Can Probabilistic Coherence be a Measure of Understanding?

Victor GIJSBERS

Received: 12/06/14
Final Version: 21/11/14

BIBLID 0495-4548(2015)30:1p.53-71
DOI: 10.1387/theoria.11733

ABSTRACT: Coherence is a measure of how much our beliefs hang together. Understanding is achieved when we see that something is not just a brute, isolated fact. This suggests that it might be possible to use the extant probabilistic measures of coherence to formulate a measure of understanding. We attempt to do so, but it turns out that a coherence theory runs into trouble with the asymmetry of understanding. We identify four difficulties and show how they have been solved by a unification approach to explanation. We also identify four advantages of the coherence approach, and assess the possibilities of reconciling the strengths of the two approaches.

Keywords: coherence, understanding, explanation, unification, asymmetry

RESUMEN: La coherencia es una medida sobre cuánta interdependencia hay entre nuestras creencias. La comprensión se obtiene cuando vemos que algo no es un hecho bruto, aislado. Esto sugiere la posibilidad de usar las medidas probabilísticas de coherencia al uso para elaborar una medida de la comprensión. Exploraré esta posibilidad, a pesar de que una teoría de la coherencia tropieza con problemas aquí dado el carácter asimétrico de la comprensión. Identifico cuatro dificultades al respecto y muestro cómo han sido resueltas mediante un enfoque de la explicación basado en la noción de unificación. También identifico cuatro ventajas en la aproximación desde la coherencia y examino qué perspectivas hay de reconciliar los puntos fuertes de ambos planteamientos.

Palabras clave: coherencia, comprensión, explicación, unificación, simetría

1. Introduction

Philosophers have used the concept of coherence to analyse several of the central notions of epistemology: truth (e.g., Blanshard 1937), knowledge (e.g., Davidson 1986), and especially justification (e.g., BonJour 1985). Not all of these attempts have been equally successful, but the concept of coherence remains central to contemporary epistemology. Starting with Shogenji (1999), there has been a lively debate about ways in which coherence can be given a formal probabilistic definition, with important contributions being Olsson (2002), Fitelson (2003), Bovens & Hartmann (2003) and Douven & Meijs (2007).

Meanwhile, philosophers of science have become more interested in the notion of understanding (see De Regt, Leonelli & Eigner 2009 and De Regt 2013 for good overviews). Understanding is supposed to be the result of successful explanation, but it may also be available from other sources, such as visual models (Lipton 2009) or unification (Gijsbers 2013). No real attempt at formalisation has been made, although Schurz (1999) and Gijsbers (2014) suggest that Schurz and Lambert's (1994) formal measure of unification might capture at least part of it.

Several authors have suggested that coherence and understanding are closely related (e.g., Kvanvig 2003, Riggs 2003, Elgin 2012). These proposals, however, have not taken full advantage of the formal developments in the field, and have remained relatively vague. The aim of the current paper is to explore whether the formal probabilistic coherence of a subject's beliefs can be used as a measure of her understanding of specific phenomena being the case.[1] Ideally, such a coherence theory of understanding would take the form of a formula like this:

$$U_S(p) = [\ldots]$$

where $U_S(p)$ is subject S's understanding of phenomenon p, and the right side of the equation applies a formal measure of coherence to S's body of beliefs in general and p's place in it in particular.

We will not succeed in finding such a formula in this paper. Instead, we will be using several simple coherence theories of understanding to assess how promising such theories actually are. Of special interest to us will be the problem of asymmetry – well-known from the literature on explanation, and reappearing, although in slightly different form, in our intuitions about understanding – and the demands that a solution to that problem leads to. We identify four difficulties and show how they have already been solved by Schurz and Lambert's unificationist approach to explanation. On the other hand, we also identify four advantages of the coherence approach. We then assess the prospects for unifying the strengths of the two theories, and conclude that the most promising suggestion is that both capture a part of our idea of understanding.

We start, in section 2, by exploring several reasons for thinking that coherence and understanding must be closely linked. We also clarify the context of the current investigation by pointing out some of the issues and theories that we will ignore. In section 3 we introduce two of the formal measures of coherence and two possible ways of linking these measures to understanding. In section 4, we introduce the problem of asymmetry, and show that our measures cannot handle it. In section 5, we discuss whether our formal theories of coherence can be improved. Four problems and potential solutions are identified. In section 6, we compare Schurz and Lambert's unificationist theory of explanation with our coherence approach and argue that both have advantages which would be hard to replicate in the framework of the other. In section 7, we conclude that a pluralistic approach is the most promising.

2. *Coherence and understanding: love at first sight?*

There are several reasons for believing that coherence and understanding are closely related. True, coherence is a property of sets of propositions, while understanding is something that applies to individual elements of such a set. And coherence is often —though not invariably— supposed to be an objective property of a set of propositions, while understanding is

[1] My focus, then, will be on what has been called *explanatory understanding* rather than *objectual understanding*; see Kvanvig 2009 and Khalifa 2013. However, as has just been indicated, it is certainly open to doubt whether explanatory understanding must result from explanations, which makes this terminology less than ideal.

something that cannot exist without a subject. But these differences also exist between coherence and justification, and are easily circumvented – after all, our claim is not that the two concepts are *identical*. Thus, one possible way to spell out the (hypothetical) relation between understanding and coherence is this:

> A subject S with a set of beliefs SB understands the fact that p if and only if p is in SB, and SB is more coherent than SB – p.

Here the idea is that a subject understands p just in case she believes p, and p coheres with her other beliefs in the sense that removing p makes her total set of beliefs less coherent. We will call this way of spelling out the relation between coherence and understanding the "addition option", since it looks at what happens to the coherence of a set of beliefs when p is added to it. An example: I believe that studying philosophy will make you wise; that Friedrich has studied philosophy; and that Friedrich is wise. Given these three beliefs, I surely understand why Friedrich is wise. And indeed, a reasonable notion of coherence will imply that this set of three statements is more coherent than the set containing only the first two.

There are other possibilities for spelling out the relation. For instance, one could compare the coherence of the subject's current beliefs with that of her beliefs if she were to stop believing p and start believing not-p instead. We will call this way of spelling out the relation between coherence and understanding the "alternatives option". Instead of merely dropping the belief that Friedrich is wise, I replace it with the belief that he is unwise. That would presumably give me a less coherent set; and so we again reach the conclusion that I understand why Friedrich is wise.

