

# Defective Equilibrium

DANNY FREDERICK

Slate House, Hunstan Lane  
Old Leake. Boston. PE22 9RG. United Kingdom  
dannyfrederick77@gmail.com

RECEIVED: 19-06-2016 • ACCEPTED: 22-09-2016

**ABSTRACT:** I argue that the conception of reflective equilibrium that is generally accepted in contemporary philosophy is defective and should be replaced with a conception of fruitful reflective disequilibrium which prohibits ad hoc manoeuvres, encourages new approaches, and eschews all justification in favour of continuous improvement. I suggest how the conception of fruitful disequilibrium can be applied more effectively to moral enquiry, to encourage genuine progress in moral knowledge, if we make moral theory empirically testable by adopting a meta-ethical postulate which is independently plausible.

**KEYWORDS:** Ad hoc – defective equilibrium - empirical testability – growth of knowledge – justification – reflective equilibrium.

## 1. Reflective equilibrium as state and method

Reflective equilibrium has been recommended with respect to empirical science by Otto Neurath (see Neurath 1983, 94-95), with respect to logic by Nelson Goodman (see Goodman 1983, 61-64), with respect to ethics by John Rawls (see Rawls 1999, 40-46), and across the whole field of enquiry by other philosophers, including W. V. Quine (see Quine 1951, 39-43), though the *term* was coined by Rawls (1999, 18). A standard exposition of reflective equilibrium might run as follows.

We are in reflective equilibrium when the set of propositions we accept satisfies two conditions: its elements are mutually consistent; and some of its elements provide the best available explanation for some of the others (the latter are then said to ‘support’ the former). As Norman Daniels puts it:

[W]e achieve reflective equilibrium when we arrive at an acceptable coherence ... An acceptable coherence requires that our beliefs not only be consistent with each other (a weak requirement), but that some of these beliefs provide support or provide a best explanation for others. (Daniels 2011, sec. 1).

Mutual consistency is, in principle, an all-or-nothing affair: either two propositions are consistent or they are not. In practice things are not so simple because of vagueness or indeterminacy. Explanatory coherence, on the other hand, is a matter of degree. For example, suppose that we are in a reflective equilibrium in which a proposition,  $P$ , which is in our accepted set of propositions, explains a subset,  $S$ , of those propositions. Suppose, further, that a proposition,  $P'$ , which is currently outside our set of accepted propositions, explains  $P$ , and thus also  $S$ , and also explains some propositions in our set which  $P$  does not explain. Then, by adding  $P'$  to our set of accepted propositions, we will increase the coherence of our set of accepted propositions, thereby moving from one state of reflective equilibrium to a more coherent, and thus better, state of reflective equilibrium.

Philosophers usually recommend reflective equilibrium as a desirable *end-state*, so they recommend that we seek to remove any inconsistencies in the set of propositions we currently accept and that we choose the best of any available rival explanations. The process through which equilibrium is achieved is sometimes called the *method* of reflective equilibrium. It involves comparing our accepted propositions, noting any inconsistencies, and then revising either particular or general propositions to achieve greater coherence by eliminating the inconsistencies and perhaps also improving the explanations:

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

wherever necessary in order to achieve an acceptable coherence among them. (Daniels 2011, sec. 1).

One might wonder why our enquiries should be governed by the norms of reflective equilibrium theory. Since enquiry is a purposeful activity, one would expect reflective equilibrium to be commended as helpful in achieving the purpose of that activity. However, on the reasonable assumption that the purpose of enquiry is to extend our knowledge, the practice of reflective equilibrium would seem to be counter-productive. Specifically, reflective equilibrium theory has the following shortcomings:

- (i) it permits inconsistencies to be removed by ad hoc manoeuvres;
- (ii) it fails to acknowledge explicitly the essential contribution of increases in reflective *disequilibrium* to the growth of knowledge;
- (iii) it takes static equilibrium, rather than ongoing improvement, as the ideal.

I explain these shortcomings of reflective equilibrium theory in turn, making use of illustrations from the history of science, and I propose a dynamic conception of fruitful reflective disequilibrium to replace it. I outline some apparent difficulties in applying the conception of fruitful reflective disequilibrium to promote the growth of moral knowledge; then I suggest that the difficulties might be overcome if we adopted a reasonable meta-ethical postulate that renders moral theories empirically testable.

