

# Reflective Equilibrium

Hassan Masoud

Jan. 30, 2012

# Reference

- Norman Daniels: Reflective Equilibrium (SEP)
- James Young: The Coherence Theory of Truth (SEP)
- Jonathan Kvanvig: Coherentist Theories of Epistemic Justification (SEP)
- Michael DePaul: Why Bother with Reflective Equilibrium? In the book 'Rethinking Intuition'

# Introduction

“Many of us, perhaps all of us, have examined our moral judgments about a particular issue by looking for their coherence with our beliefs about similar cases and our beliefs about a broader range of moral and factual issues. In this everyday practice, we have sought “reflective equilibrium” among these various beliefs as a way of clarifying for ourselves just what we ought to do. In addition, we may also have been persuading ourselves that our conclusions were justifiable and ultimately acceptable to us by seeking coherence among them.” [Daniels, 1]

# History

- Nelson Goodman: Fact, Fiction and Forecast

First edition: 1955; Fourth (last) edition: 1983

Harvard University Press

- John Rawls: A Theory of Justice

First edition: 1971; Revised edition: 1999

Harvard University Press

# Definition (as a state)

Reflective Equilibrium: “A state in which all one’s thoughts about a topic fit together; in which there are no loose ends or recalcitrant elements that do not cohere with an overall position.”

Rawls: “the proper method of ethics should be one of trying to achieve reflective equilibrium, testing theories against judgments about particular cases, but also testing judgments about particular cases against theories, until equilibrium is achieved.”  
[Oxford Dictionary of Philosophy]

# Definition (as a method)

“The method of reflective equilibrium consists in working back and forth among our considered judgments (some say our “intuitions”) about particular instances or cases, the principles or rules that we believe govern them, and the theoretical considerations that we believe bear on accepting these considered judgments, principles, or rules revising any of these elements whenever necessary in order to achieve an acceptable coherence among them.” [Daniels, 2]

# The goal

“We arrive at an optimal equilibrium when the component judgments, principles, and theories are ones we are un-inclined to revise any further because together they have the highest degree of acceptability or credibility for us.” [Daniels, 3]

# Applications

“The method of reflective equilibrium has been advocated as a coherence account of justification (as contrasted with an account of truth) in several areas of inquiry, including inductive and deductive logic as well as ethics and political philosophy.” [Daniels, 3]



# Coherentism

- The coherentist theories of *truth*
- The coherentist theories of *justification*

# The coherence theory of truth

“A coherence theory of truth states that the truth of any (true) proposition consists in its coherence with some specified set of propositions.” [Young, 1]

The correspondence theory: the relationship between a proposition and **objects**

The coherence theory: the relationship between a proposition and other **propositions**

# Coherence relation

- Consistency
- Entailment
- Mutual explanatory support

# Specified set of propositions

- The largest consistent set of propositions currently believed by actual people
- Those propositions which will be believed when people like us (with finite cognitive capacities) have reached some limit of inquiry
- the propositions which would be believed by an omniscient being

# Motivations for coherentism

- Metaphysical motivation
- Epistemological motivation

# Metaphysical motivation for coherentism

“Idealists do not believe that there is an ontological distinction between beliefs and what makes beliefs true. From the idealists’ perspective, reality is something like a collection of beliefs. Consequently, a belief cannot be true because it corresponds to something which is not a belief. Instead, the truth of a belief can only consist in its coherence with other beliefs.”

[Young, 4]

# Epistemological motivation for coherentism

- 1- A coherence theory of justification leads to a coherence theory of truth. There is no guarantee that a perfectly coherent set of beliefs matches objective reality.
- 2- We cannot get outside our set of beliefs and compare propositions to objective facts. We can only know that a proposition coheres with a set of beliefs. We can never know that a proposition corresponds to reality.

# Objections to coherentism

- The specification problem
- The transcendence problem



# The specification problem

Coherence theorists have no way to identify the specified set of propositions without contradicting their position.

Possibilities:

- Propositions which correspond to the fact
- The most comprehensive system (based on size, simplicity, empirical adequacy)

# The transcendence problem

A coherence theory of truth is unable to account for the fact that some propositions are true which cohere with no set of beliefs.

Truth transcends any set of beliefs. 'Jane Austen wrote ten sentences on November 17th, 1807' is either true or false. If it is false, some other proposition about how many sentences Austen wrote that day is true. No proposition, however, about precisely how many sentences Austen wrote coheres with any set of beliefs and we may safely assume that none will ever cohere with a set of beliefs.

