dialectica

Vol 77 : 1
March 2023

A Generalization of the Reflection Principle

Wolfgang Spohn

doi:10.48106/dial.v77.i1.04

Wolfgang Spohn. 2023. "A Generalization of the Reflection Principle." *Dialectica* 77(1): 1–1. doi:10.48106/dial.v77.i1.04.

A Generalization of the Reflection Principle

Wolfgang Spohn

This paper generalizes (probabilistic) auto-epistemology by amending the original forward-looking reflection principle of van Fraassen (1984), which is about learning or favorable epistemic changes in general, by a new, but similar backward-looking principle, which is about forgetting and other unfavorable epistemic changes. The generalization is argued to be a completion by defending what is called the no-neutrality condition. Due to the similarity, analogous consequences are provable for both principles. This fact is utilized for a plausibility check of the new principle. Finally, it is argued that this generalization should not be considered as a special case of the expert principle.

Van Fraassen (van Fraassen 1984, 244) has introduced the *reflection principle*, as he called it: *Given* your tomorrow's probabilities are such and such, your present conditional probabilities should be the very same. Hence, this principle may be called *forward-looking*. It is *the* basic principle of probabilistic or Bayesian auto-epistemology. Various interesting consequences have been derived from it. It has met diverse criticisms, several variants have been offered in response, and it is well-known by now that it holds only under restrictions. However, the literature does not offer clear ideas whether and how it may be suitably amended or even completed. This paper tries to do better by proposing what I call the full reflection principle. It has certainly been suggested in one way or another. But I am after precise statements allowing strict inferences.

Here is how the paper proceeds: Section 1 briefly recapitulates van Fraassen's original principle. Section 2 suggests an equally strong and formally analogous *backward-looking* principle. Both combine to what I call the full reflection principle. Thereby I propose to double, as it were, and arguably complete the range of auto-epistemology. In section 3 the completeness claim is supported by what I call the no-neutrality condition. However, the completeness claim does not go so far as to offer an account

of how to iteratively apply the full principle. Section 4 looks at some well-known consequences of van Fraassen's principle, which carry over to the backward-looking principle. This is intended to serve as a plausibility check of the proposed generalization. Section 5 discusses what happens if the reflection principles are referred to auto-epistemic propositions themselves; this is often not strictly distinguished. There we will discover a slight disanalogy between the forward- and the backward-looking principle. Section 6, finally, defends my generalization against the objection that it is already contained in the familiar generalization of van Fraassen's principle known as the expert principle.

1 Van Fraassen's Reflection Principle

This paper intends to focus on the philosophical and not on the formal aspects of its topic. Still, the main claims should be stated in a formally precise way. For this purpose, we need a modicum of notation. We will refer to a prior time, 0, and a posterior time, 1. This may be today and tomorrow, or yesterday and today; we will use both readings. Everything indexed in this way will refer so as well. Thus, P_0 is to represent your actual prior and P_1 your actual posterior credences, where I suppose you to satisfy all rationality constraints we think to be pertinent. These are at least the synchronic axioms of probability and some diachronic learning rules we need not fix. Without an index the temporal reference may be any. I use π, with or without indices, as a variable for your possible probability functions. Sets of those functions represent *auto-epistemic* propositions about your own probabilities; e.g. $\{\pi_0 \mid \pi_0(A) = x\}$ represents the proposition that your prior probability for A is x, $\{\pi_1 \mid \pi_1 = Q_1\}$ represents the proposition that your posterior probability function is Q_1, etc.

These possible probability functions are about some fixed algebra of propositions concerning some worldly affairs. Not all worldly affairs; the range may be quite restricted, but need not be made explicit. The label "worldly" is to mean that these propositions are about external matters, and not about your epistemic states. (In section 5 we will consider dropping this restriction.) It also means that the propositions are about empirical, not about abstract or formal matters. The epistemology of formal sciences is a very different topic not amenable to the present methods in my view.

However, your actual P_0 and P_1 are not only about these worldly propositions, but also about all the auto-epistemic propositions just introduced and

all algebraic combinations thereof. Thus, P_0 and P_1, but not the possible π, reflect on how your prior and posterior probabilities might be.

Now we are ready to explicitly state van Fraassen's principle; I will call it the *forward-looking reflection principle*, for quite obvious reasons:

(1) $P_0(A \mid \{\pi_1 \mid \pi_1 = P_1\}) = P_1(A)$ for all worldly propositions A.

That is, *given* your posterior credence is P_1 and in particular $P_1(A)$ for A, your prior for A is also $P_1(A)$.[1] Why do I state the condition in (1) as $\{\pi_1 \mid \pi_1 = P_1\}$ and not in the usual way as $\{\pi_1 \mid \pi_1(A) = P_1(A)\}$? Because the usual version is weaker and derivable from the present version, which is clearly the intended one.[2] Note, moreover, that "$\pi_1 = P_1$" cannot be literally true, given that π_1 is only about worldly propositions, while P_1 is also about auto-epistemic propositions. Strictly speaking, I should refer to P_1 as restricted to worldly propositions. This would be too cumbersome, though. This little sloppiness will do no harm.

Let me emphasize that I am using here probabilistic terminology only for convenience; it is the most familiar one. However, the reflection principle is not restricted to Bayesian epistemology. It holds for any kind of epistemic format that allows of conditional epistemic states. Binkley (1968) has proposed a qualitative version: if you now believe that you will believe p tomorrow, you should believe p already now. This principle played a crucial role in his account of the surprise examination paradox. In Spohn (2010, 125–127) I have stated the reflection principle in terms of ranking theory. One may state the reflection principle in terms of imprecise or interval probabilities. And so on.[3] Here, I will simply assume that epistemic states are represented in a fixed formal format, and then I choose the most familiar one. I do not want to

1 In order to keep things simple, this statement is predicated on the assumption that there are only finitely many possible π under consideration. Generally, though, there are uncountably many π in play, and then the condition is likely to have probability 0 so that the conditional probability in (1) is undefined. Then we must replace (1) by Constraint 2, as (Skyrms (1980, 163)) calls it, which says for all intervals I: $P_0(A \mid \{\pi_1 \mid \pi_1(A) \in I\}) \in I$. This solves the problem, because now the condition will usually have positive probability. See also Goldstein (1983).

2 For instance, from (1), but not from the usual version, one can derive $P_0(A \mid \{\pi_1 \mid \pi_1(A) = x \,\&\, \pi_1(B) = y\}) = x$, which is desirable, too.

3 Schoenfield (2012) argues that the reflection principle *entails* probabilities to be precise—and is therefore to be rejected, because other arguments require probabilities to be imprecise. If so, one should state, I think, a reflection principle for interval probabilities: $P_0(A \mid \{\pi_1 \mid \pi_1(A) = I\}) = I$, where P_0 and the π_1 would now be interval probability functions. (Observe the difference to 'Constraint 2' in footnote 1.)

doi: 10.48106/dial.v77.i1.04

discuss the more basis issue of the most adequate representation of epistemic states. Hence, the entire paper will move within a standard probabilistic representation of epistemic states, though only *pars pro toto*.

