

LAWS ARE PERSISTENT INDUCTIVE SCHEMES

Wolfgang Spohn
Fachbereich Philosophie
Universität Konstanz
D-78457 Konstanz

Abstract: The characteristic difference between laws and accidental generalizations lies in our epistemic or inductive attitude towards them. This idea has taken various forms and dominated the discussion about lawlikeness in the last decades. Hence, ranking theory with its resources of formalizing defeasible reasoning or inductive schemes seems ideally suited to explicate the idea in a formal way. This is what the paper attempts to do. Thus it will turn out that a law is simply the deterministic analogue of a sequence of independent, identically distributed random variables. This entails that de Finetti's representation theorems can be directly transformed into an account of confirmation of laws thus conceived.

1. Introduction¹

Laws are true lawlike sentences. But what is lawlikeness? Much effort went into investigating the issue, but the richer the concert of opinions became, the more apparent their deficiencies became, too, and with it the profound importance of the issue for epistemology and philosophy of science.

The most widely agreed prime features are that laws, in contrast to accidental generalizations, support counterfactuals, have explanatory power, and are projectible from, or confirmed by, their instances. These characteristics have long been recognized. However, the three topics they refer to – counterfactuals, explanation, and induction – were little elaborated in the beginning and are strongly contested nowadays. Moreover, the interrelations between these subjects were quite obscure. Hence, these features did not point to a clear view of lawlikeness, either. In this paper, I try to advance the issue. Let me start with three straight decisions.

¹ This paper is similar in scope to Spohn (2002). That paper extends the analysis to *ceteris paribus* laws, whereas the present paper more strongly emphasizes the main point. Thus, my debts concerning the other paper extend also to this paper. Moreover, I am indebted to Eckehart Köhler for his very valuable comments and to Nils-Eric Sahlin for advice concerning Frank Ramsey's position.

The first decision takes a stance on the priority of the prime features. I am convinced that it is the inductive behavior associated with laws which is the most basic one, and that it somehow entails the other prime features. I cannot justify this stance in a few lines. Suffice it to say that my study of causation (1983) led me from Lewis' (1973) theory of counterfactuals over Gärdenfors' epistemic account of counterfactuals (cf., e.g., Gärdenfors 1981) ever deeper into the theory of induction where I finally thought I had reached firm ground. Clearly, this is not an unusual intellectual movement. In Spohn (1991) I explained my view on the relation of induction to causation and thus to explanation. However, I did not return to counterfactuals, not because I thought I had inadequate means for analyzing them, but because I felt that this subject is overlaid by many linguistic intricacies that are hard and perhaps not really important to account for. My decision finds recent support, for instance, in Lange (2000) who starts investigating the relation between laws and counterfactuals and also arrives at induction as the most basic issue.²

The second decision concerns the relation between laws and their prime features. When inquiring into lawlikeness the idea often was to search for something which *allows* us to use laws in induction, explanation and counterfactuals in the way we do. That is, given that induction is really the most basic aspect, lawlikeness should be something that *justifies* the role of laws in induction. This idea issued in perplexity; no good candidate could be found providing this justification.

There is an alternative idea, namely that lawlikeness is *nothing but* the role of laws in induction. In view of the history of inductive scepticism from Hume to Goodman – which made us despair of finding a deeper justification of induction and taught us rather to describe our inductive behavior and to inquire what is rational about it while being aware that this inquiry may produce only partial justification – this idea seems to be the wiser one. I do not mean to suggest that the lessons of inductive scepticism have been neglected.³ But it is important to be fully aware of these lessons, and hence I shall pursue here the second idea and fore-swear the search for deeper justifications. We shall see that we can still say quite a lot about rational induction.

² However, the theory of induction takes a probabilistic turn in Lange (2000, ch. 4), a move with which I disagree, as I shall explain below.

³ For instance, Lange (2000) endorses these lessons when explaining what he calls the root commitment concerning the inductive strategies associated with laws.

We are thus to study the inductive properties of laws. This presupposes some account of induction or confirmation within which to carry out the study. This is what my third decision is about. I think that on this matter philosophy of science went entirely wrong in the last 25 years. Bayesianism was always strong, and rightly so. In the 1950's and 60's much effort also went into the elaboration of a qualitative confirmation theory. However, this project was abandoned in the 70's. The main reason was certainly that the efforts were not successful at all. Niiniluoto (1972) gives an excellent survey that displays the incoherencies of the various attempts. An additional reason may have been the rise and success of the theory of counterfactuals, which answered many problems in philosophy of science (though not problems of induction) and thus attracted a lot of the motivation originally directed to an account of induction.

In any case, the effect was that Bayesianism was more or less the only remaining well-elaborated alternative. This hindered progress, because deterministic laws and probability do not fit together well. Deterministic laws are not simply the limiting case of probabilistic laws, just as deterministic causation is not the limiting case of probabilistic causation.⁴ We find a parallel in the disparity between belief, or acceptance-as-true, and subjective probability, which was highlighted by the lottery paradox and has, I think, as yet not found a convincing reconciliation. My conclusion is, though I have hardly argued for it, that Bayesianism is of little help in advancing the issue of lawlikeness.

Philosophical logic was very active since around 1975 in producing alternatives, though not under the labels 'induction' or 'confirmation'. However, these activities were hardly recognized in philosophy of science. Instead, they radiated to AI where they were rather successful. It is precisely in this area where we shall find help. Let me explain.

What should we expect an account of induction or inductive schemes to achieve? I take the view (cf. Spohn 2000a) that it is equivalent to a theory of belief revision or, more generally, to an account of the dynamics of doxastic states. This is why the topic is so inexhaustible. Everybody, from the neurophysiologist to the historian of ideas, can contribute to it, and one can deal with it from a descriptive as well as from a normative perspective.

⁴ For instance, Otte (1981) nicely documents how probabilistic theories of causation go astray with extreme 0-1-probabilities. Hence, the present attitude, as represented, e.g., by Pearl (2000), is to independently assume some account of deterministic causation, say, in terms of functional equations, which is then reflected in, but not defined by, 0-1-probabilities.