These examples give us a (still weak) reason to believe that a coherence theory of understanding is possible. But there are also deeper reasons. Consider what Douven and Meijs (2007) claim to be "our most basic intuitions" about coherence: "that coherence is a matter of hanging or fitting together, and that coherence is a matter of degree" (p. 405). Shogenji (1999) agrees (p. 338).

These are also two of our most basic intuitions about understanding. That there are degrees of understanding is uncontroversial. A folk theory about drunkenness gives me some understanding of why people can't drive straight after drinking six beers, but I can deepen that understanding by learning more about the chemical composition of beer and the neurological effects of ethanol.

Even more central is the intuition that understanding is a matter of hanging together. Kosso (2002) dramatised this intuition by coming up with figure of the Omniscienter, a being who knows every particular fact about the universe, but is unable to see any connections between them. Kosso claims, and we can surely agree with him, that the Omniscienter understands nothing. To understand something *is* to grasp certain connections; for instance, between an event and its causes, or between an event and the law of nature which it instantiates. To understand something is to see that it is not just a brute fact.

Since coherence is supposed to be a non-binary measure of how well beliefs hang together, and understanding is a non-binary state that a subject is in only when she grasps certain connections between her beliefs, the idea of a coherence theory of understanding looks very natural. More natural, perhaps, than the idea of a coherence theory of justification.

For at least at first sight, it seems to be possible for S to be justified in believing *p* even if there are no links between *p* and any other beliefs of the subject. Aren't there, the argument would run, propositions that are known non-inferentially, perhaps through sense perception? Isn't it the case that no other beliefs have to play a role in that process? And doesn't that mean that coherence – even if it can perhaps play a role in some situations – could not possibly be constitutive of justification?

This chain of reasoning can be resisted, of course (the locus classicus is perhaps Sellars 1956, though the argument can be traced back much further). But in the case of understanding, this problem does not even arise, because it is obvious that understanding *p* will *always* involve other beliefs. Non-inferential understanding, if it exists, creates no trouble for a coherence theory, because non-inferential understanding would still be a grasping of certain kinds of relations between beliefs. So the idea that understanding must be closely linked to coherence comes up naturally.

The idea that coherence and understanding are related is not new. Elgin (2012) speaks of the suggestion "that coherence is the hallmark of understanding" (though she does not endorse the suggestion without qualifications). Kvanvig (2003) suggests that a subject has understanding when she possesses "a body of information together with the grasping of explanatory connections concerning that body of information". Riggs (2003) claims that "the epistemological notion of 'coherence' and the idea of 'explanatory coherence' in particular seem to be getting very close to something characteristic of understanding."

To a certain extent, such claims are unproblematic. Understanding involves connections between beliefs, and therefore something like coherence. But it is one thing to endorse this broad claim, and something else to say that understanding must be analysed in terms of coherence, or that measures of coherence are the tools we need to construct a measure of understanding. It is this more precise and less obvious claim that we will investigate in the rest of the paper using several formal measures of coherence.

Some remarks about the limitations of this investigation need to be made in advance. Since our concern is solely to establish whether there is a relation between understanding and formal measures of coherence, we will bracket several of the questions that have been raised in the context of coherence theories of understanding. First, we will bracket the question of whether and how a coherent set of beliefs must be *tethered* to the actual facts; for instance, by being true, or by having been caused by the facts. (See Elgin 2012 for a discussion of tethering. In ignoring this question, we will also ignore the issue of whether a theory of understanding needs to satisfy anti-Gettier conditions, something that Kvanvig 2003 denies and DePaul & Grimm 2007 affirm.) Second, we will bracket the question of whether and in what sense the subject must not only have a set of coherent beliefs, but must also *grasp* that coherence (Elgin 2007). Third, we will bracket the question of whether the subject must be able to *use* the claims she understands, e.g., whether she has to be able to see some of the consequences of those claims. (Such use plays a crucial role in the well-known theory of understanding developed by Dieks & de Regt 2005.) We will simply assume that in all our examples, any requirements for understanding of these types, whatever they may turn out to be, are satisfied – that the beliefs we mention are true, grasped, usable, and so on.

3. *Introducing formal coherence theories*

Recent attempts to formalise coherence have used a probabilistic framework: the coherence of a set of propositions is defined in terms of the probabilities of those propositions or their logical combinations. Let us look at two examples. Take the set A of propositions <A_1, A_2, ..., A_n>. Shogenji (1999) then defines the coherence of that set as:

$$C(A)=\frac{P(A_1 \wedge A_2 \wedge \ldots \wedge A_n)}{P(A_1)\times P(A_2)\times \ldots \times P(A_n)}$$

that is, the joint probability of the propositions divided by their multiplied individual probabilities. Olsson (2002) proposes an alternative measure:

$$C(A)=\frac{P(A_1 \wedge A_2 \wedge \ldots \wedge A_n)}{P(A_1 \vee A_2 \vee \ldots \vee A_n)}$$

that is, the joint probability of the propositions divided by the probability of their disjunction. Both measures aim to capture the intuitive idea that a set of propositions is coherent when the propositions in the set probabilistically strengthen each other.

The literature contains several other probabilistic measures of coherence, see e.g. Fitelson (2003) and Bovens & Hartmann (2003), and, for an overview and discussion, Douven & Meijs (2007). We will focus on the measures of Shogenji and Olsson partly for reasons of simplicity, and partly because we will only use them as diagnostic tools for finding the difficulties for coherence theories of understanding. Using other measures would not, we believe, substantially alter the discussion.

In order to see how these measures work out in practice, let us look at the following set of statements:

A_1 = Studying philosophy will make you wise.
A_2 = Friedrich has studied philosophy.
A_3 = Friedrich is wise.

Suppose that A_1 is certain; that $P(A_2) = 0.02$; and that $P(A_3) = 0.1$. Since in this case all philosophers are wise, the joint probability of these three statements is equal to the probability of A_2, 0.02.

According to Shogenji's measure, the set containing all three propositions has a coherence of 0.02 / (0.1 * 0.02) = 10. This is quite a big bigger than 1, which is the neutral point on Shogenji's scale, and therefore indicates large coherence. According to Olsson's measure, the coherence of the set is 0.02 / 1 = 0.02. This number lies on a scale between 0 (completely incoherent) and 1 (completely coherent); presumably, the right interpretation of 0.02 is "quite incoherent".

We see that these two measures of coherence can yield rather different results when asked to assess the overall coherence of a set of propositions. This can be seen especially clearly when we take a set containing just one proposition. Both formulas yield the number 1, but for Olsson this indicates maximal coherence, while for Shogenji it indicates the balance point between coherence and incoherence.

However, for our purposes we are not interested in the total coherence of a set. The idea we wish to investigate is that we understand why *p* just in case that adding *p* to the rest of our knowledge increases its coherence (the addition option); or perhaps just in case that adding *p* to the rest of our knowledge leads to a more coherent set than adding not-*p* (the alternatives option).

In our example, we presumably understand A_3, that Friedrich is wise. Let us see how the two options and the two measures of coherence fare.

	Shogenji	**Olsson**
$C(A_1, A_2)$	0.02/0.02 = 1	0.02/1 = 0.02
$C(A_1, A_2, A_3)$	0.02/0.002 = 10	0.02/1 = 0.02
$C(A_1, A_2, \text{not-}A_3)$	0/0.018 = 0	0/1 = 0

Using the Shogenji measure, both the addition option (10 > 1) and the alternatives option (10 > 0) give the correct assessment that we understand A_3. Using the Olsson option, only the alternatives option (0.02 > 0) gives the correct assessment, while the addition option suggests that there is no understanding (0.02 = 0.02).

It may be objected that the epistemic situation we have sketched is unrealistic, since we assumed that the hypothesis that studying philosophy makes you wise is absolutely certain. The limit cases with probabilities 0 and 1 sometimes lead to unwanted but unimportant results in probabilistic epistemology. So let us redo the calculation with these new assumptions: $P(A_1) = 0.9$; if $P(A_1)$ is false, it is instead the case that philosophy and wisdom are unrelated, and thus A_2 and A_3 are probabilistically independent. We also assume that A_1 is probabilistically independent of A_2. The new table then looks like this:

	Shogenji	**Olsson**
$C(A_1, A_2)$	0.018/0.018 = 1	0.018/0.902 ≈ 0.01996
$C(A_1, A_2, A_3)$	0.018/0.0018 = 10	0.018/0.9118 ≈ 0.01974
$C(A_1, A_2, \text{not-}A_3)$	0/0.0162 = 0	0/0.9902 = 0

We see that the verdicts remain the same, with Olsson's measure combined with the addition option now even giving a negative rating to our understanding of A_3. (From now on, I will continue using the $P(A_1) = 1$ version of the example, because that makes the calculations much easier to follow.)

This means that the Shogenji-addition test, the Shogenji-alternatives test, and the Olsson-alternatives test are on the table as potentially correct tests for assessing understanding. Each of them allows us to set up a formula for the amount of understanding that a certain subject S has. For the addition option, we would have

$$U_S(p) = F(C(B_S, p), C(B_S))$$

with $U_S(p)$ the amount of understanding that S has of P, with C the chosen measure of coherence, and with F some function that adequately picks out the difference in coherence. (F might just be subtraction, but whether that is appropriate will depend on the exact nature of C.) The alternatives option would lead to a formula of the form

$$U_S(p) = F(C(B_S, p), C(B_S, \text{not-}p))$$

At this point, however, rather than develop these theories further, we need to investigate a general counterargument against the very possibility of probabilistic coherence being a measure of understanding.

4. An asymmetry problem

Any reader familiar with the literature on explanation may wonder whether a coherence theory of understanding based on a probabilistic notion of coherence can handle the counterarguments that were used to undermine the DN-model of explanation. Specifically, we may wonder whether a coherence theory can adequately capture the *asymmetry* of understanding, which it inherits – at least to a certain extent – from the asymmetry of explanation. We will first look at how this counterargument might go in general; then we will look at a more realistic example where the measures of understanding developed in the previous section indeed deliver the wrong answer; and finally, we consider several responses to the problem.

The general argument goes like this. A purely probabilistic measure of coherence defines coherence in terms of the probabilities of propositions and logical combinations of those propositions. But we can show that two propositions can stand in the exact same probabilistic relations to the other propositions in a set, while having a very different status in terms of our understanding of them.

Consider the following three propositions:

B_1 = For all x, F(x) → G(x).
B_2 = F(a).
B_3 = G(a).

Let $P(B_1)$ be 1. Then F(a) → G(a), and B_2 and B_3 are probabilistically indistinguishable. But they might nevertheless be distinguishable in terms of understanding, for instance when B_1 is an explanatory law stating that property G is caused by property F. For example, F could be the property "is a moving charged object", and G the property "would be deflected when moving through a magnetic field". When we know that a certain object is moving and charged, we presumably understand why it would be deflected by magnetic fields. (If one feels that B_1 alone is too small a basis for understanding, one can add the laws of Maxwell and other central elements of electrodynamics to the set without ruining the example.) But when we know that the object would be deflected by magnetic fields, we of course *know* that it is moving and charged, but we do not yet *understand why* this is the case – or at least, we do not understand much of it. For that, we would presumably need to hear something about the history of the object.

Any purely probabilistic measure of coherence, when converted into a test of understanding, must make the same pronouncement about B_2 as about B_3. But we understand B_3 and do not understand B_2, so one of these pronouncements has to be wrong. Hence, probabilistic measures of coherence cannot function as measures of understanding.

Is this asymmetry problem the result of what might perhaps be a highly unrealistic assumption, namely, that there can be different properties that are perfectly correlated, and that the probability of this correlation holding can be 1? In order to show that we are not led astray by unrealistic assumptions, and to get a better sense of where and how our measures of coherence go wrong, let us look at how the measures developed in the previous section fare in a more realistic example.

We return to Friedrich the philosopher. Presumably, we understand why Friedrich is wise, but we do not yet understand why Friedrich studied philosophy or why philosophy makes people wise – or, if we do, only to a truly minuscule extent. Do any of our tests lead to the right results here? Let us look at the coherence effects of adding A_2. (Because both measures ignore the order of the set's elements, the second line of the table will not change from the table in the previous section. Note that the prior probability of someone being a wise non-philosopher is 0.08.)

	Shogenji	**Olsson**
$C(A_1, A_3)$	$0.1/0.1 = 1$	$0.1/1 = 0.1$
$C(A_1, A_3, A_2)$	$0.02/0.002 = 10$	$0.02/1 = 0.02$
$C(A_1, A_3, \text{not-}A_2)$	$0.08/0.098 \approx 0.82$	$0.08/1 = 0.08$

Using Shogenji's measure, both the addition option and the alternatives option falsely claim that we understand A_2 to a significant extent. Using Olsson's measure, both options correctly state that we do not.

But if we think through how Olsson's measure works, that success is not very reassuring. Take the addition option first. It will *always* result in the verdict that there is no understanding. For given two sets A and B such that A is a subset of B, Olsson's measure will *never* give higher coherence to B than to A. One can easily see this by looking at the formula: if we add more propositions to the set of A_is, the numerator can only remain the same or go down, while the denominator can only remain the same or go up. This means that the combination of the addition option and Olsson's measure will get all example where there is no understanding right, but it will also get all examples where there is understanding wrong.

In the case of the alternatives option, Olsson's measure leads to the right conclusion only because $P(\text{not-}A_2 \ \& \ A_3) > P(A_2)$; that is, because in our example there are more wise non-philosophers than philosophers. Suppose that wisdom goes in decline, and now only 3% of the entire populace is wise, while the population of philosophers, all of whom are wise, remains 2%. Our new values are:

	Olsson
$C(A_1, A_3)$	$0.03/1 = 0.03$
$C(A_1, A_3, A_2)$	$0.02/1 = 0.02$
$C(A_1, A_3, \text{not-}A_2)$	$0.01/1 = 0.01$

Now the alternatives option claims that we *do* understand why Friedrich is a philosopher. Furthermore, we can change the number in the final columns as we wish by choosing different percentages of non-philosophers who are wise, even though that percentage has nothing to do with our understanding of A_2, the claim that Friedrich is a philosopher. So it seems that the Olsson-alternatives option doesn't give us a defensible measure of understanding either. None of the four options we have surveyed, then, yields the right answers even in this simple and realistic case of asymmetry.

There are several ways to respond to the general argument, and *a forteriori* to the special case described above. One is of course to abandon the project we have engaged in entirely, but that would be admitting defeat too quickly. Let us instead look at a response that attempts to undercut the argument itself. That explanation is asymmetric is generally accepted (Van Fraassen 1980 is one of the rare exceptions, and his tower argument has not received much support). But, we may wonder, is the same true of understanding? Can it not be the case that while A helps me understand B, B also helps me understand A?

Now, it seems that this response is correct to the extent that it is much harder to argue for the *radical* asymmetry of understanding than it is to argue for the radical asymmetry of explanation. Especially on a causal theory of explanation, one can claim that, since there is no backwards causation, if A is part of an explanation of B, B cannot be part of an explanation of A. The argument would be non-trivial, but might be defensible. However, if understanding can be gained by means other than explanation, it is not easy to see why we would believe that A and B can never give understanding of each other; and so there is no radical asymmetry of this kind.

But of course, the counterargument given above does not depend on the claim that understanding is radically asymmetrical. It merely depends on the claim that it is possible for two probabilistically indistinguishable propositions to be understood to a different extent. Perhaps we have *some* understanding of B_2 due to its coherence with B_1 and B_3. But it seems evident that this understanding is much *less* than our understanding of B_3. And that is all we need for our counterexample to work: there is an asymmetry in understanding that our coherence measures cannot capture.

Another way to respond to the problem is to suggest that we might need to search for measures that are more appropriate for capturing the asymmetry of understanding; perhaps we can avoid trouble and still stay close to our original ideas. This is what we will do later in the paper. But for now, we want to push the probabilistic coherence measures as far as possible.

Yet another way to respond is to note that the counterargument merely shows that our coherence measures cannot function as measures of understanding on their own. But this leaves open the possibility that we could add some other tests to them which can weed out or adjust for unwanted results. We noted in our preliminary discussion that understanding always involves hanging together, and that coherence measures how much propositions hang together. But our discussion of asymmetry suggests that understanding is *more* than hanging together, since hanging together would seem to be a prototypically symmetric relation. So perhaps coherence is a *necessary* ingredient of understanding, but not a *sufficient* one. Something else is needed, something that generates the asymmetries of understanding. In the next section, we will construct a 'wish list' of things that need to be added to our coherence measures to turn them into a satisfactory theory of understanding. Afterwards, we will pose the question whether such a theory is feasible, and whether it could still be called a coherence theory of understanding.

5. *Making coherence asymmetric: four demands*

In this section, we will argue that a coherence theory of understanding needs to be changed in four ways if it is to deal satisfactorily with the asymmetry of understanding:

1. One must ensure that understanding cannot be gained from cohering with statements lower in the hierarchy of understanding (a term to be explained shortly). For this one needs to have a way to distinguish the levels of this hierarchy.
2. One must ensure that understanding cannot be gained from duplications of content; some way of analysing propositions into their component parts and removing any duplicates is necessary for this.
3. One must find a way to ensure that understanding depends on the presence *and* the content of connecting principles.
4. One must find a way to respect the asymmetry of causal principles.

We will start by developing the idea of a hierarchy of understanding. Let us call those features of the world that primarily call out for explanation the *phenomena*; we can remain agnostic about their precise nature, but some philosophers will want to identify them with our perceptions while others may choose to identify them with singular facts that are relatively close to perceptual verification.

Now understanding consists in the grasping of certain kinds of connections. But, at least at first glance, phenomena as singular facts have only two kinds of connections to each other, neither of which is enough for understanding. First, there are spatio-temporal relations. But merely pointing out spatio-temporal relations does not make us understand phenomena (though these relations may play an important role when connected to theories that involve spatio-temporal relations). Second, there are the purely logical relations that obtain between *aggregate and component phenomena*. For instance, the phenomenon that John is a male human is an aggregate phenomenon that logically implies the component phenomenon that John is a human. But the purely logical relations between aggregates and components do not give an understanding of the components; or, if they do, only a very meagre understanding at best.

Phenomena, then, lack the connections to each other that are necessary for understanding. These connections have to be added, and this is precisely the task of theories; or rather, of that which theories express, which we can call *connecting principles*. Many things can play the role of connecting principles. Some candidates are causal laws, other laws of nature, principles of reduction, metaphysical principles, natural classifications, and the transcendental laws of human cognition. In this paper, we will be neutral with regard to the plausibility of any of these candidates.