I shall not repeat here the grounds on which (1) is widely accepted, I shall only refer to some criticism below. I should mention, though, that (1) has been considered before. Spohn (1978, 161f). and Goldstein (1983) have proposed the iteration principle (which is almost equivalent, see sections 4 and 5). Quite generally one might say that the reflection principle is implicit or well-nigh explicit in Finetti (1937)'s philosophy of probability. One incurs considerable philosophical costs if one wants to abandon it.

Gaifman (1986) has suggested a more general reading of (1). He defines P_1 to be an expert for you at the prior time 0 (regarding A) if and only if (1) holds for your P_0. In this reading (1) turns into what is called the *expert principle*, saying that you should trust the experts—though this is now tautological, because an expert has been defined as one you can trust in this way. Still, the maneuver provides quite a graphic reading of the original (1), namely as saying that you should accept your posterior opinion as an expert for you. [4] In section 6 I shall discuss whether the amendment of (1) proposed in section 2 may simply be conceived as an instance of this expert principle.

This makes (1) plausible. Surely, if your posterior probabilities have learned something, have gathered evidence, are better informed, etc., then your prior probabilities should consider them to be an expert for you. At the same time this points to the restricted applicability of (1). Your future self is not always better informed. You may forget things, you may be tired, brain-washed, confused by drugs, your judgment may be obfuscated by prejudices against better knowledge, etc. In all those cases your future opinion is not a trustworthy expert for you. This restriction has often been noticed, e.g. by Christensen (1991)) (referring to epistemic change due to psychedelic drugs), by Talbott (1991) (referring mainly to memory loss) or by Spohn (1978, 166) (with re-

4 The striking similarity between the reflection principle and Lewis' Principal Principle has often been noticed; see e.g. Spohn (2010). In these terms the Principal Principle says that chance at t is your best expert at t, and truth = chance at the end of time is your best expert at all. Christensen (2010b) discusses a principle of rational reflection, which is about your probabilities conditional not on your future probabilities, but on your current probabilities as they should rationally be. This differs from the original reflection principle only in case you suspect your future probabilities to diverge from how the current probabilities should rationally be. In any case, I stick to the original principle and won't discuss such subtly related principles. The real challenge raised by Gaifman is, of course, what to do when I am impressed by several (diverging) experts. Again, peer disagreement is not my topic here.

spect to the iteration principle). Briggs (2009, 64ff).) presents a nice list of exceptions.[5]

In my view, this restriction does not diminish the importance of the reflection principle (1). Whenever your epistemic changes, be they rational or not, do not just occur to you and you rather take a reflective attitude towards them, (1) serves as a fundamental meta-principle. It does not specify how learning precisely works. There we may consider rules like simple or Jeffrey conditionalization, minimizing relative entropy, etc. But it says, however learning works, it must satisfy (1), if reflected.

Surely, though, (simple) conditionalization is the basic Bayesian learning rule. Since van Fraassen (1984) proposed a Dutch book argument in favor of his reflection principle, critical discussions like those of Christensen (1991) and Talbott (1991) focused on the tension of this argument with the familiar Dutch book justification of conditionalization. Indeed, the relation between the reflection principle and the conditionalization rule is delicate, as displayed also in Weisberg (2007). However, (Hild (1998a) Hild (1998b)) has already shown that the reflection principle is equivalent to a rule he calls auto-epistemic conditionalization, and he has fully stated the conditions under which auto-epistemic conditionalization and simple and Jeffrey conditionalization may come apart. Therefore, I shall not pursue this connection any further.

2 A Generalization of van Fraassen's Reflection Principle

If the forward-looking principle (1) has only restricted validity, we should think about whether there are auto-epistemic principles governing the cases not covered by (1). These may be irrational epistemic changes violating diachronic principles of rationality that are supposed to be governed by the reflection principle (1), whatever they are. E.g. whenever your prejudices get a hold on you, although you are sure that they are irrelevant, this is arguably a case of epistemic irrationality. However, there may also be arational epistemic changes, which do not violate diachronic principles, but simply fall outside the scope of such principles, such as fatigue, or clouding one's judgment by getting drunk (where the only irrationality may have been to get drunk), etc. Forgetting is perhaps the most familiar case of such an arational

5 van Fraassen (1995) takes the strange recourse to say that in such cases your epistemic self ceases to exist, thus recovering general validity of (1) for *you*.

doi: 10.48106/dial.v77.i1.04

change. However, it consists not only in unlearning a fact. It can take many shapes. E.g. one may forget live possibilities so that one becomes sure of the remaining possibilities. Or one can forget about evidential relations, say, that a characteristic smell is a sign of a poison. And so on.

Such changes seem to be a matter of empirical psychology and not of a normative theory of epistemic rationality. Isn't it an empirical question what we are prone to forget, how our beliefs are influenced by prejudices or drugs, etc.? Yes. However, this does not mean that a theory of epistemic rationality cannot say anything about how we should *rationally deal* with such arational and irrational changes, when we, possibly falsely, suspect them to occur. On the contrary, if we are thinking about how to extend the reflection principle (1), this is the challenge we positively face. Hence, the question is: Is there an auto-epistemic rationality principle dealing also with those irrational and arational changes?

Yes, there is. I follow the idea of Titelbaum (2013, ch. 6): just reverse the temporal perspective![6] Place yourself at your posterior P_1 and ask whether you should consult your prior P_0, whatever it was. Certainly not in van Fraassen's cases where you are better informed in P_1. But surely in those cases where you have forgotten something, are foggy-brained, etc. in P_1. Therefore, I propose the following *backward-looking reflection principle*:[7]

(2) $P_1(A \mid \{\pi_0 \mid \pi_0 = P_0\}) = P_0(A)$ for all worldly propositions A.[8]

6 I have first proposed an essential part of the following considerations in Spohn (2017, sect. 10) in a more convoluted way at the end of a long investigation on indexical belief. The connection was Sleeping Beauty, which seems to be not only a problem about indexical belief, but also about auto-epistemology. It has been observed, e.g. by Arntzenius (2003) or Mahtani (2017), that the proposition on which the probabilities in the reflection principle are conditioned must egocentrically refer to *me* and *my future* probabilities, not to *a*'s probabilities at *t*, where *a* happens to be me and *t* happens to lie in the future. Here I ignore this line of thought. The subject's potential uncertainty about her own location is not our issue.

7 Christensen (2000, 352ff). also speaks of taking a backward-looking perspective, but he thereby means something else. He is interested in diachronic coherence in the sense of an epistemic conservativism, which seeks to preserve as much of the old beliefs as possible while *learning* something new.

8 As just mentioned, Titelbaum (2013) has already observed the symmetry between the forward- and the backward-looking case. He captured both in his principle of Generalized Conditionalization (p. 127). I should have known and noticed this in Spohn (2017). However, he refers his observation to simple conditionalization and conceives of the prior state as the conditionalization of the posterior forgetful state with respect to the forgotten proposition. On p. 133 he also considers a generalized reflection principle, but only as derived from simple conditionalization. By contrast, I think of the reflection principles as having an independent role and apply the backward-looking

(2) is as important and fundamental a meta-principle as (1). Whatever the multifarious changes of our epistemic position to the worse, if we reflect on them, (2) must be obeyed in all such cases. In a decision theoretic perspective, which I do not unfold in this paper, this meta-principle is at the bottom of our efforts to fight forgetting, e.g. by building museums and archives and even inventing scripture, and of our attempts to preserve our epistemic integrity wherever we can, e.g. by banning brain-washing.

