THREE KINDS OF COHERENTISM

Jaap Hage[*]

jaap.hage@maastrichtuniversity.nl

Abstract

This paper aims to show what makes coherentism attractive in comparison to its main competitor, foundationalism. It also aims to show that constraint satisfaction is not the most attractive way to give content to the notion of coherence. In order to achieve these purposes, the paper distinguishes between epistemic, constructive and integrated coherentism. Epistemic coherentism treats coherence as a test for knowledge about a world which exists independently (ontological realism). Constructive coherentism uses coherence as a standard to determine what the facts are in a particular domain. This is a form of ontological idealism. Usually, both epistemic and constructive coherentism apply the coherence test to only part of the positions (beliefs etc.) which a person accepts. The definition of and standards for coherence, just as usually logic and standards for belief revision are kept outside the process of making a position set coherent. Integrated coherentism differs by including everything in the coherence creating process. A set of positions is integratedly coherent if and only if it satisfies the standards included in the set of positions itself.
The paper argues that integrated coherentism best fits with the ideas underlying coherentism and that it is incompatible with coherence as constraint satisfaction in a strict sense.

1 INTRODUCTION

Coherentism has its basis in epistemology (Bonjour 1985; Lehrer 1992 and 2000; Thagard and Verbeurgt 1998; Thagard 2000; Kvanvig 2011) and ramifications into ontology (Rescher 1973; Young 2008), but has also become popular in ethical theory (Rawls 1972; Daniels 2011) and in jurisprudence (MacCormick 1978 and 2005; Dworkin 1978 and 1986; Peczenik 2008; Amaya 2007). Quite recently, coherentism has found its way into Artificial Intelligence and Law too (Amaya 2007; Araszkiewicz 2010). It is not at all obvious, however, that all these versions of coherentism amount to one and the same thing. In fact, I will argue in this

paper that it is possible to distinguish at least three kinds of coherentism, which I will call epistemic, constructive and integrated coherentism.

My aims in this paper are manifold. First I will show what makes coherentism attractive in comparison to its main competitor, foundationalism. This is a discussion about the relative values of foundationalism and epistemic coherentism. I will argue that what makes epistemic coherentism the most attractive of these two alternatives is also a reason to see epistemic coherentism as a holistic theory of knowledge. This a major step into the direction of integrated coherentism. I will also argue that it is attractive to see this holistic version of coherentism as a coherence theory of truth, not only of (the justification of) beliefs. This is an important step towards the unification of epistemic and constructivist coherentism.

Building on the foundations laid in the discussion of epistemic foundationalism, I will show how coherentism can also play a role in ethics and law. This is not at all obvious, because the issue whether an ethical theory or a theory about the contents of the law is right because it is coherent is an altogether different issue than whether a belief is justified because it is an element of a coherent belief set. The version of coherentism in the fields of ethics and law would be constructive coherentism, a version of coherentism that differs fundamentally from epistemic coherentism.

After this discussion of constructive coherentism, it is a relatively minor step towards integrated coherentism, in which epistemic and constructivist coherentism are integrated. Integrated coherentism is a truly holistic version of coherentism, in which beliefs, standards, rules, values, and everything else is seen as elements of a 'theory of everything'. Such a theory of everything would be coherent if it satisfies its own standards for a good theory.

For practical purposes, this view of coherence as a characteristic of theories of 'everything' is rather useless. Therefore I will also provide a brief indication of how such a theory might be turned into a tool that helps in improving actual theories.

To conclude, I will briefly argue why recent attempts to model coherence as a kind of constraint satisfaction (Thagard and Verbeurgt 1998; Thagard 2000; Amaya 2007 and 2011; Araszkiewicz 2010) are essentially flawed because they miss out the crucial holistic aspect of coherentism.

2 EPISTEMIC COHERENTISM

Going by our day to day experience, we are naive realists. Through our senses we obtain a 'picture' of the 'outside world' and we assume that the outside world is like we experience it. This assumption must soon be given up when we discover that our senses sometimes deceive us, and that experience might be seen as the beginning of the philosophical discipline called 'epistemology'. Epistemology addresses, amongst others, the questions what knowledge is and how claims to knowledge can be justified. (Chisholm 1989, 1; Lehrer 2000, 1). However, even when we discover that our senses sometimes deceive us, our starting points will be some form of:

1. *ontological realism*, namely the view that the facts are somehow given, independent of our beliefs about them, and of

2. *epistemic foundationalism*, namely the idea that we get our knowledge about these independent facts in the first place through our senses and that the rest of our knowledge is built on top of that.[1]

2.1   WHY EPISTEMIC COHERENTISM?

Both starting points contain grounds for doubts. Epistemic foundationalism assumes that the set of our beliefs is stratified, where 'higher' layers are supported by lower layers, but not the other way round. There must be one bottom layer which contains beliefs that are not supported by other beliefs. Usually these beliefs are taken to be either self-evident or to result from sensory perception.

There is also a relation of justification between the layers, which works in the opposite direction: the beliefs of the lower layers justify the beliefs in the higher layers by allowing the higher level beliefs to be derived from them. Only the beliefs of the bottom layer cannot be justified in this way, and then there are four possibilities (Albert 1980, 11-15):

a. The beliefs of the bottom layer are not justified and the whole pyramid of beliefs erected upon them is not justified either.
b. The beliefs of the bottom layer are assumed to be justified differently, that is not by means of a derivation from other beliefs. One possible way to see this is that beliefs based on sensory perception are assumed to be justified for the reason that they are based on sensory perception. This boils down to a form of dogmatising these beliefs: they do not need any justification in the form of an argument which bases these beliefs on other beliefs.
c. The 'bottom layer' is rejected as being the bottom layer. The beliefs in it, if they are justified, must be justified by inferring them from the beliefs one or more layers below. This line of reasoning leads to an infinite regress.
d. The chain of inference in which beliefs are derived from other beliefs is allowed to be circular or bidirectional. This gives up the idea that justification works only bottom up.

Both the fear for an infinite regress and the resistance against dogmatising the bottom layer of beliefs have caused resistance against epistemic foundationalism. The main alternative has been to accept that the direction of justification between beliefs does not have to be unidirectional. Beliefs can support each other in a holistic way and the justification of a single belief consists in its being element of a justified belief set. This still rather vague alternative for epistemic foundationalism may be called *epistemic coherentism*. It is coherentism in the broad and negative sense of not being foundationalism.

This motivation for epistemic coherentism is mainly negative. Foundationalism suffers from a number of problems - threat of an infinite regress, dogmatisation of observational beliefs - which apparently cannot be overcome well. Epistemic coherentism is then defined negatively as not being foundationalist. This means that support relation between beliefs can go both directions, and that no beliefs are exempt from criticism. But does this step towards coherentism solve all the problems? Not really. I will discuss three remaining problems:

---

[1]   Arguably, this version of foundationalism is a form of empirical foundationalism. Other forms of epistemic foundationalism are also possible, as Araszkiewicz pointed out to me.

1. Epistemic coherentism assumes that the justification of beliefs lies in the relation between these beliefs and other beliefs. However, it is in general unclear how the truth of beliefs is related to their justification.
2. Coherence is taken to be a standard for the epistemic justification of beliefs. This means that it is implicitly assumed that beliefs in a coherent set are more likely to be true than beliefs which do not cohere. But why would the coherence of a set of beliefs make it more likely that the beliefs are true?
3. The justification of beliefs would lie in the participation of beliefs in a coherent set of beliefs. But what does it mean if a set of beliefs is coherent? What does coherence ad to mere consistency?

2.2 TRUTH CONNECTION

Epistemic coherentism as sketched here is a so-called 'doxastic' theory of knowledge. That it is doxastic means that the justification of beliefs solely consists in the relation of the justified beliefs to other *beliefs*.

Doxastic theories of knowledge are a subset of internalist theories, which take it that the justification of beliefs rests solely on the relation of the beliefs to other 'internal' states of the believer, where these internal states include not only beliefs, but for instance also perceptual states (I hear the Beach Boys playing 'Wouldn't it be nice?) and memories (I remember that I was here before). Internalist theories are to be contrasted with externalist theories, according to which the 'outside world' somehow plays a role in the justification of beliefs (Pollock and Cruz 1999, 89).

All internalist theories have a problem concerning the relation between beliefs and the 'outside world'. Epistemically justified beliefs are beliefs of which it is justified to believe that they are true. If truth is taken to be dependent on correspondence with an independent 'outside world', it is not clear how coherence of beliefs, or – more generally – of internal states, would have anything to do with their truth. Suppose that a person has formed a coherent set of beliefs and that miraculously the world changes considerably.[2] This change will presumably not affect the coherence of the belief set. In other words, the belief set remains coherent. However, even if it were approximately true before the change in the world, after the change it will hardly contain a grain of truth. How could the coherence of the beliefs justify the beliefs as being true, then?