## 2. Ad hoc manoeuvres

If we discover an inconsistency within the set of propositions we accept, there are better and worse ways of eliminating it. An ad hoc manoeuvre is one that eliminates the inconsistency without teaching us anything new (cf. Popper 1959, sections 19-20). If our purpose in enquiry is to extend our knowledge, to learn something new, then we should demand that a move in the direction of reflective equilibrium that removes an inconsistency should also explain something new. The point can be illustrated with two examples from the history of science.

First, in the mid-nineteenth century, the observed motions of Uranus conflicted with the predictions of Isaac Newton's theory. There were thus incon-

sistencies between accepted observation statements and a previously successful explanatory theory. The inconsistencies could have been removed by simply rejecting the observation statements as hallucinatory, or rejecting Newton's theory, or amending Newton's theory so that it made an exception of Uranus, or positing the existence of a special force acting on Uranus which had no other effects. However, each of those manoeuvres would have been ad hoc: they would have eliminated the inconsistencies without teaching us anything new.

A better way of eliminating the inconsistencies was proposed by Urbain Leverrier. Realising that the refuted predictions followed from the conjunction of Newton's theory with accepted background knowledge, he proposed to replace some of that background knowledge. Thus, he accepted both Newton's theory and the observation statements about Uranus, but he denied that the known planets are all the planets there are: he hypothesised a new fact, namely, the existence of another planet with just the properties necessary to account for Uranus' anomalous motions in terms of Newton's theory. What saved this from being ad hoc was that the new hypothesis could be tested independently, for it implied that the new planet would be seen in a specific portion of the sky at a particular time. The hypothesis passed that test: Neptune was discovered (cf. Kuhn 1957, 261-262). Thus, Leverrier's manoeuvre did not simply remove the inconsistencies in our accepted propositions; it also explained (indeed, successfully predicted) something new, namely, the observation statements concerning the positions of Neptune.

Second, in the late seventeenth century a number of observation statements accepted by the Astronomer Royal were inconsistent with the predictions of Newton's theory. The observation statements could have been rejected as hallucinatory or the result of incompetence, or Newton's theory could have been rejected; but such moves would have been ad hoc. Instead, Newton amended background knowledge concerning the way that the Earth's atmosphere refracted light, which makes heavenly bodies appear to be some distance from where they actually are. The amended theory of atmospheric refraction explained why the previously accepted observation statements were false and also why they had seemed to be true; and it survived independent tests and thus explained something new (see Lakatos 1978, 215-216).

Rawls recognises that when judgements about particular cases are rejected because they conflict with a general principle, it would be an advantage to have an explanation for why the particular judgements seemed acceptable, but he

does not require that such an explanation be provided and he does not require that the explanation also explain something new:

When a person is presented with an intuitively appealing account of his sense of justice (one, say, which embodies various reasonable and natural presumptions), he may well revise his judgments to conform to its principles even though the theory does not fit his existing judgments exactly. He is especially likely to do this if he can find an explanation for the deviations which undermines his confidence in his original judgments and if the conception presented yields a judgment which he finds he can now accept (Rawls 1999, 42-43).

It might be rejoined that the standard accounts of reflective equilibrium recognise implicitly the superiority of adjustments to particular or general propositions which are not ad hoc, in that eliminating an inconsistency in a way which explains something new increases the coherence of the set of propositions we accept and thereby achieves a better reflective equilibrium. However, if the stricture against ad hoc manoeuvres is not stated explicitly as a *requirement*, such manoeuvres will be deemed acceptable whenever there happens to be no better, more coherent, equilibrium currently available. That removes the imperative to *increase* our knowledge; it thereby condones stagnation.

It may be complained that the prohibition on ad hoc manoeuvres is too challenging. For instance, it might be said that, if we think that all ravens are black, then discover a white raven and thus give up the theory that all ravens are black, we are doing nothing amiss; we are rather making reasonable adjustments to achieve a new reflective equilibrium. However, if the purpose of enquiry is to extend our knowledge, we should not be content with such an ad hoc manoeuvre. We should instead try ways of removing the inconsistency that promise to teach us something new. For example, we could impugn the assumption that white is the natural colour of the anomalous raven. A new hypothesis that the raven had been painted white would explain why the observation statement that the raven is white appeared to be true and (in conjunction with background knowledge) it would have testable implications concerning how the white colour could be removed. If the implications survive the tests, the hypothesis has predicted, and thus explained, something new. Alternatively, we might add a qualifying condition to the generalisation that all ravens are black that not merely allows some non-black ravens but also implies that

we will find other non-black ravens under specific circumstances that we can either discover or construct. If the modified generalisation survives testing, it will teach us something new. It is true that the proscription of ad hoc manoeuvres is a challenging demand that it will often be difficult to meet, and theorists may sometimes have to make numerous attempts to improve upon their resolution of an inconsistency in order to meet it. But wherever an inconsistency is removed without meeting the demand, that should be highlighted as a defect requiring eventual amelioration.

