unlike Levi, though, I do not see this notion as a remnant of a degenerating research programme; on the contrary, the research programme on apriority puts forth blossoms again (cf. Spohn 2000).

Be this as it may, it is not correct to interpret my defense of a priori beliefs as holding a particularly narrow and rigid notion of full beliefs. I have no full beliefs at all. I have beliefs, I need not justify them unless questioned, but of course, I can justify them even if unquestioned. That is, each proposition receives that rank which is assigned to it by my balance of reasons for and against it (in the sense explicated above), and if this balance of reasons assigns a positive belief value to it (i.e. a positive rank of disbelief to its negation), then I believe it and do not doubt it. I do not see why I should treat my beliefs as doubtful, unless they are full. But, of course, my beliefs are revisable, just as Levi's full beliefs are.

Moreover, I need not be attached to being maximally bold (in Levi's terms). As Matthias Hild has pointed out to me in conversation, I am not forced to say that A is believed relative to κ iff $\kappa(\text{non-}A) > 0$. I could as well define stricter notions of believing A , say as $\kappa(\text{non-}A) > n$ for some $n > 0$. The logic of belief always comes out the same. We might indeed say that the logic of belief is represented by the ranking structure, however we precisely map the vague notion of belief onto this structure.

Here is a difference, I think, we should not try to bridge. At least it explains our attitudes toward probability. With his standard of full belief and serious possibility Levi can delimit a realm of probability and thus develop an interactionist picture. Without such a standard I can only keep belief and probability separate, or probabilify each level of belief (as explained in the appendix).

(7) The differences go still deeper. It seems to me that even our methodologies are opposite. Levi approaches epistemology in a constructive-synthetic way, whereas I rather have an analytic-structural view. Levi has enormously detailed stories to tell about our epistemological wheelwork; this is, of course, one of the fascinating aspects of his work. There are small wheels and big wheels, and Levi meticulously explains how they interact in order to produce judgment and belief. A particularly big wheel is informational value, but it is driven by several smaller wheels like explanatory power, simplicity, generality, specific interests, etc.

I am not telling any such stories; I even refuse to do so, since I am not so sure of all these wheels. What is simplicity? What is explanatory power? What is overall informational value? And so forth. What is the measurement theory for all these magnitudes? None, I suppose. As long as such questions are not well answered, I remain in doubt whether the wheelwork is a real, hypothetical, or imaginary one. This is why I prefer to tell my structural story. Whatever the inner epistemological workings, the resulting structure must be such and such, viz. a ranking structure, if my arguments and in particular my measurement theory for ranks hold good. My way of proceeding is then just the reverse of Levi's. Starting with this structure, my hope is to be able to analyse and lay bare some of this inner workings: lawhood, explanation, causal inference, the truth-conduciveness of reasons, etc.

Clearly, both, the analytic and the synthetic method, are legitimate. Presently, I do not see a clear meeting point, but maybe the methods can be joined. In the end, overall success decides; until then both projects have to be pushed ahead as far as possible.

(8) This methodological point helps explaining why Isaac Levi and I have such difficulties to agree about the relation between Shackle's functions of potential surprise and my ranking functions. As I have explained, Levi has firmly located Shackle's functions within his account of expansion and tends to put ranking functions into the same drawer. Damped informational value governing contraction may have the same formal structure, but has an entirely different use and in-

terpretation. From Levi's constructive point of view this is how one must see the situation.

From my structural point of view I do not care whether expected epistemic utility or damped informational value generate the relevant structure. Since it is the same structure, I comprise it in my ranking functions and apply it across the board. Hence, I insist that ranking functions govern both, expansion and contraction in a unified way (indeed, unified by A, n -conditionalization), even if this must appear as mixing apples and oranges to Levi.

(9) I see a final difference, perhaps the deepest of all, which motivates the issue about methodology. It is a difference about how to conceive of justification in epistemology. In a way, Levi is very modest; he does not want to justify current beliefs, but only changes of belief. Concerning the latter, he is less modest. As he repeatedly says, his inquiry of the epistemological wheelwork is to uncover in detail the justificatory structure supporting belief change.

I do not say that Levi is pursuing an objectivistic notion of justification here. He emphasizes the context-dependence of all his parameters. His index of boldness is entirely subjective; there is no prescription of a wise choice. Informational value is basically subjective; there is no need for people to agree on it. Still, I sense the remnants of an objectivistic notion in Levi's work. Even if a lot of wheels are subjective, his hope is to identify at least some objectively valid components (like explanatory power) and at least some objectively valid connections between the components (like expansion according to expected epistemic value).

This is not my picture. I am immodest in thinking that I can also justify my current beliefs (see (6) above), but in doing so I apply a more radically subjectivistic notion of justification. Like Levi, I take it as axiomatic for rational subjects that it is reasons that drive their doxastic dynamics. But I do not pretend to have an antecedent notion of reasons that I could plug in here. Rather, I read the axiom inversely – this is why I am also reversing the methodological order – as saying that for rational subjects reasons are whatever drives their doxastic dynamics. Look at their (actual and potential) dynamics, and you know what their justifications are! Perhaps, Levi is right in calling me a skeptic.

On the other hand, the difference is not so big as it may seem. Of course, I am also discontent with an entirely subjectivistic point of view, and then I am starting with the inquiry into what I call the a priori structure of the space of reasons (see Spohn 2000). Still, I very much doubt that Levi's and my inquiry can be brought to converge.

Facing so profound differences, I find it hard to start an argument; any argument is bound to end up in a draw. Hence, the point of this paper was only to explain and clarify the dialectic situation between Isaac Levi and me for the benefit of future discussion.

Appendix

Let me briefly explain what a probabilified ranking or a ranked probability is by working up from standard probability. Let W be a set of possibilities, which we assume to be finite for the sake of simplicity. Any subset of W is a proposition.

Then, p is a *probability measure* for W iff p is a function from \mathcal{P} , the set of propositions, into \mathbf{R} (= the set of real numbers) such that for all $A, B \in \mathcal{P}$:

- (1) $p(A) \geq 0$,
- (2) $p(W) = 1$,
- (3) if $A \cap B = \emptyset$, then $p(A \cup B) = p(A) + p(B)$.

If $p(A) \neq 0$, the *conditional probability* of B given A is defined as $p(B | A) = p(A \cap B) / p(A)$.

There is a more general notion that takes conditional probability as basic. p is a *Popper measure* for W iff there is a set $C \subseteq \mathcal{P}$ (the set of conditions) such that p is a function from $\mathcal{P} \times C$ into \mathbf{R} such that for all $A, B, C \in \mathcal{P}$:

- (4) for each $C \in C$ $p(\cdot | C)$ is a probability measure for W ,
- (5) if $B \cap C \in C$, then $p(A \cap B | C) = p(A | B \cap C) \cdot p(B | C)$,
- (6) if $B \in C$ and $p(A | B) > 0$, then $A \in C$.

It is well-known (cf., e.g., Spohn 1986) that Popper measures and standard probability measures relate in the following way: (p_0, \dots, p_n) is a *sequence of probability measures* for W iff:

- (7) each p_i ($i = 0, \dots, n$) is a probability measure for W ,
- (8) for each $j = 0, \dots, n$ there is a C_j such that $p_j(C_j) = 1$ and $p_i(C_j) = 0$ for all $i < j$.

Such sequences and Popper measures are strictly equivalent in the following sense: for each Popper measure p for W there is exactly one sequence (p_0, \dots, p_n) of probability measures for W , and vice versa, such that for each $A \in \mathcal{P}$ and $C \in \mathcal{C}$:

$$(9) \quad \text{if } j = \min \{i \mid p_i(C) > 0\}, \text{ then } p(A \mid C) = p_j(A \mid C).$$

Ranked probability measures, which are my proposal for unifying belief and probability, provide a further generalization. ρ is a *ranked probability measure* for W iff there is a function κ from \mathcal{P} into $\mathbf{N} \cup \{\infty\}$ and a function π from \mathcal{P} into \mathbf{R} such that for all $A, B \in \mathcal{P}$:

$$(10) \quad \rho(A) = \langle \kappa(A), \pi(A) \rangle,$$

$$(11) \quad \kappa(\emptyset) = \infty \text{ and } \pi(\emptyset) = 0,$$

$$(12) \quad \kappa(W) = 0 \text{ and } \pi(W) = 1,$$

$$(13) \quad \kappa(A) < \infty \text{ iff } \pi(A) > 0,$$

$$(14) \quad \text{if } \pi(A) > 0, \text{ then there is exactly one } C \text{ such that } \kappa(C) = \kappa(A) \text{ and } \pi(C) = 1,$$

$$(15) \quad \text{if } A \cap B = \emptyset, \text{ then } \kappa(A \cup B) = \min \{ \kappa(A), \kappa(B) \} \text{ and}$$

$$\pi(A \cup B) = \left\{ \begin{array}{l} \pi(A), \text{ if } \kappa(A) < \kappa(B) \\ \pi(B), \text{ if } \kappa(A) > \kappa(B) \\ \pi(A) + \pi(B), \text{ if } \kappa(A) = \kappa(B) \end{array} \right\}.$$