This problem is the result of an attempt to combine two philosophical positions which are hardly compatible, namely ontological realism and epistemic internalism. Ontological realism assumes that the facts in the 'world outside' do not depend on our propositional attitudes about them.[3] For example, an ontological realist takes it that the existence of neutrons does not depend on our beliefs or theories about physical reality. Somebody who is an ontological realist about morality takes it that, for instance, murder is wrong, independent of whether this is accepted or believed by human beings, individually or collectively. Somebody who is an ontological realist (Platonist) with regard to mathematics assumes that the number 438786592 can be divided by the number 16892 even if nobody has ever thought about this possibility

---

[2] This example derives from Sosa 1991a.

[3] I adapted this characterisation of realism from Devitt 1991, 13-17. The adaptation consist in replacing the epistemic notions 'belief' and 'think' by the more comprehensive internalist notion 'propositional attitude'. The reason for this move lies in the step which I will argue in section 5.1, namely the replacement of 'belief' by 'position'.

(which seems not unlikely). I have added the last two examples, because they illustrate that one can be an ontological realist with respect to some domains, while being a non-realist with respect to other domains. One can be a mathematical realist, while being a non-realist with regard to morality, or a realist about physical reality, but a non-realist about the law.

The problem with combining ontological realism and epistemic internalism is that the former detaches the world and the facts in it from our beliefs and more generally our propositional attitudes about it, while the latter detaches the justification of our beliefs from the world. As a consequence the justification of beliefs becomes independent of the facts. Where justification of beliefs is taken as being justified in holding the beliefs to be true, this is problematic.

Where the diagnosis of a problem is clear, the possible solutions can often been seen more clearly. The present problem illustrates this point. The problem of combining ontological realism and epistemic internalism can be dissolved by giving up either one of these positions.[4] Here I will assume without much argument[5] that epistemic internalism is correct, which leaves the abandonment of ontological realism as the only available option for solving the problem.

Running ahead of the discussion of integrated coherentism, I will offer here a brief account of how ontological realism can be given up, even with respect to the physical world, without completely giving up the idea that the world exists independent of our beliefs about it. The basic idea is quite simple, namely that ontological realism is a belief like any other belief. A person may believe that the North Sea borders on England and also that this is a fact that does not depend on his beliefs about it. The same person may believe that there is nothing wrong with gay marriage, and also believe that this depends on his own opinion about this subject. This person holds – as far as the example is concerned – two first-order beliefs, one about the relative place of England and the North Sea and the other about the moral standing of gay marriages. He also holds two meta-beliefs, one about beliefs about physical reality and one about moral beliefs. He may even go so far as to assume that the former beliefs are true, while the latter beliefs have no truth value. About the beliefs on which he is a realist, he assumes that he may be wrong, depending on the facts, while about the beliefs about which he is a non-realist he believes that the categories of 'right' and 'wrong' do not really apply.

On this account, 'ontological realism' is not a theory anymore about the relation (or better: the non-relation) between our beliefs and reality. It is a theory about different ways in which we hold our first-order beliefs. About some first-order beliefs we hold realist meta-beliefs, while on others we hold non-realist meta-beliefs. To coin a term by means of which this position can easily be referred to, I propose to call it 'pseudo-realism'.[6] Pseudo-realism then is the view that a belief is taken to be true, if and only if what is believed is taken to be the case in reality.

Pseudo-realism has implications for the justification of our beliefs. If somebody is the only person who believes that the Earth is flat, and this person is a pseudo-realist about this type of first-order belief, then he has reason to question one's belief about the Earth. If everybody

---

4   It is also possible to combine the positions, and to be satisfied that there is no connection between a belief being justified as true, and the actual truth of the belief.
5   The beginning of an argument can be found in Hage 2005, 35-43.
6   The term 'quasi-realism' has already been taken by Blackburn. See Blackburn 1984, 180.

holds a different view about a fact which is taken to be about facts which are the same for everybody, this should be evidence that this person's own belief is incorrect. If somebody believes that the paintings by Degas are more beautiful than those by Mondriaan, and also believes that this is a matter of taste, then the fact that everybody else believes that Mondriaan's paintings are more beautiful is not evidence against his own belief. The others merely have a different taste.

As these examples illustrate, much that is attractive in ontological realism can be maintained if one adopts pseudo-realism. In particular it remains possible to stick to the idea that the truth must be the same for everybody, and to the expectancy that scientific endeavours will in the end converge upon the truth. Moreover, pseudo-realism can be combined with epistemic internalism, including epistemic coherentism. If the meta-belief is justified that a particular first-order belief is true, and if this first-order belief is justified too, this is a very strong justification for the first-order belief, in particular if the justification of both beliefs consists if the fact that they are elements of the same coherent belief set. Moreover, in this way the truth connection is recued in the sense that the justification of a first-order belief hangs together with the justification of the meta-belief that this first-order belief is about the 'world outside'.

2.3 THE RELEVANCE AND NATURE OF COHERENCE

On the combination of ontological realism and epistemic internalism there is no connection between the truth of a belief and its being justified. No matter which standard for coherence is adopted, the fact that a belief is justified because it belongs to a coherent belief set has no relevance at all for the truth of the belief. That explains at least to some extent why the search for 'the' proper criterion for coherence has remained fruitless.

Coherence has been explained in terms of logical and conceptual relations (Peczenik 2008, section 4.1), explanatory relations (Lehrer 2000, 100-103), constraint satisfaction (Thagard 2000, Araszkiewicz 2010 and Amaya 2011), or has been described as a 'seat of the pants feel' (Putnam 1981, 132/3).[7] The search for a criterion for coherence can amongst others be explained by the intuition that although a set of justified beliefs must be logically consistent, mere consistency does not provide a ground for assuming that the beliefs are justified. There must be something 'more' and this 'more' is dubbed coherence. However, from the fact - if it is a fact – that apart from consistency, something more is needed to transform a set of beliefs into a set of justified beliefs, it does not follow that this 'more' is a single thing. Possibly – and in my eyes even probably – there are several criteria which must be met if a set of beliefs is to be justified. Later, I will argue that there is no general answer to the question what these criteria are, but running ahead of that argument I will give some examples which illustrate different ways in which a set of beliefs can be 'coherent'.

One example is that a set of beliefs about physical laws is more coherent if some of the 'lower' level laws can be derived from higher level laws. If, for instance, the laws of the movement of the planets in our solar system can be derived from Newton's law of

---

[7] See also the contribution of Bermejo to the present volume.

gravitation[8], this makes the joint belief in both laws more 'coherent' and therefore more justified.

Another example is that an ethical theory according to which all moral judgements can be based on one (or a few) 'highest' moral principle(s), such as the principle that happiness should be maximized (Bentham 1982, 11/12; MacCormick 2005, 192), is more coherent than a particularist theory according to which every moral judgement stands on its own, without an 'underlying' moral principle (Dancy 2004).[9]

Again another example is that a theory which explains that there is no connection at all between the opening times of the cinema in Lanaken (Belgium) and the timetable of the railroads in Spain is more coherent than a theory which does explain such a connection. If there is reason to assume that two phenomena are disconnected, a theory which explains their connection would not only count as false, but also as incoherent. Falsity translates into incoherence because the reasons why we assume that the theory is false must somehow conflict with the reasons for assuming the connection.

The upshot of these three examples is that there is not one single criterion for coherence which makes that more coherent belief sets are more justified. In fact, there seems to be no single criterion for coherence and it is not obvious at all that a theory which is more coherent according to one such criterion is for that reason more justified.

3   CONSTRUCTIVE COHERENTISM

The problems around epistemic coherentism can to a large extent be explained from the fact that coherentism, which is essentially an internalist epistemic doctrine, is linked with ontological realism. These problems should therefore disappear if ontological realism is given up. That may not be a very attractive approach with regard to the physical world, but it is much more plausible with regard to the normative disciplines. Many people do not assume that there exist normative facts 'out there' which merely need to be discovered by sensory perception, or – already more plausibly – by means of reasoning.

3.1   INTRODUCING CONSTRUCTIVE COHERENTISM

The alternatives for normative realism are a full-blown non-cognitivism – there is nothing to be known – or some form of constructivism. Constructivism is the view that knowledge in a field depends on the actual or possible arguments which have been, respectively can be, given for a potential piece of knowledge. The conclusions of the best arguments are taken to be true, for the reason that they are the conclusions of the best arguments. Constructivism is only possible on the assumption that there exist standards at the hand of which arguments can be compared. Moreover, these standards themselves should not be based on the assumption that arguments which meet them lead to true conclusions, because that would be circular. Classic examples of such standards in ethics are:

---

[8]   Cf.
    http://en.wikipedia.org/wiki/Kepler's_laws_of_planetary_motion#Newton.27s_law_of_gravitation.
    Downloaded on December 27, 2011.
[9]   This example illustrates that the step from being more coherent to being more justified is not an
    obvious one, at least not in Dancy's eyes.

- the utilitarian standard that arguments should point out that a course of action promotes happiness (Bentham 1982, 11/12);
- the Kantian standard that a proposed principle should be shown to be possible as a principle that would guide everybody's behaviour (Kant 1906, 421);
- the Rawlsian standard that a principle would arguably be chosen from behind the veil of ignorance (Rawls 1972, 11-22)
- the coherentist standard that a principle should arguably lead to outcomes that are acceptable in the light of the principle (reflective equilibrium; Rawls 1972, 48-51; Daniels 2011; Sosa 1991b).

With regard to the law Dworkin (1986, 225) has proposed the standard of constructive interpretation, according to which '…propositions of law are true if they figure in or follow from the principles of justice, fairness, and procedural due process that provide the best constructive interpretation of the community's legal practice.' As we will see 'best constructive interpretation' has strong links with coherentism.

In this section I will briefly discuss reflective equilibrium and constructive interpretation as examples of coherentism in a constructivist setting. The section will be concluded with some general observations about constructive coherentism.

3.2   REFLECTIVE EQUILIBRIUM

The idea of reflective equilibrium was introduced in ethical theory by Rawls, first in a paper about method in ethics (Rawls 1951) and then later, more influentially, in his *A Theory of Justice*. This version of reflective equilibrium was to become later known as 'narrow reflective equilibrium', to be contrasted with 'wide reflective equilibrium' (Sosa 1991b; Daniels 2011) that will be discussed later in this section.

The basic idea behind narrow reflective equilibrium is that an ethical theory is justified because it leads to practical consequences which fit with one's well-considered intuitive judgements about these cases. That these judgements are well-considered implies, amongst others, that they are considered in the light of the ethical theory that leads to them. An example may clarify this.

Suppose that somebody holds the general view that the capacities of persons in society are a common asset, to the benefits of which not only the individuals who actually possess these capacities are entitled, but society as a whole. These capacities include diligence and the inclination to work. Suppose also, that in practice this view leads to the conclusion that society should adopt a system of progressive taxes which makes it possible to grant a social security benefit to, amongst others, those who lack an income because they do not feel inclined to work for their money. Suppose, moreover, that intuitively one is not in favour of giving money to those who are too lazy to work. Reflective equilibrium requires that this intuitive judgement is reconsidered in the light of one's general view about capacities being a common asset. The result may be that in the end one's intuitive judgement changes and that one comes to see that those who are too lazy to work are still entitled to the benefits created by those who happen to be diligent and willing to work for their money. It is also possible that even in the light of one's theoretical views one is still repelled by the idea that people who are too indolent to work should be rewarded through the money earned by those who did make an effort. In that case, the demand of reflective equilibrium brings along that one modifies one's general theory to ensure that it does not have the repelling consequences anymore.

More in abstract, the following is at stake. Narrow reflective equilibrium is a kind of coherence which exists between more abstract normative principles and more concrete judgements. The more abstract principles support the more concrete judgements, but also the other way round. The connection between the more abstract and the more concrete level is made by logical rules of inference and factual premises. In our example, the factual premises would imply that the principle about capacities being a common asset implies that people who are not willing to work and therefore lack a income should receive a government-provided income, and that the government should raise progressive taxes in order to make this possible.

The important thing to notice in this connection is that the factual premises are excluded from the mutual adaptation process. In narrow reflective equilibrium it is not possible to adapt the (beliefs about the) facts because they do not cohere with other factual beliefs. This limitation of what can be mutually adapted in order to gain an equilibrium is abandoned in wide reflective equilibrium. Then other beliefs, principles and judgements – 'positions' from now one, for reasons that will be explained in section 5.1 - can be included in the process of mutual adaptation, and this may even go so far as to include the inference rules of logic, both deductive and inductive, into the process of reaching equilibrium (Goodman 1979, 63-65; Quine 1986, 100; Sosa 1991b). Narrow reflective equilibrium is then a variant on wide reflective equilibrium, a variant which keeps part of the positions out of the adaptation process. Moreover, wide reflective equilibrium may, depending on how 'coherence' is defined, very well be a version of coherentism.

3.3   CONSTRUCTIVE COHERENCE

There are at least two fundamentally different ways to look at the law. On the one view, the law is essentially a social phenomenon, which exists, directly or indirectly, through being accepted by the relevant members of a community. A typical example of this view is the Hartian concept of law (Hart 1994). According to this view, the law consists of a set of rules which derive their status as legal rules from being identified as such at the hand of, in the last instance, a rule that is accepted by the officials of a legal system. Knowledge of the content of the law is on this view knowledge of a part of social reality, and this kind of knowledge would be covered by an epistemic theory such as epistemic foundationalism or epistemic coherentism.

On the other view, the law is the answer to some form of normative question, for instance the question which norms should be enforced by collective means (Hage 2011). The answer to this question is a normative judgement, and it is very well possible to take a constructivist approach to obtaining this answer. One such constructivist approach is the one taken by Dworkin (1986).

The picture that Dworkin sketches of legal justification (in hard cases) is not completely the same as, but nevertheless quite similar to the above picture of narrow reflective equilibrium. The basic idea is that a set of legal materials, cases and legislation, is used to induce legal principles from. These principles are both brought in harmony (coherence) with the legal materials and a normative theory about law and its relation to politics and ethics. The resulting principles are then used to formulate 'new' rules (which may very well be the old ones) and to solve cases with them. Just like narrow reflective equilibrium, Dworkin's constructive interpretation keeps factual information outside the process of mutual

adaptation. Moreover, just like narrow reflective equilibrium may be broadened to make it include other kinds of positions, Dworkin's constructive interpretation may be broadened to make it include other positions. Arguably, if they are both sufficiently broadened, they become both versions of constructive coherentism and they do not even differ from each other anymore.

3.4 Conclusions on constructive coherentism

A major problem of epistemic coherentism is that the connection between the coherence of a belief set and the truth of the beliefs that are coherently held is unclear. The cause of this problem is the combination of epistemic internalism and ontological realism. It can be overcome by adopting pseudo-realism, but that means effectively giving up ontological realism.

On constructive approaches to a domain, such as ethics or the law, ontological realism has already been abandoned. According to constructivism a domain with respect to which one is a constructivist does not exist independent of arguments about the contents of this domain. To state it less abstractly: what is morally right is not discovered by means of moral arguments, but constructed by means of these arguments (Hage forthcoming). If a domain about which one takes a constructivist approach is called a 'constructivist domain', it holds that in a constructivist domain, the conclusions of the best arguments are by definition true.

Constructive coherentism is then 'merely' a particular form of constructivism, namely the view that the best argument about a position consists in showing that this position is an element of a coherent set of positions. Because in a constructivist domain the conclusion of the best argument is true, there is no gap anymore between what is coherent and what is true.

Another thing that has emerged from the above discussion of constructive coherentism is that coherentism cannot only be applied to a set of beliefs about reality, but also to values and norms, including norms of logic which govern the issue what can be derived from what else. In particular the step from narrow to wide reflective equilibrium has lifted the coherentist enterprise to a new level, that of integrated coherentism.

4 Narrow and integrated coherentism

Although coherentism can also be a theory of truth, I will start my exposition of integrated coherentism as a theory of justification. The basic idea is then that a person is definitely justified in accepting a position if and only if this position is an element of the position set held by this person and if this position set is integratedly coherent. This characterisation is full of technicalities which will be explained briefly in this section. I will start with the distinction between narrow and integrated coherentism, because this lies the closest to the discussions of the previous sections.

Narrow reflective equilibrium is characterised by the fact that the sought after equilibrium is established between a limited set of positions. In Rawls' case these positions are normative: abstract principles and concrete well-considered judgements; in a physical science they may be hypotheses and observation statements; in logic they may be logical laws and concrete arguments. In all these cases, the coherentist test is applied only to a subdomain. What coherence amounts to, how it is established, why it is worthwhile, how coherence relates to truth, and, last but not least, all 'background knowledge' which is required to establish

coherence or the lack thereof is somehow taken for granted. That is to say: these factors not taken along in making the evaluated set coherent.

The total set of positions that are held is divided into two parts:

- one part – the 'coherence set' - that one attemps to make coherent and
- another part, the 'background set'- which fulfils several functions, including:
    - the provision of background knowledge which is necessary to establish coherence;
    - defining coherence and providing standards for coherence;
    - indicating what the relation is between coherence and truth;
    - making clear why coherence is worthwhile;
    - indicating how an incoherent set should be modified in order to make it (more) coherent (belief revision).

The step from narrow to wide reflective equilibrium has made it clear that the elements of the second set might just as well be subjected to the test of coherence. It depends on the reason why one wants to have a set of coherent positions which positions are included in the coherence set and which positions are assigned to the background set. It would, for instance, be possible to use values – e.g. the value of coherence – as the background against which beliefs about physical reality are tested, but it would just as well be possible to use beliefs about physical reality as the background knowledge for evaluating a set of values on its coherence.

This raises the question whether the total set of positions should be divided into two parts. Why not join the coherence set and the background set into one all-encompassing coherence set? Then this positions in this set should be used themselves to evaluate the set on 'coherence' or whatever other quality would be considered relevant in the light of this all-encompassing set as a whole. If this step has been taken, wide reflective equilibrium has changed into integrated coherentism.

The word 'coherentism' in this connection does not stand for any particular standard according to which positions might cohere, and not even for the idea of coherence anymore. It only indicates that that no set of positions is granted a privileged position in the sense that it is considered to be 'foundational' or completely immune for revision. *Coherentism is no more than the negation of traditional foundationalism.*

A less traditional form of foundationalism might still emerge in the context of a coherent theory, namely if it fits in a coherent theory that some positions should be considered as foundational etc. The difference with 'traditional foundationalism' is that 'coherentist foundationalism' is the conclusion of an coherentist style of thinking, while 'traditional foundationalism' presupposes the foundational nature of some positions.

The word 'integrated' is meant to indicate that the standards for a 'good' position set, whether they be traditional coherentist (whatever that may be) or 'coherentist foundationalist' are integrated in the position set that should be integratedly coherent. The position set itself contains the standards which it must satisfy.

Coherentist approaches can now be divided into integrated coherentism and narrow coherentist approaches. The former is characterised by the fact that it works with only one

position set, which must satisfy the standards set by itself. The latter are characterised by the fact that they distinguish between a coherence set, which must be made coherent (or which can only justify the positions included in it if it is coherent), and a background set, which contains all the rest.

Depending on how the positions are divided over the two sets, several variants of narrow coherentism can be distinguished, such as epistemic and constructive coherentism. It is also possible to make distinctions within narrow coherentism by means of the standard at the hand of which coherentism is characterised. One such standard might be explanatory coherence (Lehrer 2000, 97-122); another one is coherence as constraint satisfaction (Thagard 2000).

As these two ways of distinguishing within narrow coherentism illustrate, it is the background set that determines which kind of narrow coherentism is at stake. Because integrated coherentism does not separate the background set from the coherence set, it does not allow to distinguish between different variants of integrated coherentism. It all depends on the contents of the coherence set, and this content only allows to distinguish between different coherent sets, but not between different kinds of coherentism.

5 JUSTIFICATION

Integrated coherentism is a theory about justification. Its plausibility depends amongst others on what one takes justification to be. In this connection it is important to distinguish between what justification is and the standards by means of which justification is measured. In this section I will briefly deal with the nature of justification, without saying much about the standards that should be used for justification.

5.1 POSITIONS

At first sight there are many things that can be justified, such as acts, decisions, policies, rules, beliefs and states of affairs. On closer inspection, everything that can be justified turns out to depend somehow on decision making. For instance, acts can be justified to the extent that they are potentially the outcome of decision making (intentional acts); policies and rules can be adopted and abandoned, respectively abrogated and all of these are the outcomes of decisions. The same counts for beliefs, which can also be adopted and abandoned deliberately. And, finally, states of affairs can be justified to the extent that they are the outcome of decision making, or can be changed intentionally.

The view of justification that I will present here as a presupposition of what follows does not deal with all objects of justification, but is broader than merely a theory about the justification of beliefs. Its topic is the justification of 'positions' in general and it treats a belief as one kind of position. I will use the term 'position' as a catch-all for everything, with the exception of behaviour[10] that is amenable to justification. A position is something that is actually accepted; 'things' that are amenable to position are called 'potential positions'. Potential positions include:

- beliefs ('London is the capital of the United Kingdom'),

---

[10] Legal decisions (e.g. convict the suspect) can both be seen as behaviour, in which case it is not amenable to acceptance and as a judgement about what should be done (the suspect should be convicted), in which case it is a potential position.

- practical judgments ('I should review this paper tomorrow'),
- plans ('I will take the plane to Bologna next Saturday').
- rules ('One ought to drive on the right hand side of the road'),
- values ('Truth is valuable),
- logical standards ('If P $\rightarrow$ Q and P are both true, Q must be true'), or
- guidelines for belief revision ('If two positions are incompatible, the one that was more recently acquired should be abandoned').

At first impression one might think that these different objects of justification require different forms of justification, but this impression is only correct in the sense that different standards for justification apply to them. All forms of justification can be reduced to variants on justification of behaviour (including forbearance). This is obvious for actions, and since decisions and verdicts can be brought under the category of actions (*taking* a decision, or *giving* a verdict with this particular content), it should be obvious for decisions and verdicts too. The same counts for using rules.

It is somewhat less obvious for beliefs, but the justification of a belief with a particular content can be interpreted as the justification of *accepting* this belief content. Accepting something can, for justificatory purposes, be treated as a kind of mental action. And just as it is possible to accept belief contents, it is possible to accept goals, values and principles.

It is even possible to continue along this line, by treating the justification of the different forms of actions as the justification of accepting 'that these actions are the ones that should be performed under the circumstances' (Sartor 2005, chapter 3). In this way, all forms of justification can be treated as the justification of accepting 'something'.

As a catch-all term for things that can be mentally accepted, I will use the already introduced term 'position'. I will use the expression 'position set' for the set of all positions accepted by a person.[11] Clearly, the content of a position set may be different from person to person and may change in the course of time.

5.2   LOCAL AND GLOBAL JUSTIFICATION

In the literature on legal justification (e.g. Alexy 1978, 273-278 and MacCormick 1978, chapter 2), justification has sometimes been pictured as a deductively valid argument. In such an argument the conclusion (what is justified) must be true given the truth of the premises. The idea behind this kind of justification is that the 'justifiedness' of the premises is transferred to the conclusion, analogous to the way in which the truth of the premises is transferred to the conclusion in more traditionally conceived deductive arguments.

It seems to me that this picture is mistaken in at least two ways. First, because it suggests that 'being justified' is a characteristic of positions that is similar to truth, only somewhat 'weaker'.[12] Second, because it overlooks the essentially global nature of justification. In a

---

[11]   An interesting line of thinking would be to explore what the implications are if one adopts that position sets can also be held by social groups, or for instance by 'physical science' or 'legal doctrine'. I owe this suggestion to Gustavo Arosenema. Popper's work on 'world 3' (Popper 1972, chapter 3 and 4) goes in the same direction.

[12]   A possible explanation of this use of the term 'justified' is reluctance against the use of 'true' for sentences for the topic of which one uses a constructivist approach. Truth would then be confined to domains for which one adopts realism, and 'justified' for constructivist domains. This usage has

deductively valid argument, the conclusion must be true if the premises are true. This means that the truth of the conclusion is guaranteed by the truth of the premises, and that nothing else is relevant for this truth.[13] For instance, if the statements 'All thieves are punishable' and 'John is a thief' are both true, the statement 'John is punishable' must be true too. Whatever else may be the case, this cannot influence the truth of the statement 'John is punishable', unless it has the implication that one of the premises is false after all.[14] For instance, the fact that John is only five years old either has no impact on John's punishability, because the statement 'All thieves are punishable' is still considered to be true, or – which is more plausible – it has impact, because it makes the statement 'All thieves are punishable' false.

The general point here is that the truth of the conclusion of a sound justificatory argument (deductively valid with true premises) is not influenced by additional information. If the premises of a deductive justificatory argument are true, its conclusion must be true, and then probably also justified, *whatever else is the case*. Therefore, it is not necessary to consider additional information, because this information cannot change the conclusion anymore. Deductive justification is *local* in the sense that it needs only consider the premises and the conclusion of a deductively valid argument.

The reader who thinks this is unrealistic is probably right. Not because deductively valid justification is not local, but because deductively valid justificatory arguments are seldom offered. They are seldom offered, because they require premises the truth of which cannot be established. Take our example about John, who is a thief and punishable because of the rule that thieves are punishable. A deductively valid argument that leads to the conclusion that John is punishable requires the premise that *all* thieves are punishable. That premise can only be established if it is already known that John is punishable (that would be a necessary condition), or if there is a rule that makes all thieves punishable, without exceptions. The former demand would beg the question, because we need the premise about all thieves to justify a belief about John. The second demand is unrealistic, because rules tend to have exceptions, and these exceptions cannot be enumerated.

Real life justification is normally based on premises that support a conclusion without guaranteeing its truth. If John is a thief, this is a reason to believe that he is punishable, but there may be other reasons which invalidate this conclusion. One such a reason would be that John is only five years old. In general a position is justified if the reasons pleading for acceptance outweigh the reasons against accepting it to a sufficient degree. But this means that *all* reasons for or against acceptance must be balanced and that means in turn that justification must be global. A position which is justified in the light of a particular set of other positions need not be justified in the light of an even larger set of other positions, because this larger set may contain additional reasons against adopting it.[15]

---

the disadvantage that it gives the word 'justified' a double role. namely to stand for a person being justified in doing something, and as a counterpart for 'true'.

[13] This should be read as 'irrelevant from an argumentation-technical point of view'. The only thing that is really relevant for the truth of a conclusion is whether this conclusion corresponds with the facts. But that has nothing to do with the argument from which the conclusion follows.

[14] This may be interpreted as a reason why justification on the deductive account of it is global too. But then the global nature does not lie in the deductively valid argument itself, but in the justification of the premises.

[15] In fact, it is even more complicated, because apart from reasons against a position, there may be reasons why the reasons for adoption are in general not reasons after all, or are excluded in this

In logic there is a technical term for a similar phenomenon: *nonmonotonicity*. A logic is nonmonotonic if a conclusion that follows from a set of premises does not have to follow from a wider set of premises (Ginsberg 1987, 1-2). Analogously we can say that justification is nonmonotonic because a position that is justified in the light of a set of other positions, does not have to be justified in the light of a still wider set of positions (Hage 2005, chapter 1).