It might be objected that the notion of teaching us something new, on which the identification of ad hoc manoeuvres depends, is vague. After all, any ad hoc adjustment of a theory will teach us how the unadjusted theory can be made consistent with the proposition(s) with which is inconsistent; so it will teach us *something* new. However, such a novelty is not novel enough: an adjustment that is not ad hoc successfully predicts or explains something, or solves a problem, that is independent of the problem it was introduced to solve. Ad hocness is a matter of degree, so we can expect some borderline cases; and whether something is explanatory or novel or a solution to a problem often involves qualitative considerations and thus judgement, so we can expect some disputed cases. There are, however, many clear-cut cases. It seems clear that the successful Newtonian adjustments outlined above are not ad hoc, as they entailed unexpected empirical predictions which survived testing; and the same goes for the hypothesis that the raven was painted white, if it survives testing. It seems clear, too, that amending Newton's theory to make an exception for Uranus would have been ad hoc. The fact that there can be no general algorithm for ad hocness does not detract from the notion's usefulness.

### 3. Creative disequilibrium

One thing we know from the history of science is that the growth of knowledge is brought about by people who create reflective disequilibrium by discovering or generating an inconsistency. That sets in train attempts to eliminate the inconsistency by making modifications to our accepted body of theory; and those equilibrating modifications will in turn make further contributions to the growth of knowledge, so long as they are not ad hoc. That may be done in three connected ways.

First, it is characteristic of science that accepted theories are subjected to experimental tests. But an experimental test of a theory is a serious attempt to refute it, that is, an attempt to produce an inconsistency between the theory and an experimental result. Experimental results inconsistent with a previously successful theory can be acceptably explained away if the explanations also explain something new, as in the cases of Leverrier and Newton mentioned in section 2.

Second, scientific knowledge is often augmented by means of thought-experiments which disclose inconsistencies within existing theories and thereby lead to conceptual change, as with Galileo Galilei's criticism of Aristotelian dynamics. Conceptual changes that dissipate paradoxes should also explain something additional, as Galileo's distinction between average and instantaneous speed not only resolved the paradoxes revealed in his thought-experiment but also enabled the solution of problems involving accelerated motion (cf. Kuhn 1977).

Third, we increase our scientific knowledge by proposing new theories that contradict previously successful theories, as Johannes Kepler's astronomy contradicted Nicolaus Copernicus's system, as Newton's mechanics contradicted both Kepler's and Galileo's theories, and as Albert Einstein's relativity theories contradicted Newton's (cf. Kuhn 1957; Popper 1983, 75, 131-149). Where a new theory contradicts a currently successful theory, removing the inconsistency will require that at least one of them is rejected (at least in its current form). Rejecting the new theory on grounds of tradition, or conservatism, would be ad hoc, as would rejecting the old theory because it is old. But if the new theory can explain not only the success of the old theory but also something else, then accepting the new in preference to the old is not ad hoc. For example, Newton's theory not only contradicted Kepler's and Galileo's theories, it also explained why they were successful (their predictions about celestial and terrestrial motions, respectively, were approximately accurate) and it explained other things besides, such as the motions of the tides (see Popper 1983, 139-145, 190-191).

We should add a couple of qualifications. First, explaining the success of a prior theory does not necessarily mean explaining everything it explained. Some of the questions answered by the prior theory may have been artefacts of the assumptions of the theory, so that when a new theory jettisons those assumptions it does not answer those questions but rather explains them away as pseudo-problems (as relativity treats 'what is the absolute velocity of the

earth?"). Second, some problems solved by the prior theory which are not thus explained away might remain unsolved by the new theory for some time; but until they are solved, perhaps by further novel developments of theory, the new theory cannot be regarded as having fully superseded the prior theory.