However, as is well known from the discussion of the difference between lawlike and *accidental patterns*, not every general statement can give us (significant) understanding of the particular phenomena that it implies.[2] The accidental truth that "everyone in this room

[2] It is possible that in cases like these —and some of the other cases below— a small amount of understanding is given. The asymmetry we are interested in in this section need not be the asymmetry between full and no understanding, but can be an asymmetry between much and little understanding. Since our coherence theory of understanding admits of degrees, and therefore must get degrees of understanding right, such cases are relevant.

has a beard" added to the particular fact that Victor is in the room doesn't give us much, if any, understanding of why Victor has a beard.

In general, the direction of understanding seems to be a hierarchy like this:

connecting principles → phenomena → accidental patterns.

Connecting principles like causal laws make us understand phenomena. They generally do not do this by themselves, but with the help of other phenomena; the connecting principles form the background, as it were, against which one phenomenon (Friedrich's being a philosopher) can increase our understanding of another phenomenon (Friedrich's being wise). For them to play this role, it is essential that we see the principles as expressive of real connections and not just as accidental patterns. If we had any reason to suspect that the correlation between being a philosopher and being wise was spurious, just an accident, much of our understanding of Friedrich's wisdom would evaporate.

Accidental patterns, in turn, can be understood by being derived from particular phenomena. We understand why everyone in this room has a beard when we understand that it is an accidental truth that follows from the particular facts that only Victor, John and Hassan are in the room, and that Victor, John and Hassan all have beards. (In general, it seems intuitively true that if we understand some body of beliefs, we then also understand all of its logical consequences.)

A good theory of understanding should be able to correctly tell us that connecting principles give far greater understanding of the phenomena than accidental patterns do; and it might have to do this by explicitly incorporating a the idea of a hierarchy of understanding. This would solve at least some cases of asymmetry. (This is demand #1 from the list above.)

What we can try to do, then, is take a test like the Shogenji-addition or the Olsson-alternatives test, and add the following stipulation: *p* is understood if and only if the test is passed *and* this passing of the test is at least partly due to *p*'s relations with propositions higher up in the hierarchy of understanding. (For ease of expression, we will ignore the difference between propositions and the principles, patterns and phenomena that make them true.) So if *p* is an accidental pattern, and it coheres because of the presence of phenomena, *p* is understood. But if *p* is a connecting principle, and it coheres because of the presence of phenomena, it is not understood, because phenomena are lower in the hierarchy of understanding.

A problem confronts us immediately. Coherence is a global property of sets of propositions. Given any new proposition *p*, we can check whether adding it to a set *S* increases or decreases the coherence of that set. But how can we see *what parts* of *S* are responsible for the increase or decrease? Does it even make sense to ask such a question? Isn't it the set *as a whole* that is responsible?

To make this worry more concrete, let us consider the set C which consists of the following five propositions:

C_1 = F-ness → G-ness
C_2 = F(a)
C_3 = F(b)
C_4 = G(a)
C_5 = F(a) & G(b)

Here C_1 is a connecting principle which implies that all Fs are Gs, while C_5 is an accidental pattern that does not explain either of its conjuncts. Let us assume that the probability of anything being an F is 50%; the probability of anything being a G is 50%; that the probability of C_1 is 50%; and that if C_1 is false, being F and being G are probabilistically unrelated. (It follows from these probabilities that if C_1 is true, anything that is G has a 100% chance to also be F.)

Using this background set of beliefs, we want to consider whether we understand

$$D = G(b).$$

We presumably do, and indeed, using the Shogenji measure we find:

	Shogenji
C(C)	0.125 / (0.5*0.5*0.5*0.5*0.25) = 8
C(C + D)	0.125 / (0.5*0.5*0.5*0.5*0.25*0.5) = 16
C(C + not-D)	0 / (0.5*0.5*0.5*0.5*0.25*0.5) = 0

so according to both the Shogenji-addition and the Shogenji-alternatives test, we understand D. (We will ignore the Olsson measure for now, since for the current discussion the difference between the two measures is unimportant and the calculations are already complicated enough.) We now want to apply our new extra criterion, so we must determine which part of C is responsible for the increase in coherence when we add D. One way to do that is to see whether the coherence still increases if we ignore one of the propositions in C. This should be a rough, if imperfect, guide to which propositions in C are responsible for the coherence.

But when we try to apply this procedure, we run into a question that we have not had to answer until now, namely, the question of what happens to conditional probabilities when a set of beliefs does not fully reflect the probabilistic reality. Take, for simplicity, this set:

$C_3 = F(b)$
$D = G(b)$.

Does it cohere? One way to think about it is that, no, it does not. We just have two particular facts, and they are not connected in any way. So we might want say that $P(C_3 \& D) = P(C_3) \times P(D) = 0.5 \times 0.5 = 0.25$; and then we would find that the set is indeed not coherent.

However, that would be the wrong move to make, since removing C_1 from our set of propositions doesn't alter the *objective probabilities* underlying the situation. The value of $P(C_3 \& D)$ doesn't depend on whether C_1 is or is not part of the set of propositions we are evaluating. Whether it is part of that set or not, $P(C_1) = 0.5$, and $P(C_3 \& D) = 0.5 \times P(C_3) + 0.5 \times P(C_3) \times P(D) = 0.375$. So the Shogenji coherence for the set $\{C_3, D\}$ is not 1, but 1.5. A set of just two seemingly unrelated observations can be coherent if there is an objective probability that the observations are probabilistically linked.

For our purposes, this is a terrible result. It threatens to undermine the entire idea that we could use coherence to measure how well the principles in a certain set of propositions allow us to understand the phenomena in that set. For such a set could cohere even if none of the principles connect any of the phenomena. Worse, consider a situation in which a principle Y connecting F-ness to G-ness is probabilistically correlated with a different principle X connecting H-ness to J-ness. (Perhaps both are part of the same overarching theory, and therefore likely to be true or false together.) Then X, C_3 and D would cohere, even in the strong sense that removing any of the three, or replacing any of the three by its opposite, will decrease coherence. And this seems to suggest that X gives us understanding of D, and indeed almost as much as Y. But it seems unlikely that X could give us more than a very tiny amount of understanding of C_3 and D.