Nonmonotonicity and the global nature of justification go hand in hand. The 'normal' justification of a position is always relative to a position set of limited size. To get rid of this relativity, one needs to idealise and to assume that it is possible to consider a position in the light of *all* other positions which are actually held. A person would be *definitely justified* in holding a position if and only if he would be justified in holding this position in the light of all (other) positions he holds. In coherentist terminology this would mean that a person is definitely justified in holding a position if and only if the position is an element of an integratedly coherent position set held by this person. Obviously, definite justification and integratedly coherent position sets are unrealistic notions for practical purposes, but they are useful as tools for thinking about the nature of justification.

6   POSITIVE AND NEGATIVE COHERENCE

6.1   COMPREHENSIVENESS

The idea behind integrated coherentism is that the standards for evaluating the coherence of a position set are included in the position set itself. A corollary of this starting point is that a position set is taken to be *comprehensive*: all positions a person accepts are assumed to be part of the position set that is judged on its coherence. This means that not only first-order beliefs about the external world would be included but also meta-beliefs, such as the belief that another belief is, or is not, about an independently existing external world, standards for the validity of arguments, guidelines for belief revision in case a position set is (found to be) incoherent, ontological and epistemological views, religious beliefs, values, (other) norms etc.

According to integrated coherentism, such a comprehensive position set should satisfy the standards set by itself. This means that it includes every position that should be accepted in the light of the rest of its content (the counterpart of logical closure), and does not include any position that should be rejected according to the rest of its own content (the counterpart of logical consistency). Notice that the notion of consistency is not used in this characterisation, because both the relevance and the nature of consistency are to be determined by the position set itself.

Notice also that, for the same reason, logical notions are not mentioned in the characterisation. It is likely that such logical notions will play a role in position sets that are actually held by human beings, but that is not something which is to be included in the characterisation of coherence. That it is not included has everything to do with the fact that integrated coherentism does not single out some positions for being part of the background set in which coherence is defined.

---

special case. More details on the 'logic of reasons' in Hage 1997, chapter IV, and in Hage 2005, chapter 3.

6.2   SPONTANEOUS POSITIONS

In section 2.2 the objection against epistemic coherence theories was discussed that a coherent position set may be isolated from reality. A set of positions may be coherent while all positions contained in it are false. There are three ways to tackle this objection. One way is to argue that it rests on a misunderstanding of justification. The second focuses on how actual position sets develop in time, while the third focuses on the role of meta-beliefs.

I will be brief about the first refutation of the objection. This holds that the objection confuses truth and justification. That a position is untrue is no objection against a person being justified in accepting this position. For instance, a person may be justified in believing that it is presently raining in a faraway country, because a normally reliable internet site tells her so. That the internet site had a software problem with as a consequence that it reported falsely about the actual weather does not change her being justified in her belief. This would be different if a person knows, or should have known, that a belief is untrue. But then the problem is not the falsity of the belief, but the acceptance of the belief that an accepted position is false, which amounts to inconsistency of the position set.

A second refutation of the criticism that coherence does not really justify acceptance of a position is based on a consideration how a coherency test of justification will operate in practice. One does not come up with a coherent set of positions from scratch. Normally one starts from an already existing set, a 'base' (Raz 1992) . The contents of a *real position set*, that is a set that is entertained by some real person, will have two kinds of determinants, rational ones and a-rational ones. The rational determinants make that an existing set should be corrected – new positions should be added and existing ones should be removed – because rationality requires this given the rest of the set.[16] The a-rational determinants cause 'spontaneous' changes to the contents of the set. New positions are added as the effect – notice the causal terminology - of perception, memory, intuition, or any other factors which cause what a person accepts. A person may, for instance, accept something because he mistakenly believes that this is rational in the light of what else he believes. Existing positions are removed, because they are forgotten or abandoned for irrational or a-rational reasons.

These irrational or a-rational influences on a position set are relevant because position sets are biased toward the past. Whether a new position should, from the rational perspective, be added, or an existing one removed, depends on the present contents of the set. To see why, one should notice that a particular position can have one of three statuses in the light of (the rest of) a position set:

1. it should be adopted (if not already present) because this is rational; the position is *acceptable*;
2. it should be removed (if it is already present) because this is rational; the position is *rejectable*;
3. it is neutral in the sense that it should neither be adopted nor removed; the position is *suspendable*.

---

[16]   Notice that the contents of a position set determine what *should be* added, removed or changed. The position set itself does not lead to the required modifications. Moreover, the definition of integrated coherence does not state anything about the way an integratedly coherent set should come about. it only provides a test on a position set to determine whether it is coherent.

As long as a position set is not coherent – that means in practice: always - the judgment whether a particular position should be added or removed should rationally be made on the basis of the full position set, unless there are reasons to assume that some elements are rejectable, because then these elements should not be taken into account. (This runs already ahead of the argument in section 7.2 about default reasoning.)

Suspendable positions remain in the set, even if there is no reason to adopt them, and codetermine what is acceptable. So suspendable elements can be justified because they belong to a coherent position set, without being justified by particular reasons in the set. This may, for instance, hold for beliefs caused by perception, or for evaluative judgments caused by 'intuition'. As long as there is no reason to reject them, they will be maintained. Moreover, suspendable elements can also play a role in determining whether other elements can remain in the set and are therefore justified.

Since rational modifications can only take place on the basis of an existing position set, the original elements of any realistic position set must have been spontaneous positions, that is neutral elements. This is one part of the argument why real position sets do not 'hang in the air'. Real sets stem from sets of spontaneous positions, and there is where we should look for the desired 'contact with reality'.

With the view that suspendable positions that are part of a position set should remain there, a position is adopted that is known in the epistemological literature as *negative coherentism* (Pollock and Cruz 1999, 70-71). The demand for coherence is used to weed out positions that do not fit. It is not required that every element in the position set is supported by other elements. If the latter were required, the coherentist view would be positive. The advantage of negative coherentism over positive coherentism is that it can account for positions that are based on perception, (moral) intuition, and memories. In particular the possibility to account for perceptive input makes that the link between the positions in a coherent position set and reality, that is brought about by perception, *can* be maintained. However, whether it *will* be maintained depends on the other position in the set, including beliefs about the relation between the 'external world' and perceptive states. Integrated coherentism allows for an important role for perceptive states, but does not require it.

7 META-BELIEFS

The role of meta-beliefs as a possible way to refute the 'isolation argument'(Pollock and Cruz 1999, 74-75) was already mentioned briefly in section 2.2 in the discussion of pseudo-realism. Since this refutation plays an important role in integrated coherentism, that discussion will be elaborated a little here.

7.1 PSEUDO-REALISM

The central question in connection with the 'isolation argument' is how a coherent position set relates to the 'world outside'. To begin with, we believe that there is such a world, and we also believe that through our senses this world influences our spontaneous positions. It should be noted, however, that these beliefs, like all of our other beliefs, belong to our position set, and can be corrected on the basis of the rest of the set. The existence of an external world is like the existence of so-called 'theoretical entities'. Entities like electrons cannot be perceived, but their existence is derived from other things that we can perceive. With the external world it is a little different: we believe that we can perceive it, but its

existence is still theoretical in the sense that we postulate its existence on the basis of perceptual impressions. The world is, so to speak, necessary to explain our perceptions of it.

This should not be read as stating that we believe the external world to exist only because that would explain our beliefs about it. Our cognitive apparatus is such that we spontaneously believe that (most of) our perceptive impressions are impressions of the external world. E.g. we do not only spontaneously believe that the sun is shining, but also that the sun is shining in the 'outside' world. In fact, the latter assumption is so natural that we automatically take the first belief to be identical to the second. However, we are able to question this spontaneous belief, and if it is questioned a reason for adopting it (again) is that the external world explains both the existence of our spontaneous beliefs about it, and the convergence of (some of) our beliefs about it with the corresponding beliefs of other persons.

In this connection, agreement plays an important role. If different persons have the same beliefs, this may be taken as a reason to assume that there exists an independent object of belief that causes the unanimity of the beliefs about this object. For instance, the 'objective' existence of a table explains why we all see the table and believe that it is there. Similarly, the existence of some moral disagreements and their resistance against factual information fuels the meta-belief that there is no 'objective' moral reality analogous to the physical reality.

Agreement in beliefs needs not be explained by an objective external world, however. A second type of explanation would be that agreement on a position is the outflow of the functioning of our cognitive apparatus. Think in this connection of mathematics. Mathematicians tend to agree on many results of their science, but only some of them attribute this agreement to an objectively existing world of mathematical objects, such as for instance natural numbers. Another example would be that our moral intuitions are, at least to some extent, innate (cf. Hauser 2008).

A third possible explanation of agreement is that a position is the result of a procedure that is designed in such a way that it leads to the same outcome for (almost) everybody. A legal example would be that legal arguments based on the same rules and cases and using the same canons of interpretation and argumentation lead in easy cases to the same outcomes for almost everybody.[17]

From the fact that agreement on a particular position may be a sign that this position reflects an outside world that is the same for everybody, it does not follow that where agreement is lacking a position does not reflect the outside world. For instance, we believe that the position 'There is water on the moon' reflects the outside world, but there was (at the moment of writing this paper) no agreement yet amongst scientists whether this position is true. However, because we assume that the position reflects the outside world, we tend to believe that an increase in relevant knowledge should, in the end, lead to agreement.[18] Where we do not even expect that an increase in knowledge would lead to agreement, we apparently do not assume that the position reflects the outside world.

---

[17] This may even be so by definition, if easy cases are defined as those cases which lead to agreement amongst those who argue by the rules of law. This position was argued in Hage, Leenes and Lodder 1994.

[18] This comes near to Pierce's circumscription of truth: 'Truth, what can this possibly mean except it be that there is one destined upshot to enquiry with reference to the question in hand.' Quotation taken from Kirkham 1992, 81.

7.2   DEFEASIBLE COHERENTISM

For practical purposes, the insight that somebody is justified in accepting a position if this position belongs to the integratedly coherent position set held by this person, is useless. Actual position sets will never be integratedly coherent, if only because they would most likely be infinitely large.[19] If coherentism is to be used as a real test on the justification of positions, a less demanding version should be devised.

One such less demanding version will be described here very briefly. For ease of exposition I will dub it *defeasible coherentism*. The starting point is again the idea that justification should be based on coherence and that the set which is to be made coherence also contains the standards by means of which it is evaluated. What must be dropped is the demand that this set is comprehensive, because a comprehensive set would be unmanageable. Therefore the coherence of the position which is to be justified is not to be with the full comprehensive set, but merely with a 'suitable' subset thereof. This subset should include the position that is to be justified and it should be integratedly coherent.

The crucial extra demand on the subset is that it is 'suitable'. A suitable subset is a subset which does not give a wrong picture of the full set where the position that is to be justified is concerned. Ideally, the subset should only contain this position if the full set would also contain it. Such a small position set is with respect to this position *adequate* for the comprehensive set. Obviously this adequacy cannot be established without having the disposal over the full set, and that is precisely what will not be available.

The solution to this problem can be found in the use of meta-beliefs and the characteristics of negative coherence. Somebody can hold a belief which, *as far as she can see*, fits in the set of all her positions. This means that she believes that her belief fits in her actual comprehensive belief set. She does not know whether this meta-belief about the adequacy of her subset is true, but she nevertheless has it. Moreover, as long as she does not have counter-evidence, she is defeasibly justified in sticking to it.

Not having counter-evidence means in this connection: not being aware of counter-evidence and not having counter-evidence in the same defeasible sense that the subset of her positions of which she is aware contains reasons not to trust it.

An example may make this clearer. A public prosecutor believes that John is a thief and that the rule exists that thieves are punishable. She also uses a rather traditional logic for rule application, including modus ponens style arguments. Given this limited position set, she should adopt the belief that John is punishable. Actually she also believes that minors are not punishable, but she has no reason to assume that John is a minor. Therefore she actually believes that her small position set is adequate for her actual comprehensive set where the belief that John is punishable is concerned. Moreover, as long as she is not aware of counter-evidence (there is no counter-evidence in her subset) she is justified in sticking to this meta-belief. The subset is defeasibly coherent and therefore, this public prosecutor is defeasibly justified in believing that John is punishable.

---

[19]   An integratedly coherent position set should contain everything which should be accepted in the light of what else is accepted. Probably this involves the logical closure of the position set, and logical closure under deduction means infinitely large.

8   EVALUATION

Amaya (2011) has made a brief list of six 'serious problems' which face coherence theories in law, both the constructivist and the epistemic ones. It may be useful to see how integrated coherentism fares with regard to these problems.

VAGUENESS

Coherence theories often leave the notion of coherence undefined, and that is not very helpful if it comes to the justification of legal judgements. Integrated coherentism might be seen as the so-manieth illustration of this criticism, because it leaves the determination of what counts as coherent to the position set that is to be made coherent itself.

And yet, this criticism would not be appropriate. It would be comparable to criticising the word 'good' for being too vague. 'Good' is not vague at all. It is a very general term of recommendation. That is does not specify the grounds on which the recommendation is made does not make its meaning unclear, but merely leaves the conditions for its application unspecified. Analogously, integrated coherentism is clear in the sense that it requires a position set to meet its own conditions for a good position set, without specifying what these conditions are. Requiring that a theory of coherentism specifies the conditions of coherence would be comparable to requiring a theory about the meaning of 'good' specifies what makes things good.

HOLISM

Coherence theories, at least some of them (Dworkin 1986; Peczenik and Hage 2004), require that legal judgements are justified on the basis of a theory about the whole system of beliefs about law and political morality. A similar requirement exists for epistemic justification. That makes the justification of judgement very difficult and moreover, such a theory is descriptively inadequate.

The accusation of holism applies certainly to integrated coherentism, but in my opinion holism is a strength, rather than a weakness. Admittedly, requiring coherence of a comprehensive position set makes it impossible to apply full-blown integrated coherentism as a tool for justifying legal judgements. But who said that it is a drawback of an analysis of justification that it cannot be applied as a practical tool? This is comparable to criticising a theory about the nature of prime numbers for making it impossible to enumerate all the prime numbers.

An analogous rebuttal applies to the criticism that (integrated) coherentism is descriptively inadequate. Theories about what justification amounts to are not meant to be descriptions of social practices, but abstract analyses of normative standards. Normative, because they specify which standards must be met if a position is to count as justified. Abstract, because the standards are mentioned only in an abstract way. In general, the demand that epistemic theories are descriptively adequate confuses prescription and description.

And finally, as was shown in section 7.2, integrated coherentism can be complemented by a theory of defeasible justification which is not holistic, at least not in a way that hampers practical application of integrated coherentism to the process of justifying legal judgments.

Coherent theories may be circular, for instance because belief in the occurrence of particular phenomena is justified on the basis of a physical law, while acceptance of the law is justified on the basis of these phenomena.

This objection is a straw man. I have never seen a coherence theory which would allow this simple mistake. Moreover, any feasible criterion for coherence would exclude it.

In integrated coherentism the problem is easily solved. Either such circular arguments are - for as yet unclear reasons – deemed acceptable, and then there is no problem if they are allowed. Or – which is much more likely – such arguments would not be allowed, and then the standards contained in an integratedly coherent theory would not allow them.

CONSERVATISM

Coherence theories have a tendency towards conservatism, because they make the justification of new elements depend on their coherence with a pre-existing structure.

Yes, coherence theories make the justification of new elements depend on their coherence with a pre-existing structure, just as they make the continued existence of the pre-existing structure dependent on its coherence with new positions. Mutual dependence makes the new depend on the old, and the other way round. That is exactly as it should be, and it is unclear why this might be mentioned as a problem.

What might have been mentioned is the tendency of conservatism which is attached to negative coherentism. Negative coherentism allows positions to be maintained as long as there are no coherence-based reasons to abandon them. This allows that arbitrary beliefs and moral views to survive, merely because there is as yet no counter-evidence. Integrated coherentism is committed to this negative coherentism, so this 'criticism' would apply to integrated coherentism.

The reason why integrated coherentism is a form of negative coherentism is that this is the best way to allow sensory input and moral 'intuitions' to play a role in the justification of positions. Coherentism is most plausible as a theory about when a position set needs no reconstruction anymore and is in that sense justified. It is not so that any arbitrary set of positions is justified, merely because it satisfies its own standards. It is a person who is justified in holding the position she holds, because her position set has reached a stage in which there is no need any more to modify it because of a lack of coherence.

IT IS UNCLEAR WHAT COHERENCE-BASED INFERENCE IS

There exists no clear account of the inference-patterns that legal decision makers may use while seeking coherence in law.

This is not an objection against (integrated) coherentism as a theory of justification in itself, but rather a complaint that such a theory does not solve all the problems which a legal decision maker may face. (Integrated) coherentism aims to specify what it means that a person is justified in accepting a set of positions. It does not aim to specify how this set is arrived at. In fact, as far as the justification of the positions in the set is concerned. it does not matter at all how the position set was arrived at. Demanding otherwise risks to confuse heuristics and legitimation.

Why should coherence be endorsed with justificatory force in a legal setting, granted that coherence can justify?

For integrated coherentism this question is easy to answer. There is no separate issue of legal justification, but only one general issue of being justified in accepting positions. Some of these position will concern legal issues, but for the purpose of justification they are treated just like other positions. So for integrated coherentism, this question makes little sense.

For narrow coherentism, things may lie differently, because then the issue plays a role which positions are to be included in the coherence set and which ones in the background set. Then it depends on one's theory about the nature of law whether coherentism should play a role in legal justification. Obviously there is much more to be said on this issue, but equally obviously, this is not the place to do so.

In general we can conclude that the problems mentioned by Amaya either are not real problems for coherentism *as a theory about when a position is justified*, or that integrated coherentism does not suffer from the mentioned problem.

9   COHERENCE AS CONSTRAINT SATISFACTION

To give a good impression of integrated coherentism it is useful to contrast it with a modern alternative. In this connection I have selected the work of Amaya again, because it contains a good insight in what is involved in coherentism and applies this insight to the law.[20]

Amaya has developed her own coherence theory for law, a theory based on constraint satisfaction. In this paper I will focus on the constraint satisfaction part of Amaya's views, but honesty commands that she recognises three other (?) elements in het model of legal reasoning, which I will only briefly mention. These elements are:

1. *Inference to the best explanation*. 'The claim is that legal decision-makers reach justified beliefs about the law and facts under dispute by first generating a number of alternative interpretive and factual hypotheses and then by selecting one of them as justified.' A brief comment to this element: it contains a description of what legal decision-makers allegedly do. This can only be part of a theory of justification on the assumption that the actual practice has justificatory force, that is on a form of justificatory naturalism. In fact, Amaya adopts such a form of naturalism in a footnote. (Amaya 2011, footnote 1).
2. *Epistemic responsibility*. The basic idea here is that legal reasoners have to fulfil a number of epistemic duties. The duty explicitly discussed is the duty to collect more evidence on propositions that are less than certain on one's evidence.
3. *Context*. As Amaya points out, context is very relevant to the assessment of coherence. Standards of legal justification vary with context.

I will confine myself to a brief discussion of coherence as constraint satisfaction. Amaya has taken this approach from Thagard and Verbeurgt (Thagard and Verbeurgt 1998 and Thagard 2000). She describes it as follows (Amaya 2007):

---

[20]   As this volume illustrates, Amaya's work does not exhaust the constraints satisfaction approach to legal reasoning. Earlier examples are Joseph and Prakken 2009 and Arasziewicz 2010.

'On this view, coherence maximization is a matter of maximizing the satisfaction of a set of positive and negative constraints among the elements of a given set. The idea is the following one. We start with a set E of elements, which may be propositions or other representations (goals, actions, concepts, etc.). The problem is how we can accept some elements and reject others in a way that maximizes the coherence of E. The claim is that we turn E into as coherent a whole as possible by taking into account the coherence and incoherence relations that hold between pairs of elements of E. These relations put constraints on what can be accepted or rejected. To maximize coherence, we partition E into two disjoint subsets A, which contains the accepted elements, and R, which contains the rejected elements, in a way that takes into account the local coherence and incoherence relations. For example, if a hypothesis h_1 explains e_1, we want to ensure that if h_1 is accepted, so is e_1. On the other hand, if h_1 is inconsistent with h_2, then we will make sure that if h_1 is accepted, then h_2 is rejected. The coherence problem is thus that of dividing up E into A and R in a way that best satisfies the most constraints.'

A simple example may illustrate the point. Suppose that the set E contains the following elements:

a. John is a thief.
b. John is a minor.
c. John is punishable.
d. John is not punishable.

There is a positive constraint between the elements a and c, meaning that some positive value is attached to the fact (if it is a fact) that the subset A contains both a and c. Let us assume this value equals 2. There is a negative constraint between the elements b and c, meaning that some positive value is attached to the fact (if it is a fact) that the subset A contains b and subset R contains c, or the other way round. Let us say that this value is 4. Moreover there is a negative constraint between a and d (value 1) and a positive constraint between b and d (value 2). There is a very strong negative constraint between c and d, because they are inconsistent. (If c and d are both accepted, one has to give up bi-valued logic.) Let us assume that the value of this constraint equals 1000.

Given the value of the last constraint, c and d must be in different subsets. If the one is accepted, the other one must be rejected. If c is accepted, the sum of the values is increased if a is also accepted, and if b is rejected. Id d is accepted the sum is increased if b is accepted and a rejected. The maximum sum (1007) is achieved by accepting b and d and by rejecting a and c.

This example illustrates how coherence can be given a precise criterion (maximisation of the sum of the constraint values), and how coherence can be a matter gradation (the sum can be higher or lower). Moreover, given a finite set of elements and of constraints, it is possible to compute what the most coherent theory is. These are all advantages for the practical use of a coherentist theory.

However, the example also illustrates some drawbacks of the constraint satisfaction approach. One drawback is that this approach treats constraints as bi-directional. If a supports c, then c supports a (to the same degree). On the one hand this is desirable, because on some occasions we argue from the fact that somebody is a thief to that he is punishable, while on other occasions we hypothesize that somebody might have been a

thief, because he is punishable. However, as the wording ('hypothesize', 'might have been') already suggest, the one direction of reasoning is more plausible than the other direction. It is more plausible to conclude that somebody is punishable from that he is a thief, than the other way round. The proposed representation of the constraints does not make it possible to take this into account.

Possibly this drawback can be remedied by introducing more elements and more constraints into the model, but this way to deal with the problem immediately points to a more serious drawback of the approach: it may be necessary to tinker extensively with the constraints and their values in order to obtain the 'right' results. This presupposes that the 'right' results are already given, and that the model is better or worse, depending on how good it is in producing the correct results. Then, however, this approach can only be used in epistemic contexts, where the justification of alleged knowledge is at stake. In the case of constructive coherentism, where the model defines what are the correct elements of a theory, this does not work.

It may be objected that in the constructive context it does not have to work, because any outcome of the model is correct, precisely because it is the outcome of the model. This objection leads us to what is in my opinion the most serious drawback of the constrain satisfaction model, namely that a very substantial part of the information about a domain is stored in the constraints and the values that are attached to them. In a constraint satisfaction model, the coherence of a theory consists in the way in which the ordinary elements of the theory are divided over the accepted-set and the rejected-set. In that way, the information contained in the constraints is not part of the coherent set, although it does play a role in determining what the most coherent set is. The coherence of the set is the coherence of only a part of the available information (Hage 2005, 45-47).

Essentially this criticism amounts to it that coherence as constraint satisfaction is a form of narrow coherentism which places crucial parts of the available information in the background set. In my opinion constraints should be part of the coherence sets rather than of the background sets. Moreover, it would still be better if everything is included in the coherence set, leaving the background set empty. That would amount to adopting integrated rather than narrow coherentism.

10 CONCLUSIONS

My purposes with this paper were:

- to argue why coherentism is more attractive as a theory of justification than its main competitor, foundationalism,
- to show how coherentism can play a role in moral and legal justification, and finally, on a more negative side
- to argue why coherence as constraint satisfaction, which is coming in vogue now, is less attractive.

Much of my argument has consisted of making distinctions and by showing how insight of what is at stake in coherentism 'automatically' leads to a particular variant of it, integrated coherentism. As soon as the attractiveness of integrated coherentism has become clear, the

main drawback of coherence as constraint satisfaction springs to the eye: it is a variant of narrow coherentism, the coherentist opponent of integrated coherentism.

The best way to draw the conclusions of this paper is to summarise and elaborate the distinctions that were made.

1. Major distinctions that mainly functioned on the background of this paper are those between *justification, explanation and description*. It is possible to describe how, for instance, judges go about when arguing for a particular conclusion, but such a description is not relevant for the question whether the judge was justified in adopting this decision. The exception might be that one adheres to a procedural theory of justification. Neither is it necessary that a theory of justification explains why judges take the decisions they take. A justificatory theory may play such a role, but that would require a kind of rational choice theory to explain legal decision making.

2. Possibly the main distinction in this paper is *between integrated and narrow coherentism*. Narrow coherentism strives for coherence in a limited set of positions, while the other positions held by a person are put into a background set which is either taken for granted, or is evaluated on a different basis than coherence. It is this background set that indicates why coherence is important, and what the criteria for coherence are. Integrated coherentism does not distinguish between two position sets. All positions are lumped together and are jointly evaluated on their coherence. The motivation of and the standards for coherence are part of the single position set themselves.

3. Although it might seem that the *distinction between coherentism and foundationalism* is more basic than that between the two main forms of coherentism, that is not the case from the perspective of integrated coherentism. Foundationalism is from that perspective just a variant of narrow coherentism. A set of positions (e.g. sense-based beliefs, and deductive logic) are set apart in a background set, and form a standard by means of which the positions in the other set are evaluated. This standard will in the case of foundationalism not be a coherentist one, but not even integrated coherentism requires that the position in a position set are evaluated on the basis of a coherence standard in a narrow sense.

4. If one adopts a form of narrow coherentism, it becomes easy to distinguish between *epistemic and constructive coherentism*. Epistemic coherentism applies the coherence standard to determine whether beliefs about an independently existing reality are true. This is problematic, because there is little reason to assume that coherence of beliefs is relevant for the truth of these beliefs. This 'isolation problem' is not without reason one of the main problems of coherentism. It is important to notice, however, that it is only problematic on the unlikely combination of epistemic coherentism and ontological realism (about which more soon). Constructive coherentism is the combination of two views. The one, constructivism, is the view that some facts obtain because it is the outcome of the best argument that they obtain. The other one defines the best argument in coherentist terms (e.g. reflective equilibrium, or constructive interpretation). Constructive coherentism presupposes a theory about a domain (e.g. ethics or the law) which assumes that facts in this domain depend on the best argument. Such domains are called 'constructive domains'.