# Coherentism $\neq$ Idealism

Coherentism denies some principles of realism:

- The principle of bivalence (every proposition is either true or false)
- The principle of transcendence (a proposition might be true even though it cannot be known to be true)

# Coherentist theories of justification

- The coherentist theories of justification are defined as opposed to the foundationalist theories. These theories are different responses to **the regress problem**.

# Justification

- Justification is the requirement for a true belief to be counted as a piece of knowledge.
- Knowledge = justified true belief

# The regress problem

- If any belief has to be justified in order to be accepted, what can be said about those beliefs which are intended to support the former?
- Three possibilities:
  - 1- The chain of reasoning is infinitely long
  - 2- The chain of reasoning stops somewhere
  - 3- The chain of reasoning goes in a circle

# Foundationalism vs. Coherentism

- Foundationalist theories of justification choose the second possibility. They introduce basic beliefs as the stopping point of the chain of reasoning.
- The *linear* version of coherentism chooses the third option and accepts the circularity of the chain of reasoning.
- The *holistic* version of coherentism rejects the trilemma and denies that the chain of reasoning should be linear.

# The holistic account of justification

- Neurath's ship:

“We are like sailors who on the open sea must reconstruct their ship but are never able to start afresh from the bottom. Where a beam is taken away a new one must at once be put there, and for this the rest of the ship is used as support. In this way, by using the old beams and driftwood the ship can be shaped entirely anew, but only by gradual reconstruction.”



# The holistic account of justification

- Quine's fabric:

“The totality of our so-called knowledge or beliefs, from the most casual matters of geography and history to the profoundest laws of atomic physics or even of pure mathematics and logic, is a man-made fabric which impinges on experience only along the edges. Or, to change the figure, total science is like a field of force whose boundary conditions are experience. A conflict with experience at the periphery occasions readjustments in the interior of the field. Truth values have to be redistributed over some of our statements. Re-evaluation of some statements entails re-evaluation of others, because of their logical interconnections.” [Two Dogmas]

# The main issues

- The two main issues in any coherentist theory of justification:
  - 1- The things that must cohere in order for a given belief to be justified
  - 2- The relation that must hold among these things in order for the belief in question to be justified.
- Both features can be given subjective or objective construals.

# The Things Over Which Coherence is Defined

- From a subjective viewpoint:
  - the agent's set of beliefs
- From some objective viewpoints:
  - Social version: common knowledge
  - Religious version: some theological doctrine
- Standard versions of coherentism are subjective about the items relative to which coherence is defined.
- The objective viewpoints deny the perspectival character of justification, according to which whether or not one's beliefs are justified depends on facts about oneself and one's own perspective on the world.

# The Things Over Which Coherence is Defined

- Belief is not the only subjective item to which a theorist might appeal.
- Coherentism need not prohibit the subjective system over which coherence is defined from containing experiential states.

# The relation of coherence

“The most popular **objective** approach is explanatory coherentism, which defines coherence in terms of that which makes for a good explanation. On such a view, hypotheses are justified by explaining the data, and the data are justified by being explained by our hypotheses. The central task for such a theory is to state conditions under which such explanation occurs.” [Kvanvig, 10]

# The relation of coherence

In another **objective** approach (BonJour):

- 1- Logical consistency
- 2- The extent to which the system in question is probabilistically consistent
- 3- The extent to which inferential connections exist between beliefs, both in terms of the number of such connections and their strength
- 4- The inverse of the degree to which the system is divided into unrelated, unconnected subsystems of belief
- 5- The inverse of the degree to which the system of belief contains unexplained anomalies [Kvanvig, 10]

# Problems with Bonjour's criteria

- we need to know how to weight each of these factors to provide an assessment of the overall coherence of the system.
- not all beliefs are justified to the same degree.

So, we need subjective accounts to accommodate the concept of a degree of belief.

# Subjective accounts of the relation of coherence

## Probabilistic Bayesianism:

- Probabilistic coherence: A (degree of) belief is justified if and only if it is part of a system of beliefs against which no dutch book can be made.
- Conditionalization requirement: When new information is learned, one's new degree of belief match one's conditional degree of belief on that information prior to learning it. [Kvanvig, 12]



# Probabilistic Bayesianism

- **Subjectivity:** Each person has an internal, subjective theory of evidence at a given time, in the form of conditional beliefs concerning all possible future courses of experience.
- **Coherentist:** Coherence obtains when a belief conforms to the subjective theory of evidence in question, given the other items in the set of belief.