One might hope to escape from this result by interpreting the probabilities as subjective degrees of belief of a subject. But that won't do either, because it is possible to assign subjective probabilities in such a way that the probabilistic structure of real understanding is copied even though the actual beliefs do not objectively grant such understanding. That is, it is possible to have belief assignments that do not grant understanding but that are nevertheless probabilistically indistinguishable from belief assignments that do. For example, take any situation in which the set consisting of connecting principle X and phenomena Y and Z passes some coherence-understanding test for Z. Now replace X with an arbitrary connecting principle X'. As long as the subject assigns the same probabilities to all statements involving X' as it did to those involving X (and that might well be possible even without making herself vulnerable to Dutch books), the new set will also pass the coherence-understanding test. But that means that the test must be wrong, because it is false that any phenomenon can be understood using any connecting principle whatsoever.[3]

A coherence theory of understanding must make understanding depend on two things: (1) the fact that the propositions in question are actually believed by the subject, that is, that they are part of the set whose coherence we are measuring; and (2) the objective validity of the probabilistic connections that obtain between those propositions. (This is demand #3 from the list above.) The objective interpretation of probability fails the first test, while the subjective interpretation fails the second.

I am not sure how to solve this problem, although we will see a suggestion in the next section. Let us assume that a solution can be found. Presumably, this solution will have the result that if we take set C and remove C_1 from it, the probabilistic connection between C_2 and C_4 disappears, as does that between C_3 and D. With this assumption, we can start anew our attempt to find out which part of C is responsible for the increase in coherence when we add D.

---

[3] One reviewer pointed out to me that it would be possible to hold that such belief substitutions are impossible, since the content of a belief is defined by its inferential relations to other beliefs. Hence, no two beliefs can have the same conditional probabilities assigned to them and still be other beliefs. I agree with this, but it seems to me that to avoid all examples of the kind I exploit here, one would have to assume an epistemically perfect subject. After all, the content of a belief is, on inferentialist theories, defined by the normatively correct inferential relations they have; but a normal subject can, of course, be mistaken about these. If misassignments of conditional probabilities are possible, then the kind of case I am exploiting would seem to be possible as well.

Let C-i be the set C without the statement C_i. Then, omitting calculations, we find the following table for the Shogenji coherence:

	i = 1	i = 2	i = 3	i = 4	i = 5
C(C-i)	2	4	4	4	2
C(C-i + D)	4	8	8	8	4
C(C-i + not-D)	0	0	0	0	0

What does this mean? If we take away any element of C, adding D still increases the coherence using both the addition and the alternatives tests. So we could argue that *no* statement in C is responsible for D's coherence with C. On the other hand, the coherence of C-i + D is smaller than that of C + D for all i; so we could also argue that *all* statements in C are responsible for D's coherence with C.

Perhaps we should look beyond single statements. If we take away both C_1 and C_5, the coherence increasing effects of adding D disappear completely. So perhaps we should argue that it is the combination of C_1 and C_5 that makes D cohere with C. But that is an unwelcome result, since we do not understand D because of the combination of a connecting principle and an accidental generalisation. And what may be even worse, the coherence increasing effects of adding D also disappear if we take away C_3 and C_5. But it really isn't the case that we understand D because of the combination of C_3 and C_5. (We presumably understand D because of the combination of C_1 and C_3.)

Is there a way out of these conundrums? Given that we are interested in whether or not we understand a certain phenomenon, we might perhaps be best off disregarding accidental patterns completely. After all, any coherence between the phenomenon and these patterns will be irrelevant to our understanding of the phenomenon – though it might be relevant to our understanding of the patterns. Since we are working from the assumption that on any level in the hierarchy of understanding lower-level statements are irrelevant, we can turn this into a general demand:

> When testing whether a proposition p increases the coherence of a set A, one first has to remove from A all propositions that are lower in the hierarchy of understanding than p.

Propositions on the same level must remain: it is generally only the combination of a connecting principle and some phenomena that allows us to understand another phenomenon.

So we should assess D's effects on coherence after we have purged all accidental patterns from C. That comes down to choosing C-5 as the relevant set. If D coheres with C-5, as indeed it does, can we then conclude that this coherence is due to connecting principles? For it seems as if it could also be due to just D's coherence with other phenomena.

That worry can't come true if phenomena by themselves do not cohere. And we suggested earlier in this paragraph that this is indeed the case: connecting principles are needed to connect up phenomena and allow us to understand them. So if a set of phenomena and connecting principles coheres with a new statement describing a phenomenon, that must be due to the connecting principles linking the new phenomenon to some of the other phenomena. Which is exactly what we need in order to show that this asymmetry of understanding is satisfied.

However, we do need to worry whether our current conception of a coherence theory of understanding actually *has* this desirable feature. We pointed out that there can be aggregate and component phenomena; e.g., John being a male human is an aggregate phenomenon that has as components John being male and John being human. Having statements describing all these phenomena in our set of propositions will generate coherence that doesn't have anything to do with different phenomena being linked up by connecting principles. So before we assess the coherence effect of adding a new proposition, we should eliminate these sources of spurious coherence by analysing all propositions into their smallest components, and removing any that are duplicated. We can turn this into a general demand as well (this is demand #2 from the list above):

> When testing whether a proposition *p* increases the coherence of a set A, one first has to analyse all propositions in A into their component parts, and remove all that are duplicated.

We will leave aside the difficult question of how to turn this demand into a precise formal recipe (though again we will see a suggestion in the next section).

Finally, we should note that nothing in our current conception of a coherence theory of understanding guarantees that we respect the asymmetries of causal explanation, such as the one we have seen in the case of Friedrich the philosopher. Given some general causal law linking F-ness to G-ness, we may well find that F(a) coheres with a set that contains G(a), and the other way around. But we believe that we gain more understanding of why events are the case by learning their causes, than we gain understanding of why events are the case by learning their effects. For causal principles and other asymmetric principles, we thus need some addition to our theory that compensates for any coherence in the wrong direction. (This is demand #4 from the list above.) Again, we will leave aside the difficult task of making that demand precise, and simply state that a satisfactory coherence theory of understanding must somehow do so.

To summarise, we have seen in this section that a coherence theory of understanding, if it is to deal satisfactorily with the asymmetry of understanding, must be changed in the four ways that we gave at the beginning of this section. To repeat:

1. One must ensure that understanding cannot be gained from cohering with statements lower in the hierarchy of understanding. For this one needs to have a way to distinguish the levels of this hierarchy.
2. One must ensure that understanding cannot be gained from duplications of content; some way of analysing propositions into their component parts and removing any duplicates is necessary for this.
3. One must find a way to ensure that understanding depends on the presence of connecting principles *and* the objective validity of the probability assignments.
4. One must find a way to respect the asymmetry of causal principles.

This is a hefty list, and one may wonder not just how to implement it, but also whether the theory that would be left after implementing it would still be a coherence theory. This is not an idle question; for the literature contains a theory that is *not* a coherence theory, but precisely *does* fulfil these four requirements, namely Schurz and Lambert's theory of unification. So in the next section, we will compare coherence with unification, and ask ourselves whether the idea of a coherence theory of understanding can be maintained in the face of the competition with unification.

6. Coherence and unification

Coherence theories of understanding face at least one competitor which already implements all of the four demands formulated in the previous section. Can they overcome this competition? We will first take a brief look at the competitor, Schurz and Lambert's unificationist theory of explanation (Schurz & Lambert 1994, Schurz 1999; short summaries can be found in Schurz 2014, Gijsbers 2014). We will then set out the ways in which the coherence theories might nevertheless be superior, and give a brief indication of ways in which these strengths might be developed.

Like the measures of coherence we have investigated, Schurz and Lambert's measure of unification is a global measure defined on the whole of a subject's beliefs, although in the case of Schurz and Lambert this set is first appropriately sanitised. Before measuring unification, Schurz and Lambert require us to transform the set of beliefs into the set of its *relevant elements*. This transformation is a complex logical procedure, but it comes down to analysing propositions into their component parts and removing any duplicates. The result is that all aggregate phenomena and accidental patterns are removed, thus fulfilling *demand 2*.

Unlike the measures of coherence, which try to quantify what are potentially symmetric relations of probabilistic dependence, the measure of unification is based on an asymmetric relation of assimilation. The beliefs of the subject are divided into two sets, *basic* and *assimilated* beliefs, where a belief is assimilated just in case it can be derived from basic beliefs by arguments known to the subject. Applying the theory to understanding, we can say that the assimilated statements are understood, while the basic statements are not. Statements are assimilated just in case an argument known by the subject allows their derivation from basic statements. There is a normative component at work here: the arguments must be *correct*. So in order to unify her knowledge, a subject must have general beliefs that actually do imply – in a deductive or in a probabilistic sense – her specific beliefs about the phenomena. This ensures that both the presence and the content of the subject's beliefs are essential, thus fulfilling *demand 3*.

The theory of Schurz and Lambert features a prominent distinction between *data* and *hypotheses*. Data are descriptions of what we have called phenomena, while hypotheses are all the other beliefs that survive the removal of irrelevant elements; these will all be connecting principles. Unification increases with the number of assimilated data, while decreasing with the number of basic hypotheses. (Basic data and assimilated hypotheses have no effect on unification.) This set-up allows the theory to fulfil *demand 1*.

Finally, any argument that is in accord with basic principles of causality gets a unification bonus. Since there is also a demand to maximise unification, this leads to a situation in which effects are generally understood in terms of their causes, and not the other way around. This fulfils *demand 4*.

The fact that Schurz and Lambert's theory of unification appears to have none of the weaknesses that the coherentist theories had, may appear to be a harsh blow against the prospect of using a formal measure of coherence to develop a measure of understanding.

But of course, this would only be the case if the coherence theories did not have strengths of their own. We believe we can identify four of these strengths.

First, the unification theory connects statements to each other using arguments. Coherence theories connect statements to each other through relations of probabilistic dependence. Now arguments, in the sense used by Schurz and Lambert, imply probabilistic relationships; but probabilistic relationships perhaps do not imply arguments. If this is the case, coherence theories are broader. Specifically, if understanding without argumentation is possible, coherence allows us to capture that kind of understanding in a way that the theory of Schurz and Lambert does not.

Second, coherence theories are better equipped to handle explanatory pluralism. For Schurz and Lambert, assimilation is not an all-or-nothing affair, since there are weaker and stronger types of argument that give rise to several grades of assimilation. But once a statement is fully assimilated by a strong argument, no further unification can be gained from it. By contrast, any statement can always continue to grant further coherence to new statements; it is never removed from further consideration. Thus, the unification theory seems to preclude explanatory pluralism. That is, it seems to preclude the idea that one and the same fact can be explained in radically different ways and that having all of these different explanations at our disposal increases our understanding. For an assimilated fact is assimilated, and adding more arguments of which it is the conclusion does not increase unification (though it may decrease it). But adding more connecting principles *can* increase coherence, even if a phenomenon is already strongly connected to other phenomena. If explanatory pluralism is true, coherence theories would be in a better position to capture it.

Third, Schurz and Lambert's theory states, presumably correctly, that we cannot *explain* the most fundamental of our hypotheses. But it is less clear that we do not *understand* these hypotheses. A coherence theory can make room for the plausible prospects of gaining some understanding of our most fundamental theories by seeing how they hang together with each other.

Finally, Schurz and Lambert offer only a vague and partial ordering of sets rather that a precise quantitative measure. In contrast, the coherence measures we have investigated are very precise, and assign a real number to every set of propositions. This increased precision is obviously a boon.

Given these competing sets of virtues, it is of course natural to wonder whether it would be possible to have the best of both worlds. For a coherence theory in particular, it would be ideal if the best elements of Schurz and Lambert could be lifted from their theory and incorporated into what remains fundamentally a coherentist measure. And perhaps this is possible for some parts of the theory. Although detailed and probably rather technical work remains to be done, it seems not unlikely that the distinction between phenomena and hypotheses could be used to remove as spurious all coherence that doesn't increase coherence among the phenomena. There might also be a prospect for integrating the unification bonus that Schurz and Lambert give to causation into a probabilistic measure of coherence by either developing a satisfactory probabilistic theory of causation (see Hitchcock 2012 for an overview) or developing the Bayesian work on Inference to the Best Explanation (e.g., Schupbach and Sprenger 2011) to extend to the asymmetry of explanation.

But the main difference between the approach of Schurz and Lambert and that of coherence theories is the difference between using an asymmetric relation of assimilation and using a potentially symmetric relation of probabilistic coherence. Since it is precisely the asymmetry of their relation that allows Schurz and Lambert to avoid the most difficult problems of asymmetry, one cannot simply replace their relation with one of the probabilistic measures of unification available in the literature (e.g., Myrvold 2003), which do not have this feature. And since a relation cannot be both asymmetric and not asymmetric, it is also hard to see how the relations could be 'combined'. But more on combination in the next section.