It is clear, however, that not both, (1) and (2), can be applied generally. Their applications cannot even overlap. In a case of such overlap, your prior probability should trust your posterior one, but your posterior one should reversely follow your prior one; so, if both are mutually envisaged, they must be the same. In other words, only in the case of non-change can (1) and (2) apply simultaneously. In section 5 I will provide a formal proof of this claim.

Hence, the fields of application of (1) and (2) must be disjoint (with the possible exception of non-change). Indeed, this is how I have explained (1) and (2). In order to have uniform labels, let's say that (1) applies in the case of *favorable* changes such as learning, receiving information, acquiring evidence, etc., and that (2) applies in the case of *unfavorable* changes such as forgetting, drinking too much, being influenced by prejudices one takes to be unjustified, etc. In a case of a favorable change from P to P' or an unfavorable change from P' to P let's say that P' is *superior* to P and P is *inferior* to P'.

The next question is: who is to judge changes as favorable or unfavorable and epistemic positions as superior or inferior? We who decree these principles from outside? No, I think it is preferable to subjectivize the application conditions of (1) and (2). The subject herself must assess the epistemic changes she is considering: Is my change from P_0 to P_1 favorable and P_1 superior? Then apply (1)! Is my change from P_0 to P_1 unfavorable and P_0 superior? Then apply (2)! The first instance to assess this is the subject herself.

Of course, this does not preclude that, in a second step, we have a normative argument about this assessment. Presumably, we all agree that experience is favorable and forgetting is unfavorable. But what about hunches and gut feelings? Gigerenzer (2007) is a strong plea for respecting gut feelings not only as a psychological fact, but also as a guide-line to rational decision making. Prejudices may not always be bad. What about epiphanies? Those claiming having had them feel to be in a superior position. Surely, these are difficult

principle to all kinds of arational and irrational epistemic changes, not just to the forgetting of previous certainties, as Titelbaum does by only dealing with simple conditionalization. Insofar my approach is more general.

doi: 10.48106/dial.v77.i1.04

and possibly contested issues that we may and must discuss. However, the reflection principles as such are independent from that discussion, and therefore we should keep matters separate. The intent of my subjectivizing move was precisely not to get involved into that discussion.

Let's take a slightly more general perspective for stating it. The temporal relations do not really seem to matter. The point is rather that in whatever epistemic position I am I would trust a superior position and mistrust an inferior position (where it is up to me what I take to be superior and inferior). This seems to be the gist of the principle. If so, we arrive at the following *full reflection principle*, in which the temporal location of the probabilities referred to is left open (hence no indices):

(3) for all worldly propositions A $P(A \mid \{\pi \mid \pi = P'\}) = Q(A)$, given that Q is taken to be the superior one of P and P'.[9]

This is the generalization of van Fraassen's auto-epistemology I would like to propose.

3. The No-Neutrality Condition and the Iteration Problem

Is the generalization a completion? This raises two issues. First, the principle (3) reflects only upon a single possible change. But we may certainly reflect on iterated change. (3) is silent on this and insofar still incomplete.[10] I will not be able offer a solution, but I will comment on the issue below.

Secondly, if we attend only to a single change, the reflection principle (3) is complete only if the *no-neutrality condition* holds which states that there are no neutral and no incomparable changes; there are only either favorable or unfavorable changes and nothing besides. Then, but only then, there would always be *the superior one* of P and P' (with the irrelevant exception of non-

9 The "is taken" always refers to the assessment of the subject we are talking about, not our own. In Spohn (2017), I have emphasized this by making the condition in (3) part of the conditional probability statement. However, this raises awkward questions. Are propositions of the form "P' is superior to P" part of the auto-epistemically extended algebra of propositions? Do they receive probabilities? Are these probabilities subject to change? We better avoid such questions. These propositions are outside the epistemic game we are considering. We may rather assume that the subject's superiority assessments are stable within our dynamic scenario. Therefore, I now state this condition outside the probability statement, though still in a subjectivized form.

10 The importance of this issue is underscored by the parallel case in belief revision theory, which was initially restricted to treating only single revisions and thus plagued by the iteration problem, too. I have first raised it in Spohn (1988, 112ff.). It turned out to virtually be an anomaly in the Kuhnian sense; see e.g. Rott (2009).

change), and the application condition of (3) would be complete. Does this condition hold?

Yes, I think so. I welcome favorable changes and seek superior epistemic positions (if they are not too costly) and I try to prevent unfavorable changes and to avoid inferior positions (if that is not too costly, either). At least this is so by purely epistemic standards; moral standards, e.g. may tell otherwise in special cases. Thus, a change which is neither favorable nor unfavorable would be one I don't care about. I would say then: it's nice to have the present prior P_0, and it's equally nice to have the posterior P_1 ? P_0 later on; both are fine and none of them is inferior or superior. This sounds very strange to me. This makes the change from P_0 to P_1 appear arbitrary and without good reason, and then I can't stay indifferent about the change; it must appear unfavorable to me.

To illustrate: Today I think I will be in good health next year, and tomorrow, just over sleeping and without any new information whatsoever, I think I won't. Usually, this would not be taken as a change of mind, but as an expression of a continued uncertainty. But say, today I am firmly convinced that I will be in good health next year. From this perspective it must appear arbitrary when I would have changed by tomorrow to equally firmly believing the contrary. It would be odd to presently be neutral about such a change; I should rather reject and not trust it.[11]

This is not a cogent argument. It is only to say that I cannot imagine how the no-neutrality condition could be violated. In any case, one must be aware that this condition is a crucial and substantial normative principle. If we accept it, then (3) indeed deserves the label "full reflection principle", at least regarding single changes.[12]

I think, though, that there is a deeper reason behind the no-neutrality condition. It is that ultimately there is only one standard for our epistemic states: truth. We try to approach truth and to avoid veering away from truth, however we measure the distance here. The point is that there is only one 'scale' to measure. If epistemic states would have to meet many standards on different scales, then indifferences or even incomparabilities might easily arise. Such more complex situations would certainly be relevant when we were to more generally think about what kind of person we want to be. There are many aspects in which we change for the better or the worse, and we

11 See also the arguments against arbitrary switching in White (2014, 318ff).

12 As mentioned, the case of non-change from P to P' may be ruled arbitrarily. We may say then that P is superior, or P' is, or both are. It doesn't make any difference for (3).

doi: 10.48106/dial.v77.i1.04

will often have indeterminate preferences about possible personal changes or none at all. But in the case of epistemic change our judgments seem to be unambiguous.

So far, I have only argued that there are no neutral changes. And I have excluded the possibility of incomparabilities due to a multitude of epistemic standards. However, there are easier ways for incomparabilities to arise. Surely, there are complex changes which are favorable in some respects and unfavorable in others so that, overall, the result is neither superior nor inferior, but incomparable. For instance, I learn that I have a date with the president next Friday and simultaneously forget that I have already agreed to meet the vice-president at the very same time. I propose to treat this as two changes, first a favorable and then an unfavorable one—or the other way around; it is not guaranteed that this comes to the same. Often, a temporal succession can be discerned within such a complex change, and sometimes, e.g. in my example, this move may be artificial. However, my proposal seems feasible, it avoids the need to refer favorability and unfavorability to aspects of complex changes, and it saves the no-neutrality condition. So, in any case, it is theoretically beneficial.

However, this move makes the first issue of extending the full reflection principle (3) to iterated change more pressing. To my knowledge, this issue has not been considered in the literature. Perhaps the reason is that it seemed trivial in the case of the original reflection principle. If my first epistemic state trusts the second, and the second trusts the third, already the first state can trust the third. Reversely in the case of iterated forgetting.[13] However, there are also mixed cases, and I have just alluded to them.[14] The difficult case is the one where my epistemic state first changes in an unfavorable way and then in a favorable way; e.g. first I forget some things and then I learn other things. In this case, my initial state can neither trust in the final state in the sense of principle (1), nor can it dismiss the final state in the sense of principle (2). Rather, it seems that I have to engage in a counterfactual consideration. In this case I can only trust that epistemic state that *would have* emerged had I not incurred the first unfavorable change (forgotten the one things), but still experienced the second favorable change (learned the other things). That is, I

13 Titelbaum (2013) seems to be able to treat the iterated case with the help of his principle of suppositional consistency (p. 140). But if so, this is due to the fact that the only epistemic changes he considers is the gain and loss of certainties.

14 I have discussed the various cases and their problems a bit more extensively in Spohn (2017, 408f).

would have to speculate not only about my actual epistemic states and their change, but also about my counterfactual epistemic states and their change. Hence, a general solution of this problem seems to require quite different theoretical means. It is not a task we can pursue here.

Still, it should be pursued. To emphasize its urgency: As far as I see, the issue of so-called second- or higher-order evidence is closely related. Christensen (2010a) gives a wide variety of examples. A salient structure of them (not all of them) is this: I receive a lot of ordinary (first-order) evidence on a certain matter, and I seem to draw my conclusions from it in the usual rational way. At the same time, I receive higher-order evidence (perhaps falsely) indicating that my cognitive abilities are somehow hampered. I am overly tired, I am told to have consumed a fancy drug, I am instructed that I regularly tend to overoptimism, I may be suffering from hypoxia (a realistic example from Christensen (2010b, 126)), etc. So, maybe I should correct my inferences?

In such cases, the higher-order evidence indicates that I should not trust the epistemic state I have reached. But neither can I simply rely on my prior epistemic state before the change, as the backward-looking reflection principle (2) would have it. As above, such cases are mixtures of two different epistemic movements. On the one hand, there is the first-order evidence which I should trust. On the other hand, there is my alleged epistemic handicap which suggests an accompanying unfavorable change and should make me think about what my inferences would have been without the handicap. So, the issue of higher-order evidence would also profit from solving the iteration problem. However, for the reasons indicated, I don't further address this problem. We should be content with treating the pure cases.[15]

4. Are the Consequences of The Full Principle Acceptable?

The full reflection principle (3) seems to be intuitively plausible and philosophically important. If so, we should also check for its consequences, or at least some of them. The consequences of the original principle (1) are well known. We may follow here Hild (1998a), but need not do it very far. Formally, the consequences of the generalized (3) are obviously analogous. So, the strategy in this section will be to develop the formal analogy and to check whether the results are also intuitively acceptable.

15 In response to such examples Briggs (2009, 71) proposes a principle of distorted reflection: $P_0(A \mid \{\pi_1 \mid \pi_1(A) = x\}) = x - D$, where D is a factor measuring the "expected departure from conditionalization on veridical evidence" (regarding A). This may be a correct qualification. But again, it takes two steps at once, and we should first get clear about the single steps.

doi: 10.48106/dial.v77.i1.04

The first thing to do, perhaps, is to unpack again what we have packed into the condensed abstract statement of (3). It contains in fact five different principles, depending on the temporal and the superiority relations between P and P' (where, once and for all, each principle quantifies over all *worldly* propositions A).

One case is where $P = P_0$ is the prior and $P' = P_1$ the posterior. If P_1 is superior to P_0, we get:

(3a) $P_0(A \mid \{\pi_1 \mid \pi_1 = P_1\}) = P_1(A)$, given that P_1 is taken to be superior to P_0.

This is our original forward-looking reflection principle (1). Since the proviso takes care of the main objections against the principle, there is no need to further discuss it.

However, P_1 may also be inferior to P_0. Then we get something we have not yet explicitly stated:

(3b) $P_0(A \mid \{\pi_1 \mid \pi_1 = P_1\}) = P_0(A)$, given that P_1 is taken to be inferior to P_0.

Given tomorrow's inferior opinion I stick to my prior opinion. E.g. today I believe that I have a date with my dentist on Tuesday next week. It is only reasonable, then, to stick to this belief, given I am confused tomorrow and think the date is next Wednesday. Christensen (1991) imagines an agent having swallowed a hefty dose of a certain drug and then being asked: "What do you think the probability is that you'll be able to fly in one hour, given that you'll then take the probability that you can fly to be .99?" (p. 234). He answers in place of the agent: "The sane answer to the above question is clearly one that gives a very low probability to the agent's ability to fly one hour from now, even on the supposition that she will at that time give it a very high probability" (p 235). This is clearly an instance of (3b). Hence, Christensen may be said to have anticipated the intention of the full principle (3).

Another case is where $P = P_1$ is the posterior and $P' = P_0$ the prior. Again, this splits up into:

(3c) $P_1(A \mid \{\pi_0 \mid \pi_0 = P_0\}) = P_0(A)$, given that P_1 is taken to be inferior to P_0,

which is our backward-looking reflection principle (2). And into:

(3d) $P_1(A \mid \{\pi_0 \mid \pi_0 = P_0\}) = P_1(A)$, given that P_1 is taken to be superior to P_0,

which we have not yet explicitly stated, either. It says that, given my prior opinion, I stick to my posterior opinion, if it has been acquired through a favorable change. Again, this seems to go without saying.

There is finally the case where P and P' refer to the same time, so that $P = P' = P_i$ ($i = 0, 1$), where we may, as mentioned, define the superiority relation either way. Thereby we get a *synchronic reflection principle*, which is independent of the previous diachronic principles:

(3e) $P_i(A \mid \{\pi \mid \pi = P_i\}) = P_i(A)$ $(i = 0, 1)$.

This may look odd. But it simply says that your present opinion, whether prior or posterior, is presently an expert for you. You presently don't know better than you actually know. Of course, this does not preclude that you accept other hypothetical experts as well. Since it seems to differ from the other principles, it has also received a separate discussion. We need not go into it now.[16]

What is the relation between the five parts (3a – e) of the full reflection principle (3)? As far as I see, they are independent. As already observed, the synchronic principle (3e) must be independent from the other diachronic principles. Moreover, the principles (3a + d) for favorable changes and the principles (3b + c) for unfavorable changes are independent as well, simply because they refer to disjoint conditions. Maybe the two principles about favorable changes are related? And likewise those about unfavorable changes? However, I have not discovered any relation and I think that the five parts (3a – e) are indeed independent. In the next section, however, I shall indicate how things may change.

Let's look at some consequences of the original reflection principle (1), or (3a), in order to check whether their formal generalization is also intuitively plausible. The first is the *iteration principle* already mentioned:

(4a) for all worldly propositions A $P_0(A) = \Sigma\, \pi_1(A)$? $P_0(\pi_1)$, where $P_0(\pi_1)$ is the subject's prior auto-epistemic probability for π_1 being her posterior, where the sum is taken over all her possible posteriors π_1 taken to be superior to P_0, and where $\Sigma\, P_0(\pi_1) = 1$, i.e. P_0 is sure to undergo a favorable change.

(4a) is entailed by (3a)[17]; for a possible reversal see below. In other words, your prior opinion is always a weighted mixture of all the posterior opinions you may favorably reach, where the weights are given by your prior opinion

16 van Fraassen (1984, 248) takes it to be "uncontroversial." However, Christensen (2007) after calling (3a) a principle of epistemic self-respect and quoting a lot of support for it (pp. 322f.), puts forward putative counter-examples. They trade, I think, on a subtle ambiguity of the inner and outer P_i in (3a). There, the outer P_i has some (second-order) information about the inner P_i and thus the two may come to diverge.

17 If we suppress the additional condition about superiority and stick to the shorter notation used in (4a), (3a) says $P_0(A \mid \pi_1) = \pi_1(A)$. The formula of the total probability says that $P_0(A) = \Sigma\, P_0(A \mid \pi_1)$? $P_0(\pi_1)$. By inserting the first equation into the second we get (4a).

doi: 10.48106/dial.v77.i1.04

about reaching these posteriors.[18] This is, I think, a deep epistemological insight.

In the same way, the backward-looking reflection principle (2), or (3c), entails the following *reverse iteration principle*:

(4b) for all worldly propositions A $P_1(A) = \Sigma\, \pi_0(A)\, ?\, P_1(\pi_0)$, where the sum is taken over all possible priors π_0 taken to be superior to P_1 and again $\Sigma\, P_1(\pi_0) = 1$, i.e. P_1 is sure to have undergone an unfavorable change.

The proof is analogous to the one of (4a). Is (4b) plausible? Yes. If, in your posterior P_1 you have forgotten about something, you will usually not remember, either, what your past opinion about that thing has been. Still, you might auto-epistemically wonder what your past opinion has been. And this guess work is coherent only if it satisfies (4b). For instance, you cannot coherently say: "Oh, I have forgotten my date with the dentist; I guess it's next Wednesday. But I think that yesterday I was still quite sure that it is next Tuesday." You may reversely take this as support for the backward-looking reflection principle (2).

One may think that there is a difference between (4a) and (4b). (4a), but not (4b), is grounded, as it were, in experience. The standard instantiation of (4a) is simple conditionalization: You learn exactly one member of a partition of evidential (worldly) propositions, about which you have some prior expectations. And since your possible posteriors are just the conditionalization of your prior with respect to these evidential propositions, you have the very same expectations about these posteriors. Such a grounding is, however, entirely missing in the case of (4b). You may remember your past probabilities, but if you have forgotten them, your present opinion about them is mere guesswork without such grounding.

However, the case is not as asymmetric as it seems. (4a) is not only made for simple conditionalization. It holds as well, e.g. for Jeffrey conditionalization, where learning results in some new posterior probabilities for the partition of evidential propositions. And then the posterior is not grounded in a specific evidential proposition, but in your possibly vague seemings concerning this evidential partition. Then, however, your expectations about these seemings are not much better off than in the past-oriented case. So, (4a + b) is auto-epistemic business justified by full reflection (3). Such specific grounding is welcome, but not required.

18 Here, de Finetti's heritage is particularly salient. Recall his famous representation theorem saying that your (prior) probabilities are symmetric or exchangeable (as they should be) if and only if they are a unique mixture of all the statistical hypotheses they might converge to.

5. Reflection Applied to Auto-Epistemic Propositions

To check out further consequences, we must attend to the way how Hild (1998a) presents the principles. He tacitly assumes an innocent-looking generalization; i.e. from the outset he applies van Fraassen's principle (1) also to auto-epistemic (and mixed) propositions A. Let's briefly consider in this section what happens when we thus generalize our principles (3a – e) and drop their restriction to worldly propositions. Note that this also requires us to drop the restriction of the possible probability measures π to worldly propositions. Hence, equations like $\pi = Q$ can now be literally and not only sloppily true.

A first consequence would be that the synchronic reflection principle (3e) turns out to be equivalent to what Hild calls *auto-epistemic transparency*:[19]

(5) $P_i(\{\pi \mid \pi = P_i\}) = 1$ $(i = 0, 1)$.[20]

In other words, in each second-order epistemic state reflecting also on your first-order state you know, or are sure, what your present first-order state is. In doxastic logic, this is sometimes called 'positive introspection' or the BB thesis "if you believe that p, then you believe that you believe that p", first discussed in Hintikka (1962, 123ff)., amply attacked, and amply defended. Let us not engage in this discussion now.[21]

Another consequence of the reflection principles (3a + e) thus extended is *perfect memory*:

(6) $P_1(\{\pi_0 \mid \pi_0 = P_0\}) = 1$, given that P_1 is taken to be superior to P_0.[22]

19 Christensen (2007) and Weisberg (2007) split this up into two principles called confidence and accuracy by Christensen and luminosity and transparency by Weisberg.

20 *Proof:* For the one direction, take A in (3e) to be the auto-epistemic proposition $\{\pi \mid \pi = P_i\}$. Reversely, P_i is identical with P_i conditional on a proposition with probability 1.

21 Besides the arguments referred to in footnote 15, Christensen (2007) casts doubt on auto-epistemic transparency by questioning that we have certain knowledge about our precise subjective probabilities. However, this rather questions the representation of epistemic states by precise probabilities, which we have assumed at the outset of this paper. That is, I tend to assume that beliefs and epistemic states in general are conscious mental states (in the sense of intentional or higher-order consciousness), similar to phenomenally conscious pains. So, if I do not know my precise probabilities, I don't have them, just as I don't have pains when I am not aware of them. It is this assumption, I think, this is the motivation behind the BB thesis and its kin.

22 *Proof:* Take A in (3a) to be the auto-epistemic proposition $\{\pi_0 \mid \pi_0 = P_0\}$. Thus $P_0(\{\pi_0 \mid \pi_0 = P_0\} | \{\pi_1 \mid \pi_1 = P_1\}) = P_1(\{\pi_0 \mid \pi_0 = P_0\})$, given that P_1 is taken to be superior to P_0. Auto-epistemic transparency (5) says that $P_0(\{\pi_0 \mid \pi_0 = P_0\}) = 1$. Hence, $P_0(\{\pi_0 \mid \pi_0 = P_0\} \mid \{\pi_1 \mid \pi_1 = P_1\}) = 1$ as well. So, finally, $P_1(\{\pi_0 \mid \pi_0 = P_0\}) = 1$, given that P_1 is taken to be superior to P_0. (Cf. Hild (1998a, 353).)

doi: 10.48106/dial.v77.i1.04

This is a suspiciously strong consequence. However, given the extension of the principles to auto-epistemic propositions, one could say that whenever I have become uncertain about my former epistemic state, I have forgotten my former attitude towards some proposition. So, my uncertainty must be the result of an unfavorable change.

By a similar proof, the extended (3c) + (3e) implies an analogous principle of perfect foresight. My prior knows for sure what my inferior posterior will be. This is obviously absurd. Hence, the auto-epistemic extension of the backward-looking principle (3c) must be rejected. The proof of (6) displays where the analogy breaks down. Inserting $\{\pi_1 \mid \pi_1 = P_1\}$ for A in (3c) would mean that I would trust my former superior P_0 concerning my present inferior state. But regarding my own present state, I am always in an optimal epistemic position, as confirmed by auto-epistemic transparency (5); in this respect I need no lessons from my better self. My position may be inferior only with respect to worldly propositions (and auto-epistemic propositions referring to other times). This is why we must not extend (3c) to auto-epistemic propositions (or at least not to the simultaneous ones).

The envisaged extension also helps a bit regarding the relations among the five parts of (3). That is, we may see now that the extended (3a + e) entail not only (6), but also (3d), simply because (6) says that the condition of (3d) has probability 1. For the same reason as before, we must not exploit this observation for a corresponding derivation of (3b) from (3c + e); (3b) seems to remain independent.

Moreover, we may note that Hild's extension strengthens the relation between reflection and iteration. We observed already that (3a) and (3c), respectively, entail (4a) and (4b). With the extension we may reverse the entailment: given auto-epistemic transparency (5) or the equivalent synchronic reflection principle (3e), the iteration principle (4a) implies the forward-looking reflection principle (3a).[23] Thus, under the same assumptions, the reverse iteration principle (4b) entails the backward-looking (3c), as seems unobjectionable.

Finally, this extension helps us to a formal proof of my informally justified claim in section 2 that the forward- and the backward-looking principle (1) and (2) can apply simultaneously only in the case of non-change. Given that (1) and (2) apply also to auto-epistemic propositions, we have:

23 *Proof*: We have $P_0(A \text{ and } \{\pi_1 \mid \pi_1 = P_1\}) = \Sigma\, \pi(A \text{ and } \{\pi_1 \mid \pi_1 = P_1\})$? $P_0(\pi)$ (by the extended (4a), where the sum is taken over all possible posteriors π) $= P_1(A)$? $P_0(\{\pi_1 \mid \pi_1 = P_1\})$ (by auto-transparency, because $\pi(\{\pi_1 \mid \pi_1 = P_1\}) = 1$ only for $\pi = P_1$, and otherwise $= 0$) (cf. Hild (1998a, 354)).

(7) Given auto-transparency (5), if (a) $P_0(A \mid \{\pi_1 \mid \pi_1 = P_1\}) = P_1(A)$ and (b) $P_1(A \mid \{\pi_0 \mid \pi_0 = P_0\}) = P_0(A)$ hold for all propositions A, then $P_0 = P_1$.[24]

In conclusion, we have found a slight divergence among our principles in this extension, a divergence we could justify. In the main, however, the parallel between the forward-looking and the backward-looking perspective and thus between the parts of the full reflection principle (3) stands. We have not discovered any incoherence.

6. The Full Reflection Principle and the Expert Principle

In section 1, I pointed already to the expert principle, which is the most common generalization of van Fraassen's principle (1). It may seem that the full reflection principle (3) is just another special case of the expert principle.[25] Yes, almost. At least, this holds for the backward-looking principle (2). Here, my better-informed past self may be taken to be an expert for my present forgetful self. However, not all cases of (3) are special cases of the expert principle. Let P in (3) be my probability measure and P' that of my neighbor. When I take my neighbor to be better informed, to be in a superior epistemic position (concerning a certain field), then I listen to her (in the sense of obeying (3)); these are the cases (3a) and (3c), where the temporal relation between me and my neighbor is irrelevant. But when I take her to be less well informed or in an inferior epistemic position, I do not listen to her; these are the cases (3b) and (3d). This seems to go without saying. Christensen (2000, 358). takes this for granted, too. Strictly speaking, though, it is not part of the expert principle, which says only how to deal with people taken to be at least as well-informed.

Of course, it would be no problem to pair the expert principle with a 'non-expert principle' saying that, rationally, we don't listen to persons we take to be in an inferior epistemic position. However, this would still leave us with very incomplete principles. The no-neutrality condition—which was plausible in the intrasubjective case, at least when we can divide up complex epistemic changes into unidirectional steps—has no analogue with respect to experts. Most of my fellow humans are neither better nor less well informed than me; their epistemic state is just incomparable to mine. And then both the expert and the non-expert principle are silent. This is not an objection. It is hard

24 *Proof*: We have $P_1(A) = P_0(A \mid \pi_1 = P_1)$ (due to (a)) $= P_0(A \mid \pi_1 = P_1, \pi_0 = P_0)$ (due to auto-transparency (5)) $= P_1(A \mid \pi_0 = P_0)$ (by applying (a) conditional on $\pi_0 = P_0$) $= P_0(A)$ (due to (b).

25 I am grateful to a reviewer for raising this issue.

doi: 10.48106/dial.v77.i1.04

to give any recommendations for the incomparable cases. But it is our daily business to somehow deal with them.

One may think[26] that we can apply the treatment of intrasubjective incomparabilities suggested in section 3 also to the interpersonal case of experts. However, this is not so easy. In the intrapersonal case, we had to refer, it seemed, to counterfactual epistemic states which the subject would be in, had certain unfavorable changes not occurred. This might be manageable. In the case of my incomparable neighbor, however, the corresponding counterfactual question would be which favorable changes he would have to undergo and which unfavorable changes to avoid, till I could acknowledge him to be an expert, i.e. to be in an equal or superior epistemic position concerning the issue at hand. This is a much more sweeping and indeterminate counterfactual question. The point is that the superiority and inferiority of epistemic positions is clearly assessable on the basis of intrasubjective favorable or unfavorable changes. But it is very hard to assess as such, as an interpersonal comparison would require.

So, we have a tension here. While van Fraassen's principle (1) was clearly a special case of the expert principle, the subsumption of the full principle (3) is at least doubtful. In fact, such a subsumption was not intended in the beginning. Gaifman (1986) proposed the expert principle as a formal generalization of the reflection principle, not as a substitute of the latter's epistemological role. As such it operates only in so-called time-slice epistemology (Moss (2015)) or time-slice rationality (Hedden (2015)). A basic assumption of this approach is called impartiality[27]: "In determining how you rationally ought to be at a time, your beliefs about what attitudes you have at other times play the same role as your beliefs about what attitudes other people have." (Hedden (2015, 9)) If so, it is clear that reflection principles are unnecessarily restrictive and that the expert principle completely takes over the epistemological role of the full reflection principle (given an additional 'non-expert principle').