# Problems for coherentism

- 1- Problems related to the basing relation

“An account of the basing relation is needed to explain the difference between a situation where a person has good evidence for a belief, but believes it for other reasons, and a situation where a person holds the belief because of, or on the basis of, the evidence.” [17]

# The basing relation

A distinction:

- A doxastically justified belief: where the belief is based on the evidence for it
- A propositionally justified belief: when there is good evidence for the belief, but the belief is held on other grounds

# The basing relation

- An example:

“Let the contents of the beliefs be  $p$ ,  $q$ ,  $r$ ,  $s$ , and  $t$ . Further, let each belief imply the next in sequence, i.e.,  $p$  implies  $q$ ,  $q$  implies  $r$ , and so forth. Assume as well that  $p$ ,  $q$ ,  $r$ , and  $s$  are all justified for the person in question. If so, a person can come to justifiably believe  $t$  by inferring from  $p$  to  $q$  to  $r$  to  $s$  and then to  $t$ . Suppose, however, that there are no other inferential relationships here besides the ones already assumed. If the order of inference were from  $p$  to  $s$  to  $r$  to  $q$  and then to  $t$ , believing  $t$  would not be justified.” [18-19]

# The basing relation

- A coherentist reply:  
to distinguish between that which **justifies** a belief  
and that which is **epistemically relevant** to the  
epistemic status of belief

The example:

20 levels of supporters and defeaters

- Does this any help?!

# Problems for coherentism

## 2- The isolation objection (the input problem)

“Nothing about any requirement of coherence dictates that a coherent system of beliefs need receive any sort of input from the world or be in any way causally influenced by the world.  
(BonJour 1985, p. 108)”

# The isolation problem

- A coherentist solution:

To include experiential states among the set of things on which coherence is defined.

# The truth connection

## 3- Problems related to the truth connection

- ‘Fiction’ objection:

“A longstanding objection to coherentism can be expressed by noting that a good piece of fiction will display the virtue of coherence, but it is obviously unlikely to be true.” [22]



# The truth connection

- ‘Alternative systems’ objection:  
“There is always some coherent system to fit any belief into, so that if a person were to make sufficient changes elsewhere in the system, any belief could be justified.” [23]

An objection to this one: Can it be done so easily?!

# The truth connection

- Solutions to the truth connection problem:
  - 1- If the isolation problem can be solved, the problem of truth connection would be prevented.
  - 2- There are cases to show that justification is not a reliable guide to truth:
    - the lottery paradox
    - the preface paradox [24]

# Michael DePaul

Why bother with reflective equilibrium?

# The method of reflective equilibrium

- This method “describes the approach the vast majority of philosophers in fact follow”.
- “I’ll use moral theory as an example, but remember that philosophical inquiry into other matters, for example, knowledge, causation, reference, or the nature of belief, is similarly conducted.” [DePaul, 294]

# The method

Step 1: The philosopher must begin her inquiry regarding morality with the moral beliefs she happens to have.

- Beliefs/judgments:
  - Concerning actual things or actions
  - About imaginary or hypothetical cases
  - Regarding general principles

She should discard any belief or judgment formed in circumstances that obviously make error.

# The method

Step 2: The next task is to construct a theory that accounts for the remaining judgments.

Step 3: If she comes across any conflict between theories and considered beliefs, she tries to bring them into balance through a process of *mutual adjustment* to both her theory and her considered judgments.

# The method

“The philosopher is not bound to revise the theory so that it accords with these judgments. Rather, she must attempt to determine, via further reflection, whether it is the theory or the judgments that, all things considered, she finds more likely to be true, and then revise her beliefs accordingly.” [295]

# The method

“The philosopher must seek an even wider equilibrium. She must also consider the connections between her moral beliefs and principles and the other sorts of beliefs, principles and theories she accepts or rejects.”

The fair process of *mutual adjustment*:

“Neither moral nor epistemic beliefs nor any of the other beliefs that come into play are granted a privileged status.” [296]



# The method

- 1- To reflect upon her beliefs and the logical and evidential interconnections among her beliefs
- 2- To try to construct theories that are intuitively appealing on their own and that account for various categories of beliefs
- 3- To resolve such conflicts as are uncovered in the course of these reflections and efforts at theory construction on the basis of what comes to seem most likely to be correct as a result of still further reflection [297]

# What reflective equilibrium cannot do

1- This method provides no guarantee that it will lead inquirers to **true** beliefs. [297]

“The entire process is guided by nothing more than the inquirer’s own beliefs, judgments, and what seems to the inquirer to be correct upon reflection, given enough screwy initial belief and unusual judgments about how to resolve conflicts, an inquirer could end up accepting just about anything in reflective equilibrium.” [297]

# What reflective equilibrium cannot do

2- The method of reflective equilibrium will not even **reliably** lead inquirers to the truth. [298]

Argument:

I- Almost all philosophers and other thinkers use this method.

II- There is a considerable amount of disagreement among them.

# What reflective equilibrium cannot do

3- The method of reflective equilibrium cannot be counted on to yield **justified** belief. [299]

Argument:

I- Some theories of justification are reliabilist.

II- Justification must be truth conducive.

III- At least it should be objective.

- So, why do we follow this method?
- Because any other method of inquiry is irrational.

# Alternatives

Any alternative approach would either:

(A) Abandon reflection altogether

(B) Direct the inquirer to reflect, but to do so incompletely, that is, to leave certain beliefs, principles, theories, or what have you out of account

(C) Not allow the results of the inquirer's reflections to determine what the inquirer goes on to believe.

# Alternatives

- (A) Abandoning reflection:
  - Either to stick to what she happens to believe.
  - Or, to follow an authority without any reflection or examination.

Both cases lead to **self-contradiction**.

# Alternatives

- (B) Reflecting incompletely

Innocuous examples:

Biologists or physicists; judges or jurors

Other cases (neglecting some parts of beliefs without reflection or without any justification) lead to **self-contradiction**.



# Alternatives

- (C) Not believing what seems most likely to be true

This case more obviously leads to **self-contradiction**.

# The open question

- What about truth?
- Must there be a commitment to the coherence theory of truth in the method of reflective equilibrium?

# Daniels' article (SEP)

- The main difference between Daniels and DePaul is their view about **justificatory value** of the method of reflective equilibrium.

# Justificatory value

“The key idea underlying this view of justification is that we “test” various parts of our system of beliefs against the other beliefs we hold, looking for ways in which some of these beliefs support others, seeking coherence among the widest set of beliefs, and revising and refining them at all levels when challenges to some arise from others... By extension of this account, a person who holds a principle or judgment in reflective equilibrium with other relevant beliefs can be said to be justified in believing that principle or judgment.” [3]

# Foundationalist approaches in ethics

Some subset of our moral beliefs must be:

- Fixed or unrevisable
- Immediately or directly justified
- Justifiable independently of the rest of our moral beliefs

# Problems of moral foundationalism

- Appeal to a moral sense or faculty
- Appeal to apriority of moral intuitions

# Using RE in justification of logic

“Goodman's idea was that we justify rules of inference in inductive or deductive logic by bringing them into reflective equilibrium with what we judge to be acceptable inferences in a broad range of particular cases. No rule of inference would be acceptable as a logical principle if it was not compatible with what we take to be acceptable instances of inferential reasoning.. At the same time, we should correct or revise our views about particular inferences we initially might think are acceptable if we come to see them as incompatible with rules that we generally accept and refuse to reject because they, in turn, best account for a broad range of other acceptable inferences.” [5-6]

# A problem with Goodman's theory

- Not all elements of the everyday reasoning practices of all individuals are justifiable.

Example of fallacious inductive inference:

- The gambler's fallacy

Example of fallacious deductive inference:

- The psychological studies on the affirming the consequent and denying the antecedent



# Goodman's reply

“He insists that practice can and should be corrected as we work back and forth from tentative principles to practice, revising where appropriate, presumably eliminating the sorts of inconsistencies that some psychological studies, and our everyday experiences, reveal.” [6]

# RE in ethics and political philosophy

- Rawl's theory of "justice as fairness"
- Rawls argues that the goal of a theory of justice is to establish the terms of fair cooperation that should govern free and equal moral agents.

# Rawls' theory of "justice as fairness"

- Step 1: To choose initial principles (the context of discovery)

“The appropriate perspective from which to choose among competing conceptions or principles of justice is a hypothetical social contract or choice situation in which contractors are constrained in their knowledge, motivations, and tasks in specific ways... Under these constraints ... rational contractors would choose principles guaranteeing equal basic liberties and equality of opportunity, and a principle that permitted inequalities only if they made the people who are worst off as well off as possible.” [8-9]

# Rawls' theory of "justice as fairness"

- Step 2: To revise considered principles (the context of justification)

“The chosen principles must also match our considered judgments about justice in reflective equilibrium. If they do not, then we are to revise the constraints on choice in the contract situation until we arrive at a contract that yields principles that are in reflective equilibrium with our considered judgments about justice.” [8]

# Narrow Reflective Equilibrium

“To the extent that we focus solely on particular cases and a group of principles that apply to them, and to the extent that we are not subjecting the views we encounter to extensive criticism from alternative moral perspectives, we are seeking only *narrow* reflective equilibrium. Presumably, the principles we arrive at in narrow equilibrium best “account for” the cases examined. Others, however, may arrive at different narrow reflective equilibria, containing different principles and judgments about justice.” [10]