7. Conclusions

Can formal coherence be a measure of understanding? Our investigations have not been entirely promising. At the very least, such a theory faces a set of difficult problems when confronted with the asymmetry of understanding. On the other hand, in comparing coherence theories with Schurz and Lambert's theory of unification, we have been able to identify not just the weaknesses, but also the strengths of a coherence approach.

We suggested in the previous paragraph that the prospects for a single theory that unifies the strengths of both approaches may not be great – though it should be stressed that we have not given a strict impossibility proof. However, an easier way of achieving the goal of a satisfying measure of understanding may be available. Schurz (2014) argues that explanation is a *prototype* concept: that is, a concept defined by a number of prototype properties that normally, but not invariably, go together. If that is true, the same will presumably hold for understanding. And that means that unification and coherence can perhaps be combined into one single fuller picture of understanding; not as one theory, but as two theories, both of which capture a part of what we mean by understanding. Unification would allow us to understand the asymmetry of understanding, while coherence would help us do justice to explanatory pluralism and understanding that does not take the form of arguments.

If possible, such a reconciliation would be a happy end to this story. But working out the details of such a joint prototype theory, and thus showing that it is possible to combine these two elements into a single quantitative measure of understanding, remains to be done. In the current article, we hope to have at least been able to give some of the framework within which this work can take place.

Acknowlegdements

The author wants to thank two anonymous referees for very helpful comments on an earlier version of the paper.

BIBLIOGRAPHY

Blanshard, Brand. 1939. *The Nature of Thought.* London: George Allen and Unwin.

BonJour, Lawrence. 1985. *The Structure of Empirical Knowledge.* Cambridge, Mass.: Harvard University Press.

Bovens, Luc and Stephan Hartmann. 2003. *Bayesian Epistemology*, Oxford: Clarendon Press.

Davidson, Donald. 1986. A Coherence Theory of Truth and Knowledge. In *Truth And Interpretation, Perspectives on the Philosophy of Donald Davidson*, ed. Ernest LePore, 307-19. Oxford: Basil Blackwell.

DePaul, Michael R. and Stephen Grimm. 2007. Review Essay on Jonathan Kvanvig's The Value of knowledge and the pursuit of understanding. *Philosophy and Phenomenological Research* 74: 498-514.
Douven, Igor, and Wouter Meijs. 2007. Measuring Coherence. *Synthese* 156: 405–425.
Elgin, Catherine. 2007. Understanding and the Facts. *Philosophical Studies* 132: 33-42.
Elgin, Catherine. 2012. Understanding's Tethers. In *Epistemology: Contexts, Values, Disagreements*, eds. Christoph Jäger and Winfried Löffler, 131-146. Frankfurt: Ontos Verlag.
Fitelson, Branden. 2003. A Probabilistic Measure of Coherence. *Analysis* 63: 194–199.
Fraassen, Bas van. 1980. *The Scientific Image.* Oxford: Clarendon Press.
Gijsbers, Victor. 2013. Understanding, explanation, and unification. *Studies in History and Philosophy of Science Part A* 44: 516-522.
Gijsbers, Victor. 2014. Unification as a Measure of Natural Classification. *Theoria* 79: 71-82.
Hitchcock, Christopher. 2012. Probabilistic Causation. In *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition), ed. Edward N. Zalta.
Khalifa, Kareem. 2013. Is understanding explanatory or objectual?. *Synthese* 190: 1153-1171.
Kosso, Peter. 2002. The Omniscienter: Beauty and Scientific Understanding. *International Studies in the Philosophy of Science* 16: 39-48.
Kvanvig, Jonathan. 2003. *The Value of Knowledge and the Pursuit of Understanding.* Cambridge: Cambridge University Press.
Kvanvig, Jonathan. 2009. The value of understanding. In *Epistemic value*, eds. Adrian Haddock, Alan Millar and Duncan Pritchard, 95–111. Oxford: Oxford University Press.
Lipton, Peter. 2009. Understanding Without Explanation. In *Scientific Understanding: Philosophical Perspectives*, eds. Henk de Regt, Sabina Leonelli and Kai Eigner, 43-63. Pittsburgh: University of Pittsburgh Press.
Myrvold, Wayne. 2003. A Bayesian Account of the Virtue of Unification. *Philosophy of Science* 70: 399-423.
Olsson, Erik J. 2002. What is the Problem of Coherence and Truth? *The Journal of Philosophy* 99: 246–272.
Regt, Henk de and Dennis Dieks. 2005. A Contextual Approach to Scientific Understanding. *Synthese* 144: 137-170.
Regt, Henk de, Sabina Leonelli and Kai Eigner. 2009. *Scientific Understanding: Philosophical Perspectives*, Pittsburgh: University of Pittsburgh Press.
Regt, Henk de. 2013. Understanding and explanation: Living apart together? *Studies in History and Philosophy of Science Part A* 44: 505-509.
Riggs, Wayne. 2003. Understanding 'virtue' and the virtue of understanding. In *Intellectual virtue: Perspectives from ethics and epistemology*, eds. Michael DePaul and Linda Zagzebski. Oxford: Oxford University Press.
Schupbach, Jonah and Jan Sprenger. 2011. The Logic of Explanatory Power. *Philosophy of Science* 78: 105-127.
Schurz, Gerhard. 1999. Explanation as Unification. *Synthese* 120: 95-114.
Schurz, Gerhard. 2014. Unification and Explanation: Explanation as a Prototype Concept. A Reply to Weber and van Dyck, Gijsbers, and de Regt. *Theoria* 79: 57-70.
Schurz, Gerhard and Karel Lambert. 1994. Outline of a Theory of Scientific Understanding. *Synthese* 101: 65-120.
Sellars, Wilfrid. 1956/1997. *Empiricism and the Philosophy of Mind.* Cambridge: Harvard University Press.
Shogenji, Tomoji. 1999. Is Coherence Truth-conducive? *Analysis* 59: 338–345.

Victor Gijsbers is assistant professor of philosophy at Leiden University. His research focuses on general theories of explanation; the concepts of unification and understanding; and the relation between causation and agency.

Address: Universiteit Leiden, Instituut voor Wijsbegeerte, Postbus 9515, 2300 RA Leiden, Netherlands. E-mail: v.gijsbers@hum.leidenuniv.nl