Thus, the intellectual pioneer creates an inconsistency which he then attempts to eliminate in a way that increases overall explanatory coherence. If he succeeds, the reflective disequilibrium he introduced was temporary and was a means to improved knowledge. The growth of knowledge requires an interplay of equilibrating and disequilibrating changes.

Rawls notes that the "kind of reflective equilibrium that one is concerned with in moral philosophy," involving the idea of being "presented with all possible descriptions to which one might plausibly conform one's judgments together with all relevant philosophical arguments for them," may lead to a "radical shift" in one's views; but he goes on to say that "[t]he most we can do" is to study traditional theories "and any further ones that occur to us" (Rawls 1999, 43). In contrast, a commitment to the growth of knowledge would urge not that we leisurely wait for new possibilities to occur to us, but rather that we *actively contrive* new theories to create a disequilibrium that may lead to a radical shift in our views.

Other recent philosophers seem to have even less appetite for revolutionary change. W. V. Quine and J. S. Ullian explicitly counsel conservatism: "In order to explain the happenings that we are inventing it to explain, the [new] hypothesis may have to conflict with some of our previous beliefs; but the fewer the better" (Quine & Ullian 1978, 66); "one is not apt to be tempted by a hypothesis that upsets prior beliefs when there is no need to resort to one" (Quine & Ullian 1978, 67). Quine and Ullian do allow that the maxim of conservatism can be set aside where a revolutionary new theory offers dramatic gains in simplicity or generality (cf. Quine & Ullian 1978, 75-76); but if theorists took the maxim of conservatism seriously, they would never take the time and trouble to work out a revolutionary new theory, so they would rarely, if ever, encounter a situation in which the maxim of conservatism could be set aside. It seems that Quine and Ullian see the purpose of enquiry as being to settle upon a coherent set of propositions with which we feel comfortable. But if our purpose in enquiry is to extend our knowledge, we should welcome, and encourage, the discomforts created by theoretic innovators.

#### 4. Progress

Advocates of reflective equilibrium pay little attention to the pioneer and they usually incorporate no explicit prohibitions on ad hoc manoeuvres. Further, they portray the state of static equilibrium as an ideal. For example, Daniels (2011, sec. 1) says:

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.

David Lewis writes:

Our “intuitions” are simply opinions; our philosophical theories are the same ... and a reasonable goal for a philosopher is to bring them into equilibrium. Our common task is to find out what equilibria there are that can withstand examination, but it remains for each of us to come to rest at one or another of them. (Lewis 1983, x)

Rawls allows that it is doubtful whether one can ever reach the state of reflective equilibrium, but he still regards the state as a “philosophical ideal” (cf. Rawls 1999, 43-44). Geoffrey Sayre-McCord agrees (cf. Sayre-McCord 1996, 142).

In addition, accounts of reflective equilibrium are often combined with coherence theories of truth or of justification. Thus, Goodman opines:

[D]eductive inferences are justified by their conformity to valid general rules, and ... general rules are justified by their conformity to valid inferences. But this circle is a virtuous one. The point is that rules and particular inferences alike are justified by being brought into agreement with each other. *A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend.* The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement achieved lies the only justification needed for either. (Goodman 1983, 64)

Similarly, Rawls says that, in reflective equilibrium, “we have done what we can to render coherent and to justify our convictions... A conception of justice

cannot be deduced from self-evident premises or conditions on principles; instead, its justification is a matter of the mutual support of many considerations, of everything fitting together into one coherent view” (Rawls 1999, 18-19). Sayre-McCord agrees: “as one approaches a (wide) reflective equilibrium one thereby increases the extent to which the beliefs one holds are epistemically justified” (Sayre-McCord 1996, 143). Such accounts could be welcomed by a complacent dogmatist who is more concerned to ‘get his story straight’ than to better understand the world (for an effective critique of such accounts of justification see Stich 1998).

If, in contrast, our aim is the growth of knowledge, a rest-state of reflective equilibrium, far from being an ideal, is not even desirable. We want development, not stasis. If at some time we happened to achieve consistency and coherence in our accepted propositions, our next theoretical task should be to upset that equilibrium by seeking novel facts or paradoxical implications to refute some currently successful theory or by developing a novel theory to replace an existing one. Ideally, a reflective equilibrium would never be attained: progress toward a reflective equilibrium would always be upset by a new disequilibrating intervention followed by equilibrating efforts which are in turn challenged by further disequilibrating novelty; and so on indefinitely. Our ideal is a fruitful reflective disