Chapter 10
Two Coherence Principles[†1]

10.1 Introduction[1]

The purpose of this paper is twofold. On the one hand, it is a self-contained continuation of Spohn (1991) [here: ch. 9]. I studied there the relation between three principles of coherence and two versions of the principle of causality, thereby transferring the plausibility of the former onto the latter. Ever since then, I have wondered what more can be done to defend the coherence principles than simply appeal to their plausibility. This paper tries to give an answer which, however, is partial since I shall discuss only one of the old coherence principles.

On the other hand, a more important purpose interfered. Everyone engaged in the epistemological issue of foundationalism versus coherentism will grant that the notion of coherence is in bad shape. Since pondering the second of the present coherence principles, I thought that it offers a nice explication of the notion of coherence, which I have not found in the literature, which is perfectly precise and theoretically fruitful, and which therefore deserves to be presented. In view of the richness of the notion of coherence it would be silly to claim that this is *the* explication of the notion. The intent of this paper is rather to make this explication attractive by briefly relating it to other conceptions of coherence, by explaining the epistemological picture behind it, and by showing how one might argue for the associated principles.

The plan of the paper is this: Section 10.2 introduces some of the basics of epistemology, in particular the notion of a reason which is essential for the rest of the paper. Section 10.3 goes on to explain the two coherence principles which are the

[†1] This paper was originally published in: *Erkenntnis* 50 (1999) 155–175.

[1] I am grateful to the members of our Forschergruppe, Wolfgang Benkewitz, Ulf Friedrichsdorf, Ulrike Haas-Spohn, Volker Halbach, and Erik Olsson, for various suggestions and criticisms helping to improve the paper, to Jeffrey Knight for improving my English, and to an anonymous referee for further helpful suggestions.

subject of this paper and depicts their epistemological setting. Sections 10.4–10.7 finally offer four attempts to further ground these principles, the results of which are, briefly, that it is neither enumerative induction, nor the nature of propositions as objects of belief, nor consciousness, but rather an even more fundamental principle of rationality and an elementary theory of perception which entail these principles.

A final warning: In the course of the paper I shall make many claims which may be formally elaborated within the theory of ranking functions.[2] Here, however, I mostly dispense with formal details. This has obvious advantages. One of them is that I am not immediately committed to all the assumptions built in into the theory of ranking functions and can try instead to be more general. Thus I indicate, in an informal way only, which features of doxastic modelling are needed for the reasoning at hand. However, it may not always be clear to what extent I have avoided falling back on the features of ranking functions. Opacities of this kind belong to the drawbacks of informality which, I hope, do not outweigh the advantages.

10.2 Reasons

It seems uncontroversial to me that any kind of formal epistemology must represent a doxastic state by a function β with at least the following three features:

First, β must be defined on some set of propositions, where propositions, just by definition, are to be appropriate objects of belief. For the time being we may leave the exact nature of propositions an open question, which, of course, is much discussed; I shall only make the minimal assumption that they have Boolean structure.

Second, β must allow for degrees of belief, i.e., the range of β has to be some (usually linearly) ordered set of degrees. This condition is almost trivial in view of the fact that 1 (= belief), -1 (= disbelief), and 0 (= neutrality) also form such a set of degrees, indeed the minimal one.

Third, β must allow for conditionalization, i.e., it must assign conditional degrees of belief in some substantial, reasonable way. I am not sure how to strictly prove this, but any account of the dynamics of doxastic states I know of assumes conditional degrees of belief, and I have no idea what an alternative account could look like.

These three features immediately yield a most natural notion of confirmation, justification, or reason: A proposition A confirms, supports, or is *a reason for* a proposition B relative to a doxastic state β iff A strengthens the belief in B, i.e., if

[2] Introduced in Spohn (1988) [here: ch. 1] (where I still called them ordinal conditional functions). Ranking functions are particularly suited for more formal accounts of the present discussion, because they include a straightforward notion of belief – a point which has always been difficult for the probabilist.

the belief in B given A is stronger than given non-A.[3] We may thus define a bit more formally:

A is a reason for B given C (relative to β) iff $\beta(B \mid A \wedge C) > \beta(B \mid \neg A \wedge C)$
A is irrelevant to B given C iff $\beta(B \mid A \wedge C) = \beta(B \mid \neg A \wedge C)$ and
A is a counter-reason to B given C iff $\beta(B \mid A \wedge C) < \beta(B \mid \neg A \wedge C)$

The unconditional relations are defined by reference to the tautological condition; thus *A is a reason for B* (relative to β) iff $\beta(B \mid A) > \beta(B \mid \neg A)$. Hence, being a reason is nothing but positive relevance, and being a counter-reason is nothing but negative relevance – an old idea which reaches back at least to the discussion between Carnap and Popper about confirmation.

Which properties does the reason relation have? It follows trivially (assuming that $\neg\neg A$ is the same proposition as A) that

A is a reason for B given C iff $\neg A$ is a counter-reason to B given C.

All other properties of the reason relation depend on specific assumptions about β. The most common and useful choice is, of course, to conceive β as a probability measure. Then we obtain a reason relation which is symmetric and embraces logical consequence:

A is a reason for B given C iff B is a reason for A given C; and
if B is logically implied by A, then A is a reason for B (and vice versa), provided neither A nor B has an extreme probability.

We get many other important properties in addition, which, however, will not be relevant in the sequel. Moreover, it is worth mentioning that this probabilistic reason relation is *not* transitive.

Exactly the same properties result if we conceive β to be a ranking function. It would be interesting to find out about the properties of the reason relation if β is conceived as in the AGM-theory, as an entrenchment relation, for instance (cf. Gärdenfors 1988 or Rott 2001), as a Dempster-Shafer belief function (cf. Shafer 1976), etc. I believe, though, that the behavior of the reason relation turns out to be most satisfying relative to probability measures and ranking functions. There is no space to look closer into this issue; but I indeed think that this behavior is an unduly neglected adequacy criterion for formal representations of doxastic states.

This paper will be entirely based on the reason relation of positive relevance. It is obvious that this will bias the paper from the beginning. Are there not many other reason relations or similar notions around? So why use this one? This is a large question, but to attempt a lump-sum answer: It is my impression that those engaged in the epistemological issues I am going to address usually operate with a reason relation too vague to allow any rigorous theorizing and that alternative formal

[3] Why "given non-A" rather than "given nothing"? If we interpret β in the most familiar way as a probability measure, the two alternatives are equivalent as long as the relevant conditional probabilities are defined. However, if we interpret β, e.g., as a Popper measure or as a ranking function, a simple reflection shows my alternative to be preferable.

reason relations are less suited for these issues. A better answer, however, would first grant that no explication of the reason relation is to be expected to dominate all others and then provide an extended argument comparing the virtues of the theories built around the various explications – a task too large for a small paper. In a way, however, this paper may be seen as part of such an argument.[4] In any case, I shall simply proceed with positive relevance.

10.3 Two Coherence Principles

Since logical entailment abounds among propositions, the more embracive positive relevance does so as well. Therefore it will be most crucial to observe how much of positive relevance there is beyond logical entailment. To this end we must give a bit more structure to the propositions. I shall assume that we can discern atomic propositions and that these atomic propositions are logically independent. Or to be a bit more specific: I assume a Boolean algebra of propositions as it is usually constructed in probability theory or, e.g., in Carnap's latest inductive logic (1971/80). This construction starts from a set of *variables* (not in the logical sense, but in the sense of stochastic variables). Each variable can take values from a certain range; in the simplest case it is a yes/no variable ranging over $\{0,1\}$. A *possible world* or a *possible course of events* is a function specifying a value for each variable; this is the value the variable takes in this world or course of events. A *proposition* is any set of possible courses of events. Let U denote the set of all variables, and for $V \subseteq U$ let $P(V)$ denote the set of all propositions over V; thus, $A \in P(V)$ iff A does not discriminate outside V, i.e. iff for any world w in A all worlds differing from w only outside of V are also in A. Then, a proposition A is *atomic* iff it is about a single variable, i.e., if there is a variable X such that $A \in P(\{X\})$; thus, atomic propositions concerning different variables are logically independent. Finally, a proposition is *a posteriori* iff it is neither empty (a priori false) nor identical with the set of all worlds (a priori true).

How should positive relevance spread over the set of propositions? It is impossible to say. If U is some gerrymander, a subject's beliefs concerning U may take any form whatsoever. However, if U is the set of *all* variables within the grasp of a subject's doxastic state β (certainly an ill-defined set), we have more definite expectations. One plausible expectation is stated in the *special coherence principle*:

> For any variable X and any a posteriori proposition $A \in P(X)$ there is a proposition $B \in P(U - \{X\})$ such that B is a reason for A (relative to β).

Thus the special coherence principle says that there is some inductive support for each atomic a posteriori proposition or, more simply, that no variable is independent from all others.

---

[4] Other papers of mine (Spohn 1991, 1997c, 1997/98) [here: chs. 9, 12, and 11] may be seen as further parts of such an argument. Spohn (1997/98, sect. 2) [here: sect. 11.2] in particular, contains some remarks comparing positive relevance with other reason relations.

I refer to Spohn (1991) [here: ch. 9] for one way of expanding and strengthening the special coherence principle.[5] Here I shall take another way leading to an explication of coherence. The idea is simply that the special principle looks just as plausible when we replace the single variable X by some arbitrary set of variables. Then we get the much stronger *general coherence principle*:

> For any proper subset $V \subset U$ and any a posteriori proposition $A \in P(V)$ there is a proposition $B \in P(U - V)$ such that B is a reason for A (relative to β).

Thus the general coherence principle says that the set of all variables does not fall into independent parts. Or in graph-theoretic terms: If one represents the (conditional) dependencies and independencies given by the doxastic state β in a so-called Bayesian network,[6] the general principle requires that this network is a connected graph which cannot be separated into unconnected parts. Or to be a bit more pompous: The general principle really affirms something like the unity of science, the unity of our empirical world picture.

So far, I have only claimed that these principles are plausible; in the subsequent sections we shall have to inquire into what the deeper truth behind them might be. However, let me first ask what these principles have to do with coherence. The answer is simple; the general principle *defines* coherence:

> A doxastic state β is *coherent* iff β satisfies the general coherence principle.

Coherence *is* connectedness, integratedness. This explication is as precise and clear as the underlying reason relation; it thus compares favorably with most alternative offers.

However, what we really like to know is, of course, how the explication and the principles relate to coherence as it figures in the debate between coherentism and its alternatives, or, for short, in the "knowledge debate" (since the alternatives have arisen in the quest for the nature of knowledge). So let me introduce four rough characters: the foundationalist, the coherentist, the externalist, and the (formal) belief theorist, for want of a better term. The former three are the well-known archetypes in the knowledge debate. The primary epistemological interests of the last, however, do not lie in this debate. They are, rather, to build formal models of the statics and the dynamics of doxastic states, to develop their theory, and to somehow justify the assumptions built in into the models as rational. Which stance, if any, in the knowledge

---

[5] The explication of causation defended in Spohn (1991) [here: ch. 9] entails that the special coherence principle is equivalent to a very weak principle of causality which says that each atomic proposition has a cause or an effect in some possible world. Moreover, I present there two strengthenings of the special coherence principle, one entailing and the other being entailed by a weak principle of causality saying that each atomic fact has a cause or an effect in the *actual* world.

[6] This is a directed acyclic graph the nodes of which represent variables and the vertices of which represent conditional independencies between variables obtaining according to β insofar as *all* these independencies can be read off from the vertices by help of the so-called criterion of d-separability; cf. Pearl (1988, sect. 3.3). The theory of Bayesian nets is an utterly useful tool for the epistemologist, not only because of its graphical qualities; however, it is applicable only where conditional independence behaves as in probability measures or in ranking functions.

debate is thereby entailed is only a secondary question. The attitude Carnap finally took towards inductive logic (cf. his 1971/80) is certainly prototypical, belief revision theory and probabilistic epistemology are carried out in the same spirit, and I consider myself to be a formal belief theorist in this sense as well.

There are various agreements and disagreements among these characters. All of them have some notion of the reason relation. However, the foundationalist, the coherentist, and the externalist diverge on the properties of the reason relation in well-known ways. The belief theorist is certainly an internalist; I do not know of any belief theorist providing theoretical means for allowing external facts to be reasons for or to justify beliefs. Whether he sides with the foundationalist or the coherentist will, however, depend on his doxastic model.[7] For instance, if he takes the reason relation to be symmetric, as I did above, he thereby opposes the foundationalist who insists that basic beliefs are reasons for other beliefs, but cannot have reasons outside themselves.

There is a much deeper disagreement, though. Those engaged in the knowledge debate assume that there is not only the binary relation of one belief being a reason for another, but also a unary predicate (or quantity) of a belief being justified or warranted (to a certain degree). To put it graphically, the common picture[8] is this: The binary reason relation provides a network of channels between its relata, the thickness of which governs how much of the viscous quantity called *degree of warrant* can flow through them. By itself, however, the network is empty. It still needs to be filled with this quantity. Now the disagreement starts. The foundationalist thinks that this quantity is created in what he calls basic beliefs and then flows to the other beliefs. The externalist seeks the source of this quantity in appropriately related external facts. The coherentist either says that this quantity is bestowed on a belief in virtue of its relational coherence with all other beliefs,[9] or that this quantity is created by the network itself according to its degree of intrinsic coherence and then distributes differentially among its nodes.[10] It is clear that many mixtures are conceivable, and have indeed been suggested.

Now, the deep schism is that the belief theorist does not at all know what to make of this picture. It is hardly explicable for him and, what is worse, he has no use for it. Not that his theory of doxastic states would be complete; but a theory of warrant is not among the things he is missing. There is overwhelming evidence that the theory of belief contents requires much more sophistication. He may strive for

---

[7] And on his explication of the reason relation – he need not adopt my above proposal.

[8] It may be explicitly found in BonJour (1985, sect. 5.2) or in Plantinga (1993, ch. 4). In fact, it is built in into the set-up of the justification trilemma which drives the knowledge debate and according to which one can choose only between three unpalatable alternatives: infinite justificatory regress, circular justification, or stopping justification at some unjustified or obscurely self-justifying point.

[9] This is, roughly, the version of Lehrer (1990, pp. 147ff.).

[10] This is pure coherentism as explained by BonJour (1985, sect. 5.2) and amended later on. Plantinga (1993, p. 78) criticizes this version as pure magic; indeed it looks like *creatio ex nihilo*.

more realism by considering other kinds of degrees of belief, probability intervals for instance, instead of point probabilities, or by adding a badly needed theory of computational management of doxastic states. The theory about a priori states is severely underdeveloped in my view. The theory of doxastic changes does not say much about non-experiential changes, for instance conceptual change. The input theory of observation and experience could certainly be more detailed; and the output theory of action and behavior need not stick to decision theory. Such are the tasks for the belief theorist to complete his theory (all of which are belabored, of course). As far as I know, however, the knowledge debate has not advanced any good reason for the belief theorist to think that he needs to add a theory of warrant as well. In a way, this is not surprising since knowledge is simply not a relevant topic for the belief theorist and since the notion of justification or warrant plays its primary role precisely in the difference between true belief and knowledge.[11]

In ch. 6 on Bayesian coherentism Plantinga (1993) arrives at the same conclusion, suggesting that it is a defect of the Bayesian, or the belief theorist in general, that he is unhelpful to the knowledge debate. This is only half of the truth, however. The concern should really be mutual. Of course the belief theorist should be deeply worried about the fact that he cannot, and does not want to, say much about the notion of warrant which seems to arise so naturally and is taken so seriously by many serious philosophers. Conversely, however, the knowledge debate should be deeply worried about the fact that the notion of warrant is apparently unimportant to a large part of epistemology and to equally many equally serious philosophers. The schism is unbridged.[12]

I am explaining all this because it clearly entails that whenever a belief theorist like me starts using the terms so central to the knowledge debate, he is bound to stand crossways to that debate. The conclusion I draw from this situation is this: If the belief theorist has complete ways of theorizing, or ways to complete theorizing, without referring to the knowledge debate, this is so either because that debate is really immaterial or because it is somehow implicit in his theorizing. Since I cannot believe the former, I try to verify the latter. This is how my efforts here should be seen.

For instance, defining the reason relation as I did above is something the pure belief theorist need not do; it is merely an attempt to approach the knowledge debate. Likewise, I might progress from the binary relation to the unary predicate by saying that a belief is justified iff the balance of reasons is in its favor. However, this is no more than an insubstantial metaphor so far. The belief theorist does not have the idea of an active weighing of reasons which results in a justified belief. Rather, in his rationalized picture, a doxastic state *eo ipso* satisfies the basic laws of his doxastic model (e.g., the axioms of probability), and hence each proposition is automatically in balance, so to speak, within a doxastic state: it could not be

[11] One should note that doubts about the role of justification have also been articulated within the knowledge debate; cf. von Kutschera (1982, ch. 1) or Sartwell (1992).

[12] This schism seemed to me, on reflection, to be at the center of the conference whose results are published here. Perhaps the conference has at least spanned a rope between the sides.

believed to any other degree without violating these laws (without violating, e.g., coherence in the probabilistic sense).[13]

Finally, the belief theorist has certainly great difficulties in understanding the notion of coherence, as it figures in the knowledge debate, in a warrant-creating or warrant-conferring role (cf., however, Olsson 1999). Moreover, he certainly cannot make sense of measuring coherence by measuring (probabilistic) inconsistencies (unless he resorts to something like paraconsistent logic). However, he has no difficulties in understanding the standard examples of consistent but incoherent doxastic states which simply consist in a set of unconnected or independent beliefs.[14] Connectedness and dependence is precisely what the reason relation creates. Hence, this aspect of the notion of coherence is most adequately captured by the general coherence principle. And as such it should also be of interest within the knowledge debate.[15]

Having thus roughly clarified the setting within which the above explication of coherence is placed, I can finally turn to the main purpose of this paper, i.e., to considering on which grounds doxastic states should satisfy these coherence principles.

10.4 Justifying the Coherence Principles via Enumerative Induction?

Let me first briefly look into the relation between the coherence principles and inductive logic. Indeed, this is the only place, as far as I know, where similar relevance principles are stated.

The most important and most convincing one is the principle of positive instantial relevance (cf. Carnap 1971, sect. 13), which is the probabilistic analogue to enumerative induction and says, roughly, that the fact that one individual has a certain attribute makes it likelier that another individual has this attribute as well. This clearly entails the special coherence principle, provided that the set U of variables has an appropriate structure.[16] However, positive instantial relevance is silent on the general principle, because it does not say anything about the relation between different attributes.

---

[13] The metaphor would be more substantial if it would be possible to reconstruct the degree of belief in a proposition from the strengths of the reason relations in which it stands. However, it is easy to see that this is not possible for my above reason relation and doxastic states conceived as probability measures or ranking functions. It might be worthwhile investigating which stronger assumptions allow such reconstruction.

[14] See also the coherence conditions (3) and (4) in BonJour (1985, p. 98).

[15] In Spohn (1991, sect. 5) [here: sect. 9.5] I try to argue that this kind of coherence is closely related to explanatory coherence.

[16] The structure is appropriate if the variables are constructed from attributes, relations, or magnitudes and objects such that each attribute etc. figures in more than one variable. This condition is certainly satisfied if U is the set of variables within the grasp of a given subject, and indeed implied by what Evans (1982, sect. 4.3) calls the generality constraint.

Such relations are rather specified in Carnap's theory of the analogy influence (cf. Carnap 1980, sect. 16f.). However, it is not at all clear whether Carnap's full inductive logic would satisfy the general coherence principle. This would depend on whether all attributes are integrated in one attribute space and, if not, whether any relations between different attribute spaces are specified, and how precisely the analogy influence spreads within one attribute space. Moreover, it must be admitted that this theory of analogy has been put forward quite tentatively and that it has not met many friends in the last decades; without further scrutiny no strong case can be built on it. It therefore seems advisable to look for other ways of justifying the coherence principles.

10.5 Justifying the Coherence Principles via the Essence of Propositions?

The next possible answer, though much deeper, will also be considered only very briefly. First, equating propositions with sentence meanings seems quite innocent. What precisely meanings are is, however, an inexhaustible topic. One view, which is still popular in the wake of the verifiability theory of meaning, is to construe sentence meanings or propositions not as truth conditions, but rather as assertibility, justifiability, or acceptability conditions of sentences. There are many places in the philosophy of this century where such a view is suggested. Properly understood, this approach takes the reason relations which a proposition bears to other propositions as *individuating* this proposition,[17] though this is rarely endorsed in an explicit way.

This definition of propositions entails the special coherence principle: there can be at most one exception, i.e. at most one proposition which stands in no reason relations whatsoever. Despite my sympathies for such ideas, I think that this justification of the coherence principle is at least doubtful. My concerns are fourfold.

First, I do not know of any satisfying formal implementation of the idea. The proponents of acceptability conditions are usually stuck in metaphorical descriptions, and as far as I know, the formal literature does not address the question. If the individuation of propositions is aided by the logical relations between them, it becomes trivial because each proposition is uniquely characterized by the set of its logical consequences. However, if the undertaking is restricted to the reason relations as explained above, I do not know how it might be accomplished, how, for instance, the Boolean structure of propositions might be generated. As long as this technical task is not achieved,[18] this justification of the coherence principles does not work.

---

[17] A nice parallel would be Davidson (1969) who individuated events via the causal relations they bear to other events.

[18] I know of two attempts which get close to what would be needed, namely the ingenious proposal of Popper (1934/69, Neuer Anhang *IV) to extract the Boolean structure of propositions from the properties of conditional probabilities and the construction of Gärdenfors (1988, ch. 6) which achieves the same by starting from the properties of the dynamics of belief.

Second, this definition of propositions can avoid outright circularity only by claiming a thoroughly holistic conception of sentence meanings or propositions. To maintain such a holism is certainly difficult in view of the large and on-going philosophical debate about it.[19]

A third and related concern is that there are competing accounts of propositions which do not seem worse: for instance, the account which defines, as I did in Section 10.2, propositions as sets of possible worlds or more complex indices, or the account which takes propositions as internally structured, i.e., as somehow composed of properties, relations, and objects by various rules of composition. Thus, before this line of reasoning in favor of the coherence principles can succeed, one would have to engage in intricate arguments showing that the individuation of propositions via justifiability or acceptability conditions is to be preferred to the other ones within the given context. Here one certainly moves on very general and problematic grounds.

Finally, we have the same problem as with Carnap's inductive logic. So far, the proposed strategy does not yield the general coherence principle and I cannot see any feasible strengthening of the strategy which would do so. Hence, success is again incomplete. All this is sufficient reason for looking further.

10.6 Justifying the Coherence Principles via Consciousness?

If the general principle is so recalcitrant, we better face it directly. The general line of reasoning for it seems quite obvious. Suppose my doxastic state violates the general principle and the set of variables within my grasp divides into two independent separate parts. Where am I? Certainly, my self-picture is an indispensible part of my doxastic state,[20] there are a lot of variables about myself. Apparently, these variables cannot belong to both parts, the dividing line cannot cut through myself. Thus they are wholly within one part. But then it is hard to see how the other independent part could be within my reach. My learning seems to be restricted to the part containing me, and I could not come to believe anything about the other part.

This line of reasoning may look promising, but it is a different matter to turn it into a sound argument. Clearly, the suggestion has a Kantian ring. When I just said that at least the propositions about myself must be connected, I should probably have been so cautious to refer only to the propositions concerning my consciousness. And then we seem to be in the vicinity of Kant's profound idea that the "I think" must potentially accompany all my thoughts and ideas, i.e., in the vicinity of the transcendental unity of pure apperception which Kant declares to be the first principle of understanding lying at the base of all our judgments. So, in a nutshell,

[19] Fodor (1987, ch. 3), e.g., offers a most forceful criticism of such holism.

[20] See, e.g., Perry (1979) concerning the irreducibility of attitudes de se.

the suggestion is that we may somehow derive the connectedness of our empirical beliefs from the unity of consciousness. However, closer inspection fails to confirm this; we rather encounter a class of propositions which must be exempt from the coherence principles: facts of consciousness are not within the field of the reason relation. This is the consequence of the following considerations.[21]

The suggestion from Kant is that the relevant sort of facts of consciousness are propositions about one's own beliefs; in a sense, I simply know what I do, and do not, believe. However, it would be intuitively very strange to defend, justify, or reason for one's beliefs with the help of such knowledge. Suppose someone claims: "Clinton will resign before the end of the year," and when asked for his reasons he responds: "I believe so." Then he has certainly given no reason at all, even if the answer is, unnaturally, interpreted not as the affirmation of the original claim, but as an expression of a second-order belief. Believing to believe that A is somehow tantamount to believing that A, and therefore the former cannot be used in reasoning for the latter.[22]

This intuition should be substantiated, though. This is done by Benkewitz (forthcoming, sect. 5.3), in an extended argument. Instead of adapting this argument to the present purposes,[23] however, I shall try to confirm a fairly common thought which runs as follows: Facts of consciousness are maximally certain and, generally, maximally certain propositions cannot have, or be, reasons.

Let me start with the latter claim. Why can maximally certain propositions not have, or be, reasons? Observe first that, if A is maximally certain, it is so under any conditions; this is so at least if doxastic states satisfy an analogue to the formula of total probability, i.e., if the degree of belief in a certain proposition is in some sense a weighted mixture of the conditional degrees of beliefs of that proposition under mutually disjoint and jointly exhaustive conditions. This observation entails that no proposition can be a reason, in my sense, for a maximally certain proposition. If one further accepts the symmetry of the reason relation, then this in turn entails that maximal certainties cannot be reasons for other propositions either. But one may also argue that a maximal certainty cannot be a reason for other propositions because relative to the negation of a maximal certainty, to which the minimal degree of belief should be assigned, no conditional degrees of belief can be defined. For, if such

[21] This is not intended to disprove Kant, of course, since I shall not be concerned with the special role of "I" which is so important for Kant. However, my implication certainly is that whatever kind of unity is generated by the special role of "I", it is not the unity in terms of the reason relation.

[22] Because of this I wondered about the account of observation in BonJour (1985, ch. 6) for which this kind of reasoning is essential (and hence I tried in Spohn 1997/98 [here: ch. 11] to give a coherentist account of observation without alluding to second-order beliefs). That second-order beliefs find no place in the reason relation is reflected in BonJour's work also in the role which is played by his Doxastic Presumption, which is special since he admits that there is no further justification for the beliefs about one's own beliefs.

[23] Benkewitz argues for the more consequential thesis that in an important sense a subject cannot causally explain its own present beliefs, and it would require some explanation to show how the present thesis is implicitly contained in that argument. I am grateful to Wolfgang Benkewitz for alerting me to assertions of this kind.

conditional degrees of belief were non-trivially explained, i.e. in such a way that they may have different values, this would entail an impossible splitting-up of the minimal degree of belief into several different degrees. This reasoning establishes a large class of exceptions to the coherence principles, namely the set of maximally certain propositions all of which cannot engage into reason relations.

The next question is: Which propositions are maximally certain? There seem to be two kinds. The first kind consists of propositions which are a priori in the sense of being necessarily believed in any doxastic state capable of grasping them. All analytic propositions, like "bachelors are unmarried" or "5 + 7 = 12", are a priori. But there also are Kripkean non-analytic propositions a priori like "I exist," "I am here now," "the *F* is an *F*"(provided that "the *F*" is read referentially), and reduction sentences for dispositions (cf. Spohn 1997c, [here: ch. 12]) – a class of propositions which strongly recommends itself for further investigation. Still, the fact that such a priori propositions do not fall under the scope of the coherence principles is no cause for worry. The coherence principles are designed for empirical beliefs a posteriori, and thus this kind of exception is easily tolerable.

Besides, however, it is usually held that there is a second class of propositions which are maximally certain, namely, facts of consciousness. These comprise facts about my perceptual or experiential state such as "I am now appeared redly" (to use Chisholm's phrase) or "I am in pain now", facts about my present propositional attitudes like "I think that *A*", "I believe that *A*", "I desire *A* to be the case", or "I intend to do *A*", and maybe other kinds of facts. If these propositions are maximally certain, the strategy presently considered apparently fails.

Why, though, should we think of facts of consciousness as maximally certain? We might try to elaborate one of the following two argument sketches. Both proceed from the following starting point: What precisely are facts of consciousness? We have listed examples, but a general explication would be better. The following applies to the examples and seems generally adequate: *A* is a *fact of consciousness* iff *A* is true and necessarily equivalent with, i.e., the same proposition as the proposition that I (presently) believe that *A*.[24,†2] Moreover, it seems that in this case such necessity is a priori and hence that the two propositions are even analytically equivalent. I am well aware that in giving this explication I am opening a Pandora's box; but for the present purpose let us neglect this and just look what follows from it.

The first argument sketch is this: To believe something presumably means to believe it at least to a certain degree (analogously, to be tall for a man means something like to be taller than, say, 6'4").[25] Hence, if *A* is a fact of consciousness, it is the same as believing *A* at least to a certain degree. Believing *A* in a specific, sufficiently large degree would then be something stronger, and something different, for each different degree. But if *A* is the same as believing *A* there seems to be no room for such varying degrees of belief in *A*. This suggests that there is no proper degree

[24] Thus, facts of consciousness are the same as what Benkewitz (1999, sect. 5) calls internal contents (as opposed to external contents of beliefs).

[†2] In Spohn (2005d) I have more fully developed this analysis.

[25] This idea and its vagueness is propounded by Hunter (1996).

of belief for *A*, only an improper one, so to speak; and the only improper degree of belief (which is sufficiently large) is the one expressing maximal certainty.

The other argument sketch is this: We have already seen above that doxastic states cannot be conditionalized with respect to negations of maximally certain propositions. Likewise, it looks strange and even seems impossible – though I have no further argument for this – to conditionalize a doxastic state with respect to something which denies that very state. According to my explication of facts of consciousness, however, which declares such a fact to be part of a doxastic state, we would try to do exactly this if we try to conditionalize a doxastic state with respect to the negation of a fact of consciousness. Hence, if such conditionalization does not make sense, the above explication of the reason relation does not apply to facts of consciousness; that is, facts of consciousness cannot be reasons for other propositions.

So whether we are content with declaring that facts of consciousness are maximally certain or add one of the further arguments, the conclusion is in any case that such facts are not in the field of the reason relation and that this attempt, at least, was not the right way to get help from Kantian insights. Still, one may wonder about this conclusion. It seemed to be generally agreed that the foundationalist is right insofar as the basic beliefs he postulates have at least some justifying force; the question was rather whether all justification ultimately reduces to them and whether they are really foundational in the sense of having no justification outside themselves. Moreover, conscious phenomenal or experiential states (or the identical beliefs in them) appeared to be first-rate candidates for such basic beliefs. This appearance is false, however, if my conclusion is right. But how could it then be so plausible? Let me close this section with offering two brief thoughts for reconciliation.

First, phenomenal facts of consciousness are really quite special and can only be expressed by phrases like "it looks now *thus* to me", accompanied by a deferred ostension to my present phenomenal experience. Propositions like the one that the tomato in front of me looks red to me, or even that I am appeared redly now, may also seem to be facts of consciousness. But they are not, they are subtly different; and the subtle difference suffices to make them unexceptional and to integrate them into the circle of reason. So, they may well serve as a substitute offer to the foundationalist.[26]

Second, one must pay close attention to the dynamics of the reason relation. Doxastic states change and positive relevance changes with them. Consider the proposition that I shall be in such and such a conscious phenomenal or doxastic state in an hour. There is no problem for this proposition to be a reason for, and to find reason in, other propositions. An hour later, I am in such and such a conscious state and thus believe it to obtain with maximal certainty or in a way excluding it from the reason relation.[27] Still an hour later, my doxastic state will have changed again. Then I believe that I was in such and such a conscious state an hour ago – in a less than maximal

[26] This is more fully argued in Spohn (1997/98) [here: ch. 11]. However, I argue at the same time that these propositions are not basic in the foundationalist's strict sense.

[27] Strictly speaking, it is not the same proposition as before which I believe then, because the temporal index has shifted. However, being precise about this would only enforce my point.

degree, however, not because I have learnt new things in between, but simply because the conscious state has turned into a less than maximally certain recollection which is again justificatorily related to other propositions in both ways. Hence, even the phenomenal proposition is within the circle of reason for most of the time; it jumps out of the circle only during the dazzling moment of conscious experience.

10.7 Justifying the Coherence Principles via a Theory of Perception

Should we conclude therefore that the line of reasoning sketched at the beginning of the previous section fails? No. I suggest that we stay away from that dazzling moment and replace the subject's consciousness by his beliefs about an arbitrary perceiver who may be a third person or he himself at another time. Thereby we can turn the rough sketch into a more rigorous argument proceeding in seven steps. The first six steps deal with the special coherence principle. A simple further step will finally carry us to the general principle.

(1) The argument must start somewhere. I propose the following principle of rationality: A subject should have a variable degree of belief in any a posteriori proposition within his grasp. That is, if the subject believes in such a proposition to a certain degree, there should be a possible dynamics which leads him to believe that proposition to some other degree.

This sounds almost tautological. Recall that a priori propositions were defined above as propositions necessarily believed in any doxastic state (capable of grasping them). Hence, a posteriori propositions may or may not be believed or, more generally, may have varying degrees of belief in different doxastic states. This is weaker than the rationality principle; the different states need not be connected by a possible doxastic dynamics. Still, the principle thereby appears evident. If, by definition, varying attitudes are possible towards an a posteriori proposition, one should not be so dogmatic to fix one's attitude once and for all.

(2) Now let A be a proposition about a single variable which does not belong to the exceptions already admitted, i.e., which is a posteriori and not a fact of consciousness, and which is thus believed to a non-extremal degree. How can this degree change? Mainly by obtaining reasons for or against A, that is, by coming to believe or, more generally, by changing the degree of belief in other propositions which are positively or negatively relevant to A so that the belief in A changes its degree as well. Now, if A would violate the special coherence principle, there would be nothing that counts for or against it, there would be no way to change the degree of belief in A – in contradiction to the rationality postulate in (1).[28]

[28] Of course, we always suffer from a large and grave practical inaccessibility of reasons, simply because our experience is so restricted in space and time. The case at hand is worse, however; there, the non-existence of reasons would be irrevocably fixed in the internal structure of the doxastic state.

(3) The proof in (2) leaves a gap, however. The degree of belief in *A* may also change directly, not mediated by changes concerning other propositions. Indeed, the foundationalist will point out that this is the case with basic propositions as he conceives them, namely, as propositions which do not find any reasons outside themselves and are thus defined to violate the special coherence principle *without* necessarily being facts of consciousness. One may rejoin that this definition is empty because basic propositions are certainly used as reasons for other propositions, and the symmetry of the reason relation then entails that these allegedly basic propositions have reasons as well. However, this rejoinder has two shortcomings. First, nothing has been said so far to exclude the strange case of a basic proposition which is not good for justifying anything else; and secondly, the symmetry of the reason relation is, of course, something the foundationalist cannot accept. So, the proof in (2) needs some amendment.

(4) To this end we should first ask: What are the basic propositions in the foundationalist's sense? There is no perfect agreement, as far as I see, but the usual answer is that basic propositions are *directly perceived* propositions. What, in turn, are these? Some say, or think they are forced to say, that directly perceived propositions are facts of consciousness having purely phenomenal qualities as their contents. However, directly perceived propositions then reduce to a class of exceptions which we have already seen not to serve the foundationalist's purposes. So we may dismiss this reduction.

There is a more fruitful notion of direct perception according to which other propositions can be directly perceived as well. It runs as follows: Let us first assume that we can distinguish doxastic changes caused by perception from other doxastic changes (due to new concepts, drugs, forgetfulness, etc.). Changes through perception are usually accounted for by rules of conditionalization.[29] Now it is easy to check that these rules have the following property: Given the prior and the posterior doxastic state, and given that the change from the former to the latter was governed by a rule of conditionalization, the minimal set of propositions relative to which conditionalization was applied is uniquely determined; we may call this minimal set the *source* of the change. It seems then appropriate to say that the proposition(s) directly perceived in a perception is (are) just the proposition(s) in the source of the change brought about by the perception.

It must be emphasized that it is possible, but certainly exceptional that facts of consciousness are directly perceived in this sense. Usually, directly perceived propositions are public and in principle perceivable for many observers. Moreover, directly perceived propositions then stay firmly within the circle of reason; there is no need to exempt them from the circle. Their distinctive role rather lies, according to the account given, in the role they play in doxastic changes.

[29] In probabilistic terms these rules are simple conditionalization and generalized conditionalization as introduced by Jeffrey (1965, ch. 11). These rules can also be stated in terms of ranking functions; cf. Spohn (1988, sect. 5) [here: sect. 1.5].

(5) The next step is to introduce the following standard theory of perception: If x directly perceives that A (and if A is not a fact of consciousness), then A is a cause of the fact that x believes (more firmly than before) that A. Despite many disagreements concerning the theory of perception this seems to be one uncontested corner-stone.

A more contested question is, among others, whether this theory can be turned into an analysis of (direct) perception. The answer must be negative, it seems; there are certainly many propositions not directly perceived (or not perceived at all) for which this causal relation also obtains. People have tended then to require that it is this causal relation which must be somehow direct. However, this only leads to completely assimilating directly perceived propositions to facts of consciousness. But this seems wrong: the directness does not lie in the causal relation, but in the kind of belief change, as is also expressed in the familiar assertion that the directly perceived is non-inferentially known.[30]

(6) Now we may finally close the gap left in (2) and noticed in (3). The gap was that A may also be a basic proposition, i.e., a directly perceivable proposition which is not a fact of consciousness, which may directly change its degree of belief, and which thus appears to have no reason. This appearance is, however, refuted in five steps. First, since A is directly perceivable, it is possible that some observer x directly perceives that A. Suppose, secondly, that I believe in the above uncontested theory of perception. Then I believe that, given that x perceives that A, A is a cause of x's belief that A (where, however, we should exclude the case where the perception is my own present one). Thirdly, we may assume that whenever I believe that B is a cause of C, then B is also a reason for me for C (in the sense defined above), and vice versa.[31] This entails, fourthly, that under the condition that x perceives that A, A is a reason for me for assuming that x believes that A, and vice versa. If A is far-fetched, this condition will be far-fetched, too. Still, it is a posteriori and its falsity not maximally certain. Then, fifthly, some further mild assumptions[32] will turn the conditional reason relation into an unconditional one. Hence, the special coherence principle holds even for all directly perceivable or basic propositions.

How did I thereby avoid the two shortcomings noted at the end of (3)? First, I refuted the strange case of a basic proposition which is not a reason for anything else by specifying for each basic proposition another proposition for which it is a reason. And second, I think the foundationalist can concede that an effect is a reason to infer the cause, just as the cause is a reason to infer the

[30] Cf., e.g., Armstrong (1968, p. 234). I am well aware that steps (4) and (5) move on highly controversial grounds. However, in pursuit of the argument I want to give it may be legitimate to cut just one aisle through these grounds.

[31] Indeed, this assumption is a theorem of my theory of causation, given some weak restrictions; cf. Spohn (1991, p. 188 and notes 51, 54, and 55) [here: pp. 230f.]. Because of its plausibility I take this theorem rather as confirming that theory.

[32] Cf. again Spohn (1991, p. 188 and notes 54 and 55) [here: pp. 230f.].

effect, i.e., that at least in the case considered the reason relation is indeed symmetric.

(7) This may seem an improperly long-wided argument in favor of a fairly weak principle. The only excuse I have for proposing it is that I see no other argument extending to the general coherence principle as well. But now the extension is straightforward.

Consider any partition $\{V, U - V\}$ of the set U of all variables, and let $S = P(V)$ be the set of all propositions over V and $T = P(U - V)$ the set of all propositions over $U - V$. Then one possible case is that S contains all directly perceivable propositions and T none. In this case, however, the reason relation must relate S and T. Otherwise, nothing whatsoever could be found out about propositions in T, nothing could change my degree of belief in propositions in T. Again this contradicts the rationality postulate stated in (1). The same holds for the case where T contains all directly perceivable propositions and S none.

The final case is where both S and T contain directly perceivable propositions. Hence assume that $A \in S$ and $B \in T$ are directly perceivable. According to the above theory of perception, A is a cause of the fact that a given perceiver x believes that A, and B is a cause of the fact that x believes that B. Then, the trick goes, x also believes that $A \wedge B$. Both A and B are then partial causes of x's belief that $A \wedge B$.[33] Hence, if I believe in this theory of perception, then, as in (5), A as well as B are reasons for me for the proposition that x believes that $A \wedge B$, and vice versa.

Where, however, is the proposition that x believes that $A \wedge B$? It may not be totally clear which variables describe the doxastic state of x (at a certain time t). Let us try the two most plausible proposals. The most coarse-grained procedure would be to assume a single variable with a rich range consisting of all possible states x might be in (at t). But then both A and B are reasons to assume x to be in a certain doxastic state. There is thus at least one reason relation between S and T, since this rich variable must be either in V or in $U - V$. The most fine-grained procedure would be to assume a separate variable for each proposition taking all the possible degrees of belief of x (at t) as possible values. Then it is the variable for x's degree of belief in $A \wedge B$ which must be either in V or in $U - V$. And again there must be a proposition in S and another in T which are related by the reason relation. This finishes my proof of the general coherence principle.

Let me briefly sum up: I hope to have made clear the relevance of the two coherence principles discussed here and thus also the relevance of providing some argument for them. Clearly, I have offered only an argument sketch; but I believe that the steps and premises I have suppressed do not invalidate my argument. There were, however, a number of important premises. Some of them were linguistic, consisting in the explications of crucial notions I have used in the course of the argument. But there was also a substantial premise, namely, the rationality principle stated in (1). Moreover, I have introduced two assumptions. First, the proof of the special

[33] I use here "partial cause" for emphasis and not as a new term. Here, as in every-day language, "cause" always means "partial cause".

coherence principle assumed that the subject believes in the theory of perception mentioned in (5). Second, the extension to the general coherence principle additionally relied on (the subject's belief in) the capability of an arbitrary perceiver to form conjunctive beliefs. In this way the line of reasoning envisaged at the beginning of Section 10.6 and modified at the beginning of this section could be made to work. Whether this is a trivial or a significant result I do not dare to assess.