If $\kappa(A) < \infty$, the *conditional ranked probability* of B given A is defined as $\rho(B \mid A) = \langle \kappa(B \mid A), \pi(B \mid A) \rangle$, where $\kappa(B \mid A) = \kappa(A \cap B) - \kappa(A)$ and, if $\kappa(A \cap B) < \infty$ and C is the proposition such that $\kappa(C) = \kappa(A \cap B)$ and $\pi(C) = 1$, $\pi(B \mid A) = \pi(A \cap B \cap C) / \pi(A \cap C)$.

It is easy to see that such a ranked probability measure ρ is equivalent to a *ranked* sequence of probability measures which is like an above sequence satisfying (7) and (8) with the only difference that the sequence is not indexed by $\{0, \dots, n\}$, but by the finite ranks occupied by ρ , i.e., by $\{\kappa(A) \mid A \in \mathcal{P}, \kappa(A) < \infty\}$.

Ranked probability measures look like a hybrid, but, as emphasized, they show a unified behavior. One may conjecture the unified behavior is due to the fact that ranked probability measures are similar to non-standard probability measures, since ranks may be conceived as orders of magnitudes relative to some infinitesimal. This conjecture is certainly true, but the precise extent of the similarity still awaits clarification.

References

- Gärdenfors, Peter (1988), *Knowledge in Flux. Modeling the Dynamics of Epistemic States*, Cambridge, Mass.: MIT Press.
- Gärdenfors, Peter, and Hans Rott (1995), "Belief Revision", in: D.M. Gabbay, C.J. Hogger, J.A. Robinson (eds.), *Handbook of Logic in Artificial Intelligence and Logic Programming, vol. 4, Epistemic and Temporal Reasoning*, Oxford: Oxford University Press, pp. 35-132.
- Halpern, Joseph Y. (2003), *Reasoning about Uncertainty*, Cambridge, Mass.: MIT Press.
- Hempel, Carl G. (1962), "Deductive-Nomological vs. Statistical Explanation", in: H. Feigl, G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science, vol. III, Scientific Explanation, Space, and Time*, Minneapolis: University of Minnesota Press, pp. 98-169.
- Hintikka, Jaakko (1962), *Knowledge and Belief*, Ithaca, N.Y.: Cornell University Press.
- Jeffrey, Richard C. (1965), *The Logic of Decision*, Chicago: University of Chicago Press.
- Jeffrey, Richard C. (1992), *Probability and the Art of Judgment*, Cambridge: Cambridge University Press.
- Kyburg, Henry E. jr. (1961), *Probability and the Logic of Rational Belief*, Middletown, Conn.: Wesleyan University Press.
- Levi, Isaac (1967), *Gambling with Truth*, New York: Knopf.
- Levi, Isaac (2004), *Mild Contraction: Evaluating Loss of Information Due to Loss of Belief*, Oxford: Oxford University Press.
- Shackle, George L.S. (1949), *Expectation in Economics*, Cambridge: Cambridge University Press.
- Shackle, George L.S. (1969) *Decision, Order, and Time in Human Affairs*, Cambridge: Cambridge University Press, 2nd ed.
- Shafer, Glenn (1976), *A Mathematical Theory of Evidence*, Princeton: Princeton University Press.
- Spohn, Wolfgang (1983), *Eine Theorie der Kausalität*, unpublished Habilitationsschrift, München.
- Spohn, Wolfgang (1986), "The Representation of Popper Measures", *Topoi* 5, 69-74.
- Spohn, Wolfgang (1988), "Ordinal Conditional Functions: A Dynamic Theory of Epistemic States", in: W.L. Harper, B. Skyrms (eds.), *Causation in Decision, Belief Change, and Statistics*, Dordrecht: Kluwer, pp. 105-134.
- Spohn, Wolfgang (1990), "A General Non-Probabilistic Theory of Inductive Reasoning", in: R.D. Shachter, T.S. Levitt, J. Lemmer, and L.N. Kanal (eds.), *Uncertainty in Artificial Intelligence 4*, Amsterdam: Elsevier, pp. 149-158.
- Spohn, Wolfgang (2000), "Über die Struktur theoretischer Gründe", in: J. Mittelstraß (ed.), *Die Zukunft des Wissens. Akten des 18. Deutschen Kongresses für Philosophie*, Berlin: Akademie Verlag, pp. 163-176.