# Wide Reflective Equilibrium

“In a *wide* reflective equilibrium, for example, we broaden the field of relevant moral and non-moral beliefs (including general social theory) to include an account of the conditions under which it would be fair for reasonable people to choose among competing principles, as well as evidence that the resulting principles constitute a feasible or stable conception of justice, that is, that people could sustain their commitment to such principles.” [12]

## But not *too* wide?

“The device of the contract is thus in reflective equilibrium with certain background theories that themselves contain moral beliefs... If Rawls were trying to justify the structure of the contract by appeal to theories that themselves were completely non-moral, then he would be offering the kind of independent justification for the principles that would characterize them as foundational, so the claim that the background theories are themselves moral is part of the rationale for concluding that Rawls is clearly rejecting foundationalism.” [12-13]

# Justice as political

- “In *A Theory of Justice*, Rawls seemed to think that all people might converge on a common or *shared* wide reflective equilibrium that included “justice as fairness,” the conception of justice for which he argues.” [13]
- “In his later work, *Political Liberalism* (Rawls 1993), Rawls abandons the suggestion that all people might converge in the same, shared wide reflective equilibrium that contains his conception of justice.” [14]



# Justice as political

“Complexity, uncertainty, and variation in experience lead human reason, when exercised under conditions of freedom, of the sort protected by the principles of justice as fairness, to an unavoidable pluralism of comprehensive moral and philosophical views. This unavoidable fact of reasonable pluralism makes one key feature of justice as fairness untenable, namely the account Rawls gave of the stability of his preferred conception of justice.” [14]

# Stability

“The test for stability is to ask if people raised under this view would conform to it over time with less strain of commitment than other conceptions would face. In effect, passing the test shows it is worth adopting this view because it will not prove so fragile that it is not worth the effort to institutionalize.” [14]

# The new method

- The problem: The plurality of ideas would prevent the society from reaching a common point.
- To address this problem, Rawls recasts justice as fairness as a “freestanding” political conception of justice on which people with different comprehensive views may agree in an “overlapping consensus.”

# Political Reflective Equilibrium

“The *public* justification of such a political conception involves no appeal to the philosophical or religious views that appear in the comprehensive doctrines that form this overlapping consensus. Instead, we might think of this process of working back and forth among **the key shared ideas in the public**, democratic culture and the articulated features of the political conception of justice as a political reflective equilibrium.” [15]

# Political Reflective Equilibrium

In *A theory of Justice*, shared philosophical arguments justify key elements of the method, but in *Political Liberalism*, the support for each of the elements must derive from the distinctive features of the various comprehensive views. For example, “a Kantian, a Millian, and a religious person who believed in free faith might all support, but for quite different reasons, the idea that agents were free in the sense of being capable of forming and revising their conceptions of the good life.” [18]

# Criticisms of Reflective Equilibrium

## 1- Objection from a utilitarian point of view

Central to the method of reflective equilibrium in ethics and political philosophy is the claim that our considered moral judgments about particular cases carry weight, if only initial weight, in seeking justification.

# Criticisms of Reflective Equilibrium

- A traditional criticism of utilitarianism is that it leads us to moral judgments about what is right that conflict with our “ordinary” moral judgments.
- A utilitarian response to this claim is that these judgments as pre-theoretical “intuitions” that probably result from cultural indoctrination and thus reflect superstition, bias, and mere historical accident. On this view, moral intuitions or judgments should have no evidentiary credentials and should play no role in moral theory construction or justification. [20]

# Criticisms of Reflective Equilibrium

## 2- Rejection of Rawls's Constructivism

Rawls claims that his view of justice is **constructivist**, meaning that he appeals to some general claims about the nature of persons as well as some empirical facts about human behavior or institutions as part of the justification for the principles of justice.



# Criticisms of Reflective Equilibrium

- Cohen's Objection:

Constructivism combines considerations of justice with other considerations (both empirical and moral). As a result, it does not tell us what justice itself requires.

# Criticisms of Reflective Equilibrium

- A reply to Cohen:

A possible reply to Cohen is that his view about constructivism collapses into his controversial metaethical claim that **principles of justice cannot rest on general facts about human behavior or anything else.**

# Criticisms of Reflective Equilibrium

## 3- Epistemological criticisms

- The vagueness of the concept of coherence
- The inseparability of coherence accounts of justification from coherence accounts of truth
- The overemphasis on the human rationality

Thank You!