In his defense of time-slice rationality, (Hedden 2015 chs. 8 – 9) makes crucial use of an assumption called uniqueness (by Feldman 2007): "Given a

26 Suggested by the same reviewer.

27 The other basic assumption is synchronicity: "All rationality requirements are synchronic." (Hedden (2015), [9]). See also Moss (2015, 177) for a statement of the two principles. The principle of impartiality seems to have been stated first by Christensen (2000, 363f).). Moss (2015, 178), and Hedden (2015, 56), happily observe that van Fraassen's reflection principle satisfies the assumption of synchronicity, insofar as it speaks only about P_0 ant its conjectures about P_1. Of course, this observation carries over to the other versions.

body of total evidence, there is a unique doxastic state that it is rational to be in" (Hedden (2015, 130)). This reveals an entirely different picture of normative epistemology than the one pursued here. It is that there is a unique prior—*the ur*-prior, as it were—and then all epistemic change is due to a change of the body of total evidence.[28] Given this, we do not need any diachronic rules for epistemic change. We always refer back to the *ur*-prior and then consider how the given body of total evidence operates on it. All change is in that body and only there; the body may get larger (through learning) or smaller (through forgetting).[29] My fellow humans are in the very same situation. They rationally proceed from the very same *ur*-prior; and they differ from me only in their total evidence. This is why they count just as much as my future or past epistemic states. Or rather, nobody counts; only the bodies of total evidence count. Certainly, this would simplify our epistemological business considerably.

The assumption of uniqueness also makes the notion of an expert very easy. Among rational subjects, a is an expert for b simply if a has at least as much evidence as b. If a's and b's evidence only overlap, they are in incomparable states. But joining their evidence would result in an expert for both. With this easy notion of an expert the above idea of paralleling intersubjective incomparabilities with intrasubjective incomparable changes might be less problematic.

The alternative is to deny uniqueness. Hedden (2015, 129) calls this permissiveness. But what is the dialectic situation here? Pace Hedden, it is not that the one must defend uniqueness and the other permissiveness (by showing two *ur*-priors to be equally rationally acceptable). In my view, the burden of proof lies with the defender of uniqueness. And a proof should constructively indicate how the unique *ur*-prior looks like. The literature is not so promising. The only positive attempt I know of is objective Bayesianism as proposed by Williamson (2005). [30] By contrast, I tend to take centuries of skepticism to suggest that such a proof will fail.[31] Obviously, this is too big an issue to be

28 Or the total evidence need not appeal to any prior at all. But then the *ur*-prior is unique as well, namely empty.

29 This picture also motivates Titelbaum (2013)'s framework of gain and loss of certainties.

30 To be precise, Williamson does not need to refer to an ur-prior. He rather proposes a unique way of responding to any given total body of evidence, which does not depend on an underlying ur-prior. Rather the ur-prior would be the response to the empty evidence.

31 Hedden (2015, 134) himself (as well as Kelly (2014, 309)) points to the alleged failure of Carnap's project of inductive logic, which also started out searching for the unique prior. I should add, though, that in ch. 8 Hedden admits that uniqueness may force one to allow for indeterminate

doi: 10.48106/dial.v77.i1.04

discussed now.[32] The point is only this: In the absence of such a proof we should not proceed from the assumption of uniqueness. A positive defense of permissiveness is not really required.[33]

Kelly (2014) usefully distinguishes statements of uniqueness that have interpersonal import, as he calls it, from those that have not. Intrasubjective uniqueness only requires that, given my background or my personal ur-prior I have only one rational way of responding to the evidence. I have no quarrel with this. However, uniqueness with interpersonal import requires that there is only one and the same rational way for everybody for responding to the evidence. This is the version intended and critically discussed above.

The question then is how to pursue normative epistemology without the assumption of uniqueness. Just in the way as it is done traditionally, and here as well, namely by stating synchronic principles of epistemic rationality and diachronic principles. The latter can only refer to a subject's prior and posterior and a piece of total evidence in between, but *not* to an *ur*-prior and a body of total evidence reaching back to the indefinite time of the *ur*-prior.[34] In this conception, the reflection principles as discussed here have a natural place, and the goal of stating a complete dynamic is important, while atemporal expert principles do not directly add to it and need not be complete. Given uniqueness, we can also distinguish inferior and superior epistemic positions, simply by looking at the size of the bodies of total evidence underlying them. However, when looking at epistemic dynamics in the traditional way, then, as argued, we must also classify single changes as favorable and unfavorable (and if we cannot do this objectively, we leave it to the subject herself, as proposed here). As mentioned, favorable and unfavorable changes do not only consist

and/or imprecise probabilities. However, considering other epistemic formats shifts the discussion still further. I have explained why I focus here on precise probabilities only.

32 I admit that the issue can also be discussed in the abstract without constructive proposals for an ur-prior. See e.g. the exchange between White (2014) and Kelly (2014). It is clear that my sympathies lie here with Kelly. However, he is still too obliging, I find; he does not raise the point about the burden of proof.

33 Recall also this: In the final section of Lewis (1980), Lewis discovers a tension between his Principal Principle and Humean Supervenience. He considers resolving the tension by assuming what is now called uniqueness. But he shies away from this solution which he finds "not very easy to believe." As is well known, he modified the Principal Principle later on.

34 When we model a dynamic process, in physics, meteorology or wherever, we do it in a form of a law saying how one state of the system modelled changes into the subsequent state, possibly under the influence of external factors, and we can do this in discrete or in continuous time. So, this is also the natural format for normative epistemology as well where we try to say what a rational epistemic dynamic should look like.

in gaining and losing evidence or certainties; they may take various other forms not easily subsumed under the picture motivated by uniqueness. In the perspective pursued here, expert principles become relevant only because we take listening to experts to induce favorable change, unlike listening to non-experts. This is why I think that the reflection principles have an independent value. They can be substantially subsumed under the expert principles only within a questionable epistemological picture.

References[*]

Wolfgang Spohn
0000-0002-3213-8907
Department of Philosophy
University of Konstanz
wolfgang.spohn@uni-konstanz.de

Arntzenius, Frank. 2003. "Some Problems for Conditionalization and Reflection." *The Journal of Philosophy* 100(7): 356–370, doi:10.5840/jphil2003100729.

Binkley, Robert W. 1968. "The Surprise Examination in Modal Logic." *The Journal of Philosophy* 65(5): 127–136, doi:10.2307/2024556.

Briggs, Rachael A. 2009. "Distorted Reflection." *The Philosophical Review* 118(1): 59–85, doi:10.1215/00318108-2008-029.

Christensen, David. 1991. "Clever Bookies and Coherent Beliefs ." *The Philosophical Review* 100(2): 229–247, doi:10.2307/2185301.

—. 2000. "Diachronic Coherence versus Epistemic Impartiality ." *The Philosophical Review* 109(3): 349–371, doi:10.1215/00318108-109-3-349.

—. 2007. "Epistemic Self-Respect." *Proceedings of the Aristotelian Society* 107: 319–337, doi:10.1111/j.1467-9264.2007.00224.x.

—. 2010a. "Higher-Order Evidence." *Philosophy and Phenomenological Research* 81(1): 185–215, doi:10.1111/j.1933-1592.2010.00366.x.

—. 2010b. "Rational Reflection." in *Philosophical Perspectives 24: Epistemology*, edited by John Hawthorne, pp. 121–140. Hoboken, New Jersey: John Wiley; Sons, Inc., doi:10.1111/j.1520-8583.2010.00187.x.

Finetti, Bruno de. 1937. "La prévision, ses lois logiques, ses sources subjectives." *Annales de l'Institut Henri Poincaré* 7: 1–68. Translated as Finetti (1964).

—. 1964. "Foresight: Its Logical Laws, Its Subjective Sources." in *Studies in Subjective Probability*, edited by Henry E. Kyburg Jr. and Howard E. Smokler, 2nd ed., pp.

[*] I am deeply indebted to three reviewers, one from the *Journal of Philosophy* and two from *Dialectica*. They considerably helped me in various respects. I also acknowledge support by the German Research Foundation, Grant EXC 2064/1, No. 390727645, and Grant SP279/21-1, No. 420094936.

93–158. Garden City, New York: Krieger Publishing Co. Translation of Finetti (1937), reprinted in Kyburg and Smokler (1980, 53–118), doi:10.1007/978-1-4612-0919-5_10.

GAIFMAN, Haim. 1986. "A Theory of Higher Order Probabilities." in *TARK 1986. Theoretical Aspects of Reasoning about Knowledge: Proceedings of the First Conference*, edited by Joseph Y. HALPERN, pp. 275–292. San Francisco, California: Morgan Kaufmann Publishers, Inc. Reprinted in Skyrms and Harper (1988, 191–220).

GIGERENZER, Gerd. 2007. *Gut Feelings: Short Cuts to Better Decision Making*. London: Penguin Books.

GOLDSTEIN, Matthew. 1983. "The Prevision of a Prevision." *Journal of the American Statistical Association* 78(384): 817–819, doi:10.2307/2288190.

HEDDEN, Brian. 2015. *Reasons without Persons. Rationality, Identity, and Time*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780198732594.001.0001.

HILD, Matthias. 1998a. "Auto-Epistemology and Updating." *Philosophical Studies* 92(3): 321–361, doi:10.1023/a:1004229808144.

—. 1998b. "The Coherence Argument Against Conditionalization." *Synthese* 115(2): 229–258, doi:10.1023/a:1005082908147.

HINTIKKA, Jaakko. 1962. *Knowledge and Belief*. Ithaca, New York: Cornell University Press.

KELLY, Thomas. 2014. "Evidence Can Be Permissive." in *Contemporary Debates in Epistemology*, edited by Matthias STEUP, John TURRI, and Ernest SOSA, 2nd ed., pp. 298–311. Contemporary Debates in Philosophy n. 3. Oxford: Wiley-Blackwell. First edition: Sosa and Steup (2005), doi:10.1002/9781394260744.ch12.

KYBURG, Henry E., Jr. and SMOKLER, Howard E., eds. 1980. *Studies in Subjective Probability*. 2nd ed. Garden City, New York: Krieger Publishing Co., doi:10.2307/2344211.

LEWIS, David. 1980. "A Subjectivist's Guide to Objective Chance." in *Studies in Inductive Logic and Probability. Volume II*, edited by Richard C. JEFFREY, pp. 263–294. Berkeley, California: University of California Press. Reprinted, with a postscript (Lewis 1986b), in Lewis (1986a, 83–113).

—. 1986a. *Philosophical Papers, Volume 2*. Oxford: Oxford University Press, doi:10.1093/0195036468.001.0001.

—. 1986b. "Postscript to Lewis (1980)." in *Philosophical Papers, Volume 2*, pp. 114–132. Oxford: Oxford University Press, doi:10.1093/0195036468.001.0001.

MAHTANI, Anna. 2017. "Deference, Respect and Intensionality." *Philosophical Studies* 174(1): 163–183, doi:10.1007/s11098-016-0675-6.

MOSS, Sarah. 2015. "Time-Slice Epistemology and Action under Indeterminacy." in *Oxford Studies in Epistemology*, volume V, edited by Tamar Szabó GENDLER and John HAWTHORNE, pp. 172–194. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780198722762.003.0006.

Schoenfield, Miriam. 2012. "Chilling Out on Epistemic Rationality: A Defense of Imprecise Credences (and Other Imprecise Doxastic Attitudes)." *Philosophical Studies* 158(2): 197–219, doi:10.1007/s11098-012-9886-7.

Skyrms, Brian. 1980. *Causal Necessity: A Pragmatic Investigation of the Necessity of Laws*. New Haven, Connecticut: Yale University Press.

Skyrms, Brian and Harper, William L., eds. 1988. *Causation, Chance, and Credence. Proceedings of the Irvine Conference on Probability and Causation, vol. 1*. The University of Western Ontario Series in Philosophy of Science n. 41. Dordrecht: Kluwer Academic Publishers.

Sosa, Ernest and Steup, Matthias, eds. 2005. *Contemporary Debates in Epistemology*. 1st ed. Contemporary Debates in Philosophy n. 3. Malden, Massachusetts: Basil Blackwell Publishers. Second edition: Steup, Turri and Sosa (2014).

Spohn, Wolfgang. 1978. *Grundlagen der Entscheidungstheorie*. Kronberg/Ts.: Scriptor Verlag.

—. 1988. "Ordinal Conditional Functions: A Dynamic Theory of Epistemic States." in *Causation in Decision, Belief Change, and Statistics. Proceedings of the Irvine Conference on Probability and Causation, vol. 2*, edited by William L. Harper and Brian Skyrms, pp. 105–134. The University of Western Ontario Series in Philosophy of Science n. 42. Dordrecht: Kluwer Academic Publishers, doi:10.1007/978-94-009-2865-7_6.

—. 2010. "Chance and Necessity: From Humean Supervenience to Humean Projection." in *The Place of Probability in Science. In Honor of Ellery Eells (1953–2006)*, edited by James H. Fetzer, pp. 101–132. Boston Studies in the Philosophy of Science n. 284. Dordrecht: Springer Verlag.

—. 2017. "The Epistemology and Auto-Epistemology of Temporal Self-Location and Forgetfulness." *Ergo* 4(13): 359–418, doi:10.3998/ergo.12405314.0004.013.

Steup, Matthias, Turri, John and Sosa, Ernest, eds. 2014. *Contemporary Debates in Epistemology*. 2nd ed. Contemporary Debates in Philosophy n. 3. Oxford: Wiley-Blackwell. First edition: Sosa and Steup (2005).

Talbott, William J. 1991. "Two Principles of Bayesian Epistemology ." *Philosophical Studies* 62(2): 135–150, doi:10.1007/bf00419049.

Titelbaum, Michael G. 2013. *Quitting Certainties. A Bayesian Framework Modeling Degrees of Belief*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780199658305.001.0001.

van Fraassen, Bas C. 1984. "Belief and the Will." *The Journal of Philosophy* 81(5): 235–256, doi:10.2307/2026388.

—. 1995. "Belief and the Problem of Ulysses and the Sirens." *Philosophical Studies* 77(1): 7–37, doi:10.1007/bf00996309.

Weisberg, Jonathan. 2007. "Conditionalization, Reflection, and Self-Knowledge." *Philosophical Studies* 135(2): 179–197, doi:10.1007/s11098-007-9073-4.

WHITE, Roger. 2014. "Evidence Cannot Be Permissive." in *Contemporary Debates in Epistemology*, edited by Matthias STEUP, John TURRI, and Ernest SOSA, 2nd ed., pp. 312–323. Contemporary Debates in Philosophy n. 3. Oxford: Wiley-Blackwell. First edition: Sosa and Steup (2005).

WILLIAMSON, Jon. 2005. *Bayesian Nets and Causality.Philosophical and Computational Foundations*. Oxford: Oxford University Press.