5. One way to deal with the isolation problem is to distinguish between *ontological realism and pseudo-realism*. Ontological realism assumes for some domain that the

objects and facts in it exist independently of what people accept or believe about it. Pseudo-realism takes existence and truth to be an outflow of a position set, and in particular meta-beliefs. A sentence is on this view considered to be true if its content is accepted (the car is red indeed) and if it is also accepted that the 'world outside' makes it true (there are independently obtaining facts entailing that the car is red) . This distinction between ontological realism and pseudo-realism has implications for the relation between the facts and the truth of sentences. Both views can adhere to the correspondence theory of truth; a sentence is true if and only if the fact that it expresses obtains. They give a different reading of the direction of this connection between fact and truth, however. Ontological realism assumes that the facts are already there and that the sentence aims to reflect this fact (word to world direction of fit; Searle 1979, 3-4; Hage 2005, 165). Pseudo-realism, on the contrary, assumes that the fact obtains because the sentence is true (world to word direction of fit). Pseudo-realism sits well together with constructivism, because of the chain: best argument → true sentence → fact.

Integrated coherentism is a very abstract theory about what a person is justified in accepting. Its abstraction has several consequences:

a. Integrated coherentism in its pure form is not useful as a practical tool to justify or modify beliefs. It is useful, however, to gain insight in what is good in coherentism, and what is not, and to see what is required if one wants to adopt some version of narrow coherentism. It is also useful as an indication how a more practical form of coherentism (defeasible coherentism) can be developed.

b. Because of its abstraction, integrated coherentism is compatible with adapted versions of foundationalism, constraint satisfaction, and ontological realism. The required adaptation is usually that the other position is embedded in an integratedly coherent position set which explains why that position is worthwhile. For example, the position may be part of a coherent theory that beliefs that are caused by sensory perception take a privileged place in the total position set and cannot be rejected, or can only be rejected for very strong reasons. If this meta-belief about beliefs based on sensory perception is part of an integratedly coherent position set, the rest of the set functions as an empirical foundationalist set.

c. Because of this possibility to combine integrated coherentism with many other views which at first sight might seem to conflict with it, it is very difficult to criticize integrated coherentism. The main criticism seems to be that integrated coherentism does not exclude much and is therefore empty. My reply to that criticism would be that mathematics is empty in the same way.

Integrated coherentism is valuable if it is taken for what it is, a very abstract theory about when a person is justified in accepting a particular position and as a demonstration of how and in what sense coherence plays a role in this. It must function as a background against which more specific epistemological theories can be developed, and as a test for the feasibility thereof. It should not be taken as an alternative for these more concrete theories.

REFERENCES

Alexy, Robert (1978), *Theorie der juristischen Argumentation. Die Theorie des rationalen Diskurses als Theorie der juristischen Begründung,* Frankfurt: Suhrkamp.

Amaya, Amalia (2007), 'Formal Models of Coherence and Legal Epistemology', in *Artificial Intelligence and Law* 15, 429-447.

Amaya, Amalia (2011), 'Legal Justification by Optimal Coherence', in *Ratio Juris* 24, 304-329.

Araszkiewicz, Michal (2010), 'Balancing of Legal Principles and Constraint Satisfaction', in R.G.F. Winkels (ed.), *Legal Knowledge and Information Systems. (Proceedings of the Jurix 2010 conference)*, Amsterdam: IOS Press, 7-15.

Blackburn, Simon (1984), *Spreading the Word. Groundings in the Philosophy of Language*, Oxford: Clarendon Press.

Bonjour, Laurence (1985), *The Structure of Empirical Knowledge*, Cambridge: Harvard University Press.

Chisholm, Roderick M. (1989), *Theory of Knowledge*, 3rd ed., Englewood Cliffs: Prentice Hall 1989.

Daniels, Norman (2011), 'Reflective Equilibrium', in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Spring 2011 Edition)*, http://plato.stanford.edu/archives/spr2011/entries/reflective-equilibrium

Devitt, Michael (1991), *Realism and Truth*, 2nd ed., Oxford: Blackwell.

Dworkin, Ronald (1986), *Law's Empire*, London: Fontana.

Ginsberg, M.L. (ed.) (1987), *Readings in Nonmonotonic Reasoning*, Los Altos: Morgan Kaufman.

Goodman, Nelson (1979), *Fact, Fiction and Forecast*. 3rd ed., Hassocks: The Harvester Press.

Hage, Jaap C. (1997), *Reasoning with Rules*, Dordrecht: Kluwer.

Hage, Jaap C. (2005), *Studies in Legal Logic*, Dordrecht: Springer 2005.

Hage, Jaap C. (2011), 'The Method of a Truly Normative Legal Science', in M. Van Hoecke (ed.), *Methodologies of Legal Research*, Oxford: Hart Publishing, 19-44.

Hage, Jaap C. (forthcoming), 'Construction or reconstruction? On the function of argumentation in the law', in press.

Hage, J.C., R.E Leenes and A. Lodder (1994), 'Hard cases: a procedural approach', *Artificial Intelligence and Law* 2, 113-167.

Hart, Herbert L.A. (1994), *The Concept of Law*, 2nd ed. Oxford: Clarendon Press.

Hauser, M.D. (2008), *Moral minds: How nature designed our universal sense of right and wrong*. London: Abacus.

Joseph, S. and H. Prakken (2009), 'Coherence-Driven Argumentation to Norm Consensus', *Proceedings of the 12th International Conference on Artificial Intelligence and Law*, New York: ACM, 58-77.

Kant, Immanuel (1906), *Grundlegung zur Metaphysik der Sitten*, Hamburg: Felix Meiner Verlag. First edition 1785. The references use the page numbers of the Akademie Ausgabe.

Kirkham, R.L. (1992), *Theories of Truth*. Cambridge: MIT Press.

Kvanvig, Jonathan (2011), 'Coherentist Theories of Epistemic Justification' in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Summer 2011 Edition)*, http://plato.stanford.edu/archives/sum2011/entries/justep-coherence

Lehrer, Keith (1992), 'Coherentism', lemma in Jonathan Dancy and Ernest Sosa (eds.), *A Companion to Epistemology*, Oxford: Blackwell, 67-70.

Lehrer, Keith (2000), *Theory of Knowledge*, 2nd ed., Boulder: Westview Press.

MacCormick, Neil (1978), *Legal Reasoning and Legal Theory*, Oxford: University Press.

MacCormick, Neil (2005), *Rhetoric and the Rule of Law*, Oxford: University Press.

Peczenik, Aleksander (2008), *On Law and Reason,* 2nd ed., Dordrecht: Springer.

Peczenik, Aleksander and Jaap Hage (2004), 'Legal Knowledge about What?' *Ratio Juris* 13, 326-345.

Pollock, John L. and Joseph Cruz (1999), *Contemporary Theories of Knowledge*, 2nd ed. Lanham: Rowman and Littlefield.

Popper, K.R. (1972), *Objective Knowledge*, Oxford: Clarendon Press.

Putnam, Hilary (1981), *Reason, Truth and History*, Cambridge: Cambridge University Press.

Quine, W.V. (1986), *Philosophy of Logic*. 2nd ed. Cambridge: Harvard University Press.

Rawls, J. (1951), 'Outline of a Decision Procedure for Ethics', *Philosophical Review* 64, 177-197. Also in J. Rawls, *Collected Papers*, edited by Samuel Freeman, Cambridge: Harvard University Press 1999, 1-19.

Rawls, J., (1972), *A Theory of Justice*, Oxford: Oxford University Press.

Raz, Joseph (1992), 'The Relevance of Coherence', *Boston University Law Review* 72. Also in Joseph Raz, *Ethics in the Public Domain*, revised edition, Oxford: Clarendon Press 1995, 77-325.

Rescher, Nicholas (1973), *The Coherence Theory of Truth*, Oxford: Oxford University Press.

Sartor, Giovanni (2005), *Legal Reasoning, A Cognitive Approach to the Law*, Dordrecht: Springer.

Searle, John (1979), *Expression and meaning. Studies in the theory of speech acts*, Cambridge: Cambridge University Press.

Sosa, Ernest (1991a), 'The Raft and the Pyramid: Coherence versus Foundations in the Theory of Knowledge', *Midwest Studies in Philosophy* 5 (1980), 3-25. Also in Ernest Sosa, -*Knowledge in Perspective. Selected Essays in Epistemology.* Cambridge: Cambridge University Press,, 165-191.

Sosa, Ernest (1991b), 'Equilibrium in Coherence?' in Ernest Sosa, *Knowledge in Perspective. Selected Essays in Epistemology.* Cambridge: Cambridge University Press, 257-269

Thagard, Paul and Karsten Verbeurgt (1998), 'Coherence as Constraint Satisfaction', *Cognitive Science* 22, 1-24.

Thagard, Paul (2000), *Coherence in Thought and Action*, Cambridge: MIT Press.

Young, James, O. (2008), 'The Coherence Theory of Truth', in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, http://plato.stanford.edu/archives/fall2008/entries/truth-coherence