Philosophers, I assume, would like to come up with a very general normative account. Bayesianism provides such an account that is almost complete. There, rational doxastic states are described by probability measures, and their rational dynamics is described by various conditionalization rules. As mentioned, however, in order to connect up with deterministic laws, we should proceed with an account of doxastic states which represents plain belief or acceptance-as-true. Doxastic logic is sufficient for the statics, but it does not provide any dynamics. Probability < 1 cannot represent belief, because it does not license the inference from the beliefs in two conjuncts to the belief in their conjunction. Probability 1 cannot do it, either, because we would like to be able to update with respect to information previously disbelieved, because disbelieved propositions would have probability 0 according to this approach, and because Bayesian dynamics does not provide an account of conditionalization with respect to null propositions (that is why I called Bayesianism almost complete). Hence, Bayesianism is unhelpful. Belief revision theory (cf., e.g., Gärdenfors 1988) was devised to fill the gap. Unfortunately, the dynamics it provides turned out to be incomplete as well (cf. Spohn 1988, sect. 3). There have been several attempts to plug the holes (cf., e.g., Nayak 1994 and Halpern 2001), but I am still convinced that ranking theory, proposed in Spohn (1983, sect. 5.3, and 1988), though under a different name, offers the most adequate account for a full dynamics of plain belief.⁵

In any case, this is my third decision: to carry out my study of the inductive behavior of laws strictly in terms of the theory of ranking functions. This framework may be unfamiliar, but the study will not be difficult, since ranking theory is a very simple theory. Still, my constructive attempts will leave little space for broader discussion. The hope is that my account appears illuminating even without extensive comparative discussion.

The plan of the paper is now almost obvious. In section 2 I shall introduce the theory of ranking functions as far as needed. Section 3 explicates lawlikeness, i.e., the difference between laws and accidental generalizations insofar as it can be expressed in ranking terms. At least a brief comparative discussion is certainly called for; this is given in section 4. Since belief in a law is analyzed not as a belief in a regularity or some more sophisticated proposition, but rather as a certain

⁵ One may think that the incompleteness of Bayesianism is filled by Popper measures. However, if one combines the lessons of Spohn (1986) and (1988), it is clear that Popper measures are just as incomplete as is AGM belief revision theory. So, even the Bayesian has reason to buy into ranking theory as suggested in Spohn (1988), sect. 7.

inductive attitude, the immediate question arises how a law, i.e., such an inductive attitude, can be confirmed in turn. This crucial question is addressed in the final section 5.

I do not deny in this paper, but only neglect that laws have important meta-physical aspects as well. I have been less negligent in Spohn (1993), where I tried to understand causal laws as objectifications of inductive schemes, and in Spohn (1997), where I discussed both aspects of reduction sentences, the laws associated with disposition predicates. But if I am right, lawlikeness is an epistemological criterion, and this is why this paper entirely focuses on the epistemological aspects of laws.

2. Ranking Functions

Let us start with a set W of possible worlds, small rather than large worlds, as we shall see soon. Each subset of W is a truth condition or *proposition*. I assume propositions to be the objects of doxastic attitudes. Thus I take these attitudes to be intensional. We know well that this is problematic, and we scarcely know what to do about the problem. Hence, my assumption is just an act of front alignment.

The assumption also entails that we need not distinguish between propositions and sentences expressing them. Hence, I shall often use first-order sentences to represent or denote propositions and shall not distinguish between logically equivalent sentences, since they express the same proposition.

That is all we need to introduce our basic notion: κ is a *ranking function* (for W) iff κ is a function from W into \mathbf{N} (the set of non-negative integers) such that $\kappa(w) = 0$ for some $w \in W$. For each proposition $A \subseteq W$ the *rank* $\kappa(A)$ of A is defined by $\kappa(A) = \min \{ \kappa(w) \mid w \in A \}$ and $\kappa(\emptyset) = \infty$. For $A, B \subseteq W$ the (*conditional*) *rank* $\kappa(B \mid A)$ of B given A is defined by $\kappa(B \mid A) = \kappa(A \cap B) - \kappa(A)$. Since singletons of worlds are propositions as well, the point and the set function are interdefinable. The point function is simpler, but auxiliary, the set function is the one to be interpreted as a doxastic state.

Indeed, ranks are best interpreted as *grades of disbelief*. $\kappa(A) = 0$ says that A is not disbelieved at all. It does not say that A is believed; this is rather expressed by $\kappa(\bar{A}) > 0$,⁶ i.e., that non- A is disbelieved (to some degree).⁷ The clause that $\kappa(w) =$

⁶ \bar{A} is the complement or the negation of A .

0 for some $w \in W$ is thus a *consistency* requirement. It guarantees that at least some proposition, and in particular W itself, is not disbelieved. This entails the *law of negation*: for each $A \subseteq W$, either $\kappa(A) = 0$ or $\kappa(\bar{A}) = 0$ or both.

The set $C_\kappa = \{w \mid \kappa(w) = 0\}$ is called the *core* of κ (or of the doxastic state represented by κ). C_κ is the strongest proposition believed (to be true) in κ . Indeed, a proposition is believed in κ if and only if it is a superset of C_κ . Hence, the set of beliefs is *deductively closed* according to this representation.⁸

There are two laws for the distribution of grades of disbelief. The *law of conjunction*: $\kappa(A \cap B) = \kappa(A) + \kappa(B \mid A)$, i.e., the grade of disbelief in A and the grade of disbelief in B given A add up to the grade of disbelief in A -and- B . And the *law of disjunction*: $\kappa(A \cup B) = \min\{\kappa(A), \kappa(B)\}$, i.e., the grade of disbelief in a disjunction is the minimum of the grades of the disjuncts. The latter is again only a consistency requirement, though a conditional one; if that law would not hold the inconsistency could arise that both $\kappa(A \mid A \cup B), \kappa(B \mid A \cup B) > 0$, i.e., that both A and B are disbelieved given A -or- B .

According to the above definition, the law of disjunction indeed extends to disjunctions of arbitrary cardinality. I find this reasonable, since an inconsistency is to be avoided in any case, be it finitely or infinitely generated. Note that this entails that each countable set of ranks has a minimum and thus that the range of a ranking function is well-ordered. Hence, the range \mathbf{N} is a natural choice.⁹

However, here we better avoid all complexities involved in infinity. Therefore I shall outright assume that we are dealing only with finitely many worlds and hence only with finitely many propositions. This entails that each world in W (or the set of its distinctive features) is finite in turn. Hence, as announced, they are small worlds. One may think that this is a strange start for an investigation of natural laws. However, an analysis of lawlikeness should work also under this

⁷ I apologize for the double negation; after a while one gets used to it.

⁸ Consistency and deductive closure have been often attacked and equally often defended. The issue of logical omniscience is indeed highly problematic. But we have already decided the issue by taking propositions as objects of doxastic attitude. I don't see viable alternatives.

⁹ It is obvious that one has various options at this point. For instance, in Spohn (1988) I still took the range to consist of arbitrary ordinal numbers, but the advantages of this generality did not make up for the complications. By contrast, Hild (t.a., sect. 3.2) does not extend the law of disjunction to the infinite case and is thus free to adopt non-negative reals as values.

It is also obvious that the issue about infinite disjunction is closely related to the discussion of the Limit Assumption in Lewis (1973, sect. 1.4). Without that assumption, it may happen that "if A were the case, then B_i would be the case" is true for infinitely many B_i which are jointly unsatisfiable. Lewis finds reason to accept this state of affairs. I prefer to accept the Limit Assumption instead.

finiteness assumption. After all, our world seems both to have laws and to be finite. Generalizing my observations below to the infinite case is urgent, but would require a separate paper.

There is no need here to develop ranking theory extensively. A general remark may be more helpful: ranking theory works in almost perfect parallel to probability theory. Take any probabilistic theorem, replace probabilities by ranks, the sum of probabilities by the minimum of ranks, the product of probabilities by the sum of ranks, and the quotient of probabilities by the difference of ranks, and you are almost guaranteed to arrive at a ranking theorem. For instance, you thus get a ranking version of Bayes' theorem. Or you can develop the whole theory of Bayesian nets in ranking terms. And so on. The general reason is that one can roughly interpret ranks as the orders of magnitude of (infinitesimal) probabilities.

The parallel extends to the laws of doxastic change, i.e., to rules of conditionalization. Thus, it is at least plausible that ranking theory provides a complete dynamics of doxastic states (as may be shown in detail; cf. Spohn, 1988, sect. 5). It is due to this fact that ranking functions are fully entitled to be referred to by my loose talk of inductive schemes in the title of this paper.

It is still annoying, perhaps, that belief is not characterized in a positive way. But there is remedy: β is the *belief function* associated with κ (and thus a belief function) iff β is the function assigning integers to propositions such that $\beta(A) = \kappa(\bar{A}) - \kappa(A)$ for each $A \subseteq W$. Similarly, $\beta(B | A) = \kappa(\bar{B} | A) - \kappa(B | A)$. Recall that at least one of the terms $\kappa(\bar{A})$ and $\kappa(A)$ must be 0. Hence, $\beta(A) > 0$, < 0 , or $= 0$ iff, respectively, A is believed, disbelieved, or neither; and A is the more strongly believed, the larger $\beta(A)$. Thus, belief functions may appear to be more natural. But their formal behavior is more awkward. Therefore I shall use both notions.

Above, I claimed that a full dynamics of belief is tantamount to an account of induction and confirmation. So, what is confirmation with respect to ranking functions? The same as elsewhere, namely *positive relevance*: A confirms or is a reason for B relative to κ iff $\beta(B | A) > \beta(B | \bar{A})$, i.e., iff $\kappa(\bar{B} | A) > \kappa(\bar{B} | \bar{A})$ or $\kappa(B | A) < \kappa(B | \bar{A})$ or both.¹⁰

There is an issue here whether the condition should require $\beta(B | A) > \beta(B)$ or only $\beta(B | A) > \beta(B | \bar{A})$, as stated. In the corresponding probabilistic case, the two

¹⁰ I believe that if epistemologists talk of justification and warrant, they should basically refer to this relation of A being a reason for B ; cf. Spohn 2001. That's, however, a remark about a different context.

conditions are equivalent if all three terms are defined, but the first condition is a bit more general, since it may be defined while the second is not. That is why the first is often preferred. In the ranking case, however, all three terms are always defined, and the second condition may be satisfied while the first is not. In that case the second condition on which my definition is based seems to be more adequate.¹¹

A final point that will prove relevant later on: Ranking functions can be mixed, just as probability measures can. For instance, if κ_1 and κ_2 are two ranking functions for W and if κ^* is defined by

$$\kappa^*(A) = \min\{\kappa_1(A), \kappa_2(A) + n\} \text{ for some } n \in \mathbf{N} \text{ and all } A \subseteq W,$$

then κ^* is again a ranking function for W . Or more generally, if K is a set of ranking functions for W and ρ a ranking function for K , then κ^* defined by

$$\kappa^*(A) = \min\{\kappa(A) + \rho(\kappa) \mid \kappa \in K\} \text{ for all } A \subseteq W$$

is a ranking function for W . The function κ^* may be called the *mixture* of K by ρ .

This is all the material we shall need. I hope that the power and beauty of ranking theory is apparent already from this brief introduction. I have not argued here that if one wants to state a full dynamics of plain belief or acceptance-as-true, one must buy into ranking theory. I did so in Spohn (1988, sect. 3). Even that argument may not be entirely conclusive. However, I guess the space of choices is small, and I would be very surprised if a simpler choice than ranking theory were to be available.

Be this as it may, let us finally turn to our proper topic, the epistemology of laws.

¹¹ A relevant argument is provided by the so-called problem of old evidence. The problem is that after having accepted the evidence it can no longer be confirmatory. However, this is so only on the basis of the first condition. According to the second condition, learning about A can never change what is confirmed by A , and hence the problem does not arise. This point, or its probabilistic analogue, is made by Joyce (1999, sect. 6.4) with the help of Popper measures.

3. Laws

Let me start with a simple formal observation. Given some ranking function κ , to believe $A \wedge B$ means that $C_\kappa \subseteq A \cap B$, i.e., $\kappa(\neg A \vee \neg B) > 0$, i.e., $\min\{\kappa(\neg A), \kappa(\neg B)\} > 0$. This, however, can be implemented in many different ways. In particular, it leaves open how $\kappa(\neg A \vee \neg B)$ relates to $\kappa(\neg A)$ and $\kappa(\neg B)$ and whether and how A and B depend on each other. Hence, if you start with believing $A \wedge B$, but now learn that $\neg A$ obtains, you may, or may not, continue to believe B , depending on the value of $\kappa(\neg B \mid \neg A)$.

Basically the same point applies to believing a universal generalization. This, I propose, is the clue to understanding laws. Let us take $G = \bigwedge x(Px \rightarrow Qx)$ as our prototypical generalization (\rightarrow always denotes truth-functional material implication). I have already simplified things by assuming the worlds in W to be finite. This entails that the quantifier in G ranges over some finite domain D . For $a \in D$, let G_a be the instantiation of G by a , i.e., $G_a = Pa \rightarrow Qa$. Now to believe G in κ means that $C_\kappa \subseteq G$, i.e., $\kappa(\neg G) > 0$, i.e., $\min\{\kappa(\neg G_a) \mid a \in D\} > 0$. Thus, the generalization is believed as strongly as the weakest instantiation.¹²

Let us assume, moreover, that this is the only belief in κ , i.e., that $C_\kappa = G$; thus, no further beliefs interfere. This entails in particular that $\kappa(Pa \wedge Qa) = \kappa(\neg Pa \wedge Qa) = \kappa(\neg Pa \wedge \neg Qa) = 0 < \kappa(Pa \wedge \neg Qa)$ for each $a \in D$ and hence that $\kappa(\neg Qa \mid Pa) > 0$, i.e., that Pa is positively relevant for Qa . In other words, under this assumption the belief in the material implication $Pa \rightarrow Qa$ is equivalent to the positive relevance of Pa for Qa .

Again, the belief in G can be realized in many different ways. Let me focus for a while on two particular ways, which I call the ‘persistent’ and the ‘shaky’ attitude. If you learn about positive instances, G_a, G_b , etc. you do not change your beliefs according to κ , since you expected them to be positive, anyway.¹³ The crucial difference emerges when we look at how you respond to negative instances, $\neg G_a, \neg G_b$, etc. according to the various attitudes:

¹² Note, by the way, that this would also hold for an infinite domain of quantification. Hence, for ranking theory there is no problem of null confirmation for universal generalizations which beset Carnap’s inductive logic.

¹³ I am using here a technical notion of positive instance: a is a positive instance of G iff G_a , i.e. $Pa \rightarrow Qa$, is true. If $Pa \wedge Qa$, a positive instance in the intuitive sense, would be learnt, the beliefs would change, of course (at least given our assumptions that nothing except G is believed in κ).

If you have the *persistent* attitude,¹⁴ your belief in further instantiations is unaffected by negative instances, i.e., $\kappa(\neg G_b) = \kappa(\neg G_b \mid \neg G_a)$ ($b \neq a$), and indeed $\kappa(\neg G_b) = \kappa(\neg G_b \mid \neg G_{a_1} \wedge \dots \wedge \neg G_{a_n})$ for any $n \in \mathbf{N}$ ($b \neq a_1, \dots, a_n$). If, by contrast, you have the *shaky* attitude, your belief in further instantiations is destroyed by a negative instance, i.e., $\kappa(\neg G_b \mid \neg G_a) = 0$ and, a fortiori, $\kappa(\neg G_{\neq a} \mid \neg G_a) = 0$.¹⁵

The difference is, I find, characteristic of the distinction between lawlike and accidental generalizations. Let us look at two famous examples. First the coins:

- (1) All German coins are round.
- (2) All of the coins in my pocket today are made of silver.

It seems intuitively clear to me that we have the persistent attitude towards (1) and the shaky attitude towards (2). If we come across a cornered German coin, we wonder what might have happened to it, but our confidence that the next coin will be round again is not shattered. If, however, I find a copper coin in my pocket, my expectations concerning the further coins simply collapse; if (2) has proved wrong in one case, it may prove wrong in any case.

Or look at the metal cubes, which are often thought to be the toughest example:

- (3) All solid uranium cubes are smaller than one cubic mile.
- (4) All solid gold cubes are smaller than one cubic mile.

What I said about (1) and (2) applies here as well, I find. If we bump into a gold cube this large, we are surprised – and start thinking there might well be further ones. If we stumble upon an uranium cube of this size, we are surprised again. But we find our reasons for thinking that such a cube cannot exist unafflicted and will instead start investigating this extraordinary case (if it obtains for long enough).

The example of Bode’s law mentioned by Köhler (2003) fits perfectly. As Köhler explains, the law appeared to be entirely accidental for a long time. This finds expression in a shaky attitude. If astronomers had discovered a new planet in our solar system, or even another solar system with localizable planets, which

¹⁴ “Resilient” might be an appropriate term as well, but I do not want to speculate whether this would be a use of “resilient” similar or different to the one introduced by Skyrms; cf., e.g., Skyrms(1980).

¹⁵ Here, $G_{\neq a}$ stands for $\bigwedge x(x \neq a \rightarrow G_x)$. Note that we have $\kappa(\neg G \mid \neg G_a) = 0$ according to both the persistent and the shaky attitude, simply because $\neg G_a$ logically implies $\neg G$.

does not obey Bode's law, we would have abandoned it, i.e., we would have lost all confidence to apply it to new cases. But, as Köhler goes on to explain, it has nowadays acquired the status of a genuine law. If such an exception were discovered today, we would seek for possible explanations of this exception, but still apply the law to new cases, i.e., we would display a persistent attitude.¹⁶

As far as I see, the difference between the shaky and the persistent attitude applies as well to the other examples prominent in the literature.¹⁷ However, my wording is certainly more determined than my thinking. According to my survey, intuitions are often undecided. There may be mixed attitudes, and somehow the attitude seems to depend on how one came to believe in the regularity; there may be different settings for one and the same generalization. At the moment, though, I am concerned with carving out what appears to me to be the basic difference, and therefore I am painting black and white.

However, it is easy to refine the distinction between the persistent and the shaky attitude and to gain a systematic overview over the range of options provided by ranking theory. Let me briefly explain this point before further discussing my explication.

How many ways are there to believe the generalization G , i.e., to realize $\kappa(\neg G) > 0$? A natural and strongly simplifying assumption is

Symmetry: For all $a_1, \dots, a_n, b_1, \dots, b_n \in D$

$$\kappa(\neg G_{a_1} \wedge \dots \wedge \neg G_{a_n}) = \kappa(\neg G_{b_1} \wedge \dots \wedge \neg G_{b_n}).$$

In obvious analogy to inductive logic, symmetry says that the disbelief in violations of a generalization depends on their number, but not on the particular instances. For $n = 1$ symmetry entails that there is some $r > 0$ such that for all $a \in D$ $\kappa(\neg G_a) = \kappa(\neg G) = r$. More generally, symmetry entails, as is easy to see, that there is some function c from \mathbf{N} to \mathbf{N} such that for any $n+1$ different $a_1, \dots, a_n, b \in D$ the equality $\kappa(\neg G_b \mid \neg G_{a_1} \wedge \dots \wedge \neg G_{a_n}) = c(n)$ holds, where $c(0) = r$. Indeed, all ranks of all Boolean combinations of the G_a are uniquely determined by the function c .

Another plausible assumption familiar from inductive logic is

¹⁶ However, it is obviously a difficult matter to describe the kind of belief change that has taken place in the change of attitude towards Bode's law. It does not appear to be just a kind of conditionalization. The ideas developed in section 5 might perhaps help here.

¹⁷ Cf., e.g., the overview in Lange 2000, pp. 11f.

Non-negative instantial relevance: For all $a_1, \dots, a_n, a_{n+1}, b \in D$

$$\kappa(\neg G_b \mid \neg G_{a_1} \wedge \dots \wedge \neg G_{a_n}) \geq \kappa(\neg G_b \mid \neg G_{a_1} \wedge \dots \wedge \neg G_{a_n} \wedge \neg G_{a_{n+1}}).$$

This is tantamount to the function c being non-increasing.

Given the two assumptions there remain not so many ways to believe G ; any non-increasing function c with $c(0) = r$ stands for one such way. Hence, the persistent attitude characterized by $c(n) = r$ for all n stands for one extreme, whereas the shaky attitude for which $c(n) = 0$ for $n \geq 1$ stands for the other. So, one may think about whether any ways in between fit the examples better than the extreme ones.

Still, the consideration shows that the two attitudes I have taken to distinguish accidental and lawlike generalizations are suited best for marking the spectrum of possible attitudes. Hence, my discussion will proceed in terms of these two extremes.

4. Discussion

The examples have suggested that we treat a universal generalization G as lawlike if we have the persistent attitude towards it and that as accidental if we have the shaky attitude towards it. Hence, the difference does not lie in the propositional content, it lies only in our inductive attitude towards the generalization or, rather, its instantiations.

This account of lawlikeness is perhaps closest to the old idea that laws are not general statements, but rather inference rules or licenses. This idea goes back at least to Ramsey (1929)¹⁸ who states very clearly: “Many sentences express cognitive attitudes without being propositions; and the difference between saying yes or no to them is not the difference between saying yes or no to a proposition” (pp. 135f.). And “... laws are not either” [namely propositions] (p. 150). Rather: “The general belief consists in (a) A general enunciation, (b) A habit of singular belief” (p. 136). Clearly, one must not read too much of the present context into Ramsey’s writings; he was occupied with the problems of his days, for instance with

¹⁸ Sahlin (1997, p. 73) explains that Ramsey’s view of general propositions was obviously influenced by Hermann Weyl.

rejecting Russell's acknowledgment of general facts and with thus finding a different interpretation for general sentences.¹⁹ Still, his attempts to construct a pragmatic theory of belief, which are also reflected in my brief quotations, are surprisingly modern. There does not appear to be a 70 years gap when Lange says in his much more elaborate book (2000) that "the root commitment that we undertake when believing in a law involves the belief that a given inference rule possesses certain objective properties, such as reliability" (p. 189).²⁰

From a purely logical point of view, however, it was always difficult to see the difference between accepting $\bigwedge x(Px \rightarrow Qx)$ as an axiom and accepting the inference rule "for any a , infer Qa from Pa ". The only difference is that the rule is logically weaker; the rule is made admissible by the axiom, but the axiom cannot be inferred with the help of the rule. What else beside the unproductive logical point could be meant by the slogan "laws are inference rules" was always hard to explain. Clearly, there is also no firm difference concerning confirmation, falsification, etc.; if a general sentence can be falsified by a counter-instance, then a rule can be shown to be invalid by a case in which it leads from truth to falsity.

Still, one might say that the inference-license perspective puts more emphasis on what to do in the single case. And this emphasis is not mere rhetorics. It is reflected, I think, in my central notion of persistence and thus finds a precise induction-theoretic basis. It should be clear here that ranking functions *are* (possibly very complex) inference rules. If inference is not understood deductively, but more liberally as confirmation or being a reason, then ranking functions may obviously be interpreted as defeasible inference rules that are believed to be reliable, even if they are not universally valid.

Hence, the mark of laws is not their universality, which breaks down with one counter-instance, but rather their operation in each single case, which is not impaired by exceptions. Here, my account meets with Cartwright (1989) and her continuous forceful efforts to explain that what we have to attend to are the capacities and their cooperation taking effect in the single case. Her objective capacities or powers thus correspond to my subjective reasons as embodied in a ranking function. And as I tried to argue in (1993) and (2000b) the gap between the subjective and the objective and between reasons and causes is not unbridgeable.

¹⁹ See Sahlin (1991, pp. 138ff.) for more details.

²⁰ And he continues to explain the differences between the older 'inference licence' literature and his position which hide in the reference to "certain objective properties such as reliability".

So much for some important agreements. The most obvious disagreement is with Popper, of course. Given how much we have learned from Popper about philosophy of science, my account is really ironic, since it concludes in a way that it is the mark of laws that they are *not* falsifiable by negative instances; it is only the accidental generalizations that are so falsifiable. Of course, the idea that the belief in laws is not given up so easily is familiar at least since Kuhn's days, and even Popper (1934, ch. IV, §22) insisted from the outset that falsifications of laws proceed by more specialized counter-laws rather than simply by counter-instances. But I cannot recall having seen the point being stripped down to its induction-theoretic bones.

What I have said so far may provoke a confusion that I should hurry up to clarify. The persistent attitude towards $G = \bigwedge x(Px \rightarrow Qx)$ is characterized, I said, by the *independence* of the instantiations; experience of one instance does not affect belief about the others. In this way, belief about an instance G_b , i.e., the positive relevance of Pb for Qb , is persistent. But didn't we learn that one mark of lawlikeness is *enumerative induction*, i.e., the confirmation of the law by positive instances? Surely, enumerative induction outright contradicts the independence I claim.

Herein lies a subtle confusion. Belief in a law is more than belief in a proposition, it is a certain doxastic attitude, and that attitude as such is characterized by the independence in question. If I would have just this attitude, just this belief in a law, my κ would exhibit this independence. Enumerative induction, by contrast, is not about what the belief in a law *is*, but about how we may acquire or confirm this belief. The two inductive attitudes involved may be easily confused, but the confusion cannot be identified as long as one thinks belief in a law is just belief in a proposition.

However, what could it mean then to confirm a law if it does not mean to confirm a proposition? Indeed, my definition of confirmation in section 2 applies only to the latter, and to talk of the confirmation of laws, i.e., of a second-order inductive attitude towards a first-order inductive attitude, is at best metaphorical so far. Enumerative induction or falsificationism do not seem to make sense within this setting. This is a crucial point, and I shall devote the final section to it where I shall make a proposal for reconstructing enumerative induction and the falsification of laws. But so far I am concerned only with the attitude in which the belief in a law itself consists.

Is my explanation of lawlikeness a deep one? No, it is just as plain as, for instance, that of the counterfactual theorist who says that lawlikeness *is* support of counterfactuals or that a law *is* a universally quantified subjunctive conditional. Analysis has to start somewhere, and it acquires depth only by showing how to explain other features of laws by the basic ones. That is a task I cannot pursue here.²¹ But I would like to insist that, as a starting point, the present analysis is to be preferred. There are good reasons for feeling uneasy about starting with subjunctives or a similarity relation between worlds. By contrast, ranking theory is a very plain theory with a very obvious interpretation.

The only doubt one may have about my starting point may concern its sufficiency as a basis of analysis. In particular one may feel that the crucial property of laws is one which justifies the inductive attitude I have described, say, some kind of material or causal necessity. Maybe. But I am sceptical and refer to my second decision in section 1.

This does not mean that I have to sink into subjectivism, that I am bound to say that it is merely a matter of one's inductive taste what one takes to be a law. There may be objectivizations and rationalizations for our beliefs in laws. I do not intend to start speculating about this, but one very general rationalization is quite obvious. It is of vital importance to us to have persistent attitudes to a substantial extent. Something is almost always going wrong with our generalizations, and if we always had the shaky attitude, our inductions and expectations would break down dramatically and we could not go on living.

5. On the Confirmation of Laws

Above I have emphasized that one must not confuse the inductive or confirmatory attitude embodied in laws with the confirmation of the laws themselves. What the latter could be was, however, quite unclear, in contrast to the confirmation of propositions, which was already well handled by ranking functions. What can we do about this?

Fortunately, there is clear precedent in the literature. If one is aware of the close similarity between probability and ranking theory, one could notice that a law as I conceived it is nothing but a sequence of independent, identically

²¹ But see my account of causal explanation in terms of ranking functions in Spohn (1991).

distributed random variables translated into ranking terms. And if one is familiar with the work of de Finetti (1937), one sees that he addresses exactly our problem. In his famous theorems de Finetti showed that there is a one-one correspondence between symmetric probability measures for an infinite sequence of random variables and mixtures of Bernoulli measures according to which the variables are independent and identically distributed, and that the mixture concentrates more and more on a single Bernoulli measure as evidence accumulates. He thus showed to the objectivist that subjective symmetric measures provide everything he wants, i.e., beliefs about statistical hypotheses that converge toward the true one with increasing evidence.

De Finetti's issue between objectivism and subjectivism is not my concern. Ranking functions are thoroughly epistemological and have as such no objective interpretation.²² Still, we can immediately extract an account of the confirmation of laws from de Finetti's theory. Despite its artificial and formalistic appearance, the basic construction is, I find, highly illuminating.

Let us start with n mutually exclusive and jointly exhaustive properties or predicates Q_1, \dots, Q_n (these are Carnap's Q -predicates). For each $i \leq n$ we have the elementary law $G^i = \neg \forall x Q_i x = \bigwedge x \neg Q_i x$. For any proposition $A \subseteq W$ we may now count how often the law G^i is violated if A obtains; this is done by the function $v(A, i) = \text{card}\{a \in D \mid A \subseteq Q_i a\}$.²³ So, if we define the ranking function κ^i for W by $\kappa^i(A) = v(A, i)$, κ^i precisely represents the belief in the law G^i . Without any evidence, though, we do not believe in any law G^i . Our attitude towards the laws is rather represented by the ranking function ρ_0 for which $\rho_0(\kappa^i) = 0$ for each $i = 1, \dots, n$. Hence, our doxastic attitude towards the propositions $A \subseteq W$ is represented by the mixture κ_0 of the κ^i with respect to ρ_0 , as defined by

$$\kappa_0(A) = \min_{i \leq n} \kappa^i(A) + \rho_0(\kappa^i) = \min_{i \leq n} v(A, i).$$

Now, how does this attitude change by experience? Via conditionalization, as always. But let us describe this in detail. Let $\mathbf{r} = \langle r_1, \dots, r_n \rangle$ stand for any sequence of n non-negative integers, and let $r = r_1 + \dots + r_n$. Define next $E(\mathbf{r})$ to be the proposition (evidence) that among the first r objects precisely r_i instantiate Q_i ($i =$

²² But see Spohn (1993).

²³ I am still jumping between sentences and the corresponding propositions as seems convenient to me.

$1, \dots, n$); the order of instantiation is irrelevant. Clearly, $\kappa_0(E(\mathbf{r})) = \min r_i$. Let B range over propositions about the remaining objects and not the first r ones, and let $\kappa_{\mathbf{r}}$ be the ranking function that we have for those propositions after receiving evidence $E(\mathbf{r})$. Then we have:

$$\begin{aligned} \kappa_{\mathbf{r}}(B) &= \kappa_0(B \mid E(\mathbf{r})) = \kappa_0(B \cap E(\mathbf{r})) - \kappa_0(E(\mathbf{r})) \\ &= \min_{i \leq n} (v(B, i) + r_i) - \min_{i \leq n} r_i = \min_{i \leq n} (\kappa^i(B) + (r_i - \min_{i \leq n} r_i)). \end{aligned}$$

That is, if we define $\rho_{\mathbf{r}}$ by $\rho_{\mathbf{r}}(\kappa^i) = r_i - \min_{i \leq n} r_i$, we have

$$\kappa_{\mathbf{r}}(B) = \min_{i \leq n} (\kappa^i(B) + \rho_{\mathbf{r}}(\kappa^i)).$$

Hence, $\kappa_{\mathbf{r}}$ is the mixture of the κ^i with respect to $\rho_{\mathbf{r}}$. So, the evidence $E(\mathbf{r})$ makes us change our attitude towards the laws from ρ_0 to $\rho_{\mathbf{r}}$, and $\rho_{\mathbf{r}}$ represents the degrees to which the various laws have been confirmed or rather disconfirmed. If $\rho_{\mathbf{r}}(\kappa^i) > 0$, we might say that κ^i is falsified, but note that falsification is never conclusive in this construction.

The definition of $\rho_{\mathbf{r}}$ also shows what enumerative induction comes to in our context. The more instances of Q_j , i.e., counter-instances to G^j we find, the higher $\rho_{\mathbf{r}}(\kappa^j)$, i.e., the stronger the disbelief in κ^j . Or a bit more formally: Suppose we have already received evidence $E(\mathbf{r})$ and now discover a new instance a_{r+1} of Q_j . Then our total evidence is $E(\mathbf{s})$, where $\mathbf{s} = \langle r_1, \dots, r_{j-1}, r_j, r_{j+1}, \dots, r_n \rangle$. The relation between $\rho_{\mathbf{r}}$ and $\rho_{\mathbf{s}}$ is then immediately calculated: If $r_j \geq r_i$ for at least one $i \neq j$, i.e., if some other hypothesis is at least as good as G^j , then $\rho_{\mathbf{s}}(\kappa^j) = \rho_{\mathbf{r}}(\kappa^j) + 1$ and $\rho_{\mathbf{s}}(\kappa^i) = \rho_{\mathbf{r}}(\kappa^i)$ for $i \neq j$; i.e., the disbelief in κ^j increases. And if $r_j < r_i$ for all $i \neq j$, i.e., if G^j was so far *the* least violated hypothesis, then $\rho_{\mathbf{s}}(\kappa^j) = \rho_{\mathbf{r}}(\kappa^j) = 0$ and $\rho_{\mathbf{s}}(\kappa^i) = \rho_{\mathbf{r}}(\kappa^i) - 1$ for $i \neq j$; i.e., the other hypotheses become less disbelieved or possibly no longer disbelieved at all, and κ^j is thus possibly no longer believed. Either way, the negative instance $Q_j a_{r+1}$ is negatively relevant to κ^j and hence positively relevant to the complement $\{\kappa^i \mid i \neq j\}$. Moreover, the instance $Q_j a_{r+1}$ is non-negatively relevant to each individual κ^i ($i \neq j$). (It need not be positively relevant. The new instance of Q_j changes the degree of disbelief of κ^i or $\{\kappa^k \mid k \neq i\}$ only if κ^j is the best hypothesis or κ^i is the best and κ^j the second best

hypothesis according to $\rho_{\mathbf{r}}$.) Thus, our account of the confirmation of laws does justice to enumerative induction.

This is what I can offer as a translation of de Finetti's results into the framework of ranking functions. I find the translation basically plausible, and it strongly suggests following its course. One should characterize the class of ranking functions which represent mixtures of laws, and one should inquire the extent to which the representation is unique (for instance, there is an obvious one-one correspondence between the $\kappa_{\mathbf{r}}$ and the $\rho_{\mathbf{r}}$ in the above mixtures). One should look at de Finetti's representation results for the infinite as well as for the finite case (recall the finiteness assumption made in this paper). The ranking analogue to de Finetti's notion of partial exchangeability would be particularly interesting. And so forth.²⁴

On the other hand, the translation still looks artificial and quite detached from actual practice. For instance, if $\min r_i$ is large, one would tend to say that all of the laws G^i are disconfirmed by $E(\mathbf{r})$ and to conclude that none of the laws holds. One might account for this point by defining some κ^0 representing the belief in lawlessness, by mixing it into κ_0 , say with the weight $\rho_0(\kappa^0) = q$, and by finding then that as soon as $\min r_i > q$ we have $\rho_{\mathbf{r}}(\kappa^i) = 0$ only for $i = 0$. This is definitely not the only way or place to improve upon the basic result.

So, there is certainly a lot of work to do in order to extend the proposal and to apply it to more realistic cases. But the message should be clear already from the basic case just explained. The theory of mixtures provides a clear account of what it means to confirm and disconfirm not only propositions, but also inductive attitudes such as ranking functions representing belief in laws. Hence, at least in this respect, my account does not leave a principal gap. Still, it is obvious that this paper only begins to redeem the claim of its title: laws are persistent inductive schemes.

References

- Carnap, Rudolf / Richard C. Jeffrey (Eds.), *Studies in Inductive Logic and Probability, Vol. I*, Berkeley: University of California Press 1971.
 Cartwright, Nancy, *Nature's Capacities and Their Measurement*, Oxford: Clarendon Press 1989.

²⁴ See, e.g., the rich results collected in the papers in Carnap, Jeffrey (1971) and Jeffrey (1980).

- de Finetti, Bruno, "La Pr evision: Ses Lois Logiques, Ses Sources Subjectives", *Annales de l'Institut Henri Poincar e* 7, 1937. Engl. translation "Foresight: Its Logical Laws, Its Subjective Sources", in: Henry E. Kyburg jr. / Howard E. Smokler (Eds.), *Studies in Subjective Probability*, New York: John Wiley & Sons 1964, pp.93-158.
- G ardenfors, Peter, "An Epistemic Approach to Conditionals", *American Philosophical Quarterly* 18, 1981, pp. 203-211.
- G ardenfors, Peter, *Knowledge in Flux*, Cambridge, Mass.: MIT Press 1988.
- Halpern, Joseph Y., "Conditional Plausibility Measures and Bayesian Networks", *Journal of AI Research* 14, 2001, pp. 359-389.
- Hild, Matthias, *Introduction to Induction. On the First Principles of Reasoning*, to appear.
- Jeffrey, Richard C. (Ed.), *Studies in Inductive Logic and Probability, Vol. II*, Berkeley: University of California Press 1980.
- Joyce, James M., *The Foundations of Causal Decision Theory*, Cambridge: Cambridge University Press 1999.
- K ohler, Eckehart, "Comments on Wolfgang Spohn's Paper", this volume 2003, pp. #-##.
- Lange, Marc, *Natural Laws in Scientific Practice*, Oxford: Oxford University Press 2000.
- Lewis, David K., *Counterfactuals*, Oxford: Blackwell 1973.
- Nayak, Abhaya C., "Iterated Belief Change Based on Epistemic Entrenchment", *Erkenntnis* 41, 1994, pp. 353-390.
- Niiniluoto, Ilkka, "Inductive Systematization: Definition and a Critical Survey", *Synthese* 25, 1972, pp. 25-81.
- Otte, Richard, "A Critique of Suppes' Theory of Probabilistic Causality", *Synthese* 48, 1981, pp. 167-189.
- Pearl, Judea, *Causality. Models, Reasoning, and Inference*, Cambridge: Cambridge University Press 2000.
- Popper, Karl R., *Logik der Forschung*, Wien: Springer 1934.
- Ramsey, Frank P., "General Propositions and Causality", 1929, in: Frank P. Ramsey, *Foundations. Essays in Philosophy, Logic, Mathematics and Economics*, ed. by David H. Mellor, London: Routledge & Kegan Paul 1978, pp. 133-151.
- Sahlin, Nils-Eric, "Obtained by a Reliable Process and Always Leading to Success", *Theoria* 57, 1991, pp. 132-149.
- Sahlin, Nils-Eric, "'He Is No Good For My Work'. On the Philosophical Relations Between Ramsey and Wittgenstein", *Poznan Studies in the Philosophy of Science and the Humanities* 51, 1997, pp. 61-84.
- Skyrms, Brian, *Causal Necessity*, New Haven: Yale University Press 1980.
- Spohn, Wolfgang, *Eine Theorie der Kausalit at*, unpublished Habilitationsschrift, M unchen 1983.
- Spohn, Wolfgang, "The Representation of Popper Measures", *Topoi* 5, 1986, pp. 69-74.
- Spohn, Wolfgang, "Ordinal Conditional Functions. A Dynamic Theory of Epistemic States", in: William L. Harper / Brian Skyrms (Eds.), *Causation in Decision, Belief Change, and Statistics*, vol. II, Dordrecht: Kluwer 1988, pp. 105-134.
- Spohn, Wolfgang, "A Reason for Explanation: Explanations Provide Stable Reasons", in: Wolfgang Spohn / Bas C. van Fraassen / Brian Skyrms (Eds.), *Existence and Explanation*, Dordrecht: Kluwer 1991, pp. 165-196.
- Spohn, Wolfgang, "Causal Laws are Objectifications of Inductive Schemes", in: Jacques-Paul Dubucs (Ed.), *Philosophy of Probability*, Dordrecht: Kluwer 1993, pp. 223-252.
- Spohn, Wolfgang, "Begr undungen a priori – oder: ein frischer Blick auf Dispositionspr adikate", in Wolfgang Lenzen (Ed.), *Das weite Spektrum der Analytischen Philosophie. Festschrift f ur Franz von Kutschera*, Berlin: de Gruyter 1997, pp. 323-345.
- Spohn, Wolfgang, "Wo stehen wir heute mit dem Problem der Induktion?", in: Rainer Enskat (Ed.), *Erfahrung und Urteilskraft*, W urzburg: K onigshausen & Naumann 2000a, pp. 151-164.

- Spohn, Wolfgang, "Deterministic Causation", in: Wolfgang Spohn / Marion Ledwig / Michael Esfeld (Eds.), *Current Issues in Causation*, Paderborn: Mentis 2000b, pp. 21-46.
- Spohn, Wolfgang, "Vier Begründungsbegriffe", in: Thomas Grundmann (Ed.), *Erkenntnistheorie. Positionen zwischen Tradition und Gegenwart*, Paderborn: Mentis 2001, pp.33-52.
- Spohn, Wolfgang, "Laws, Ceteris Paribus Conditions, and the Dynamics of Belief", *Erkenntnis* 57, 2002, pp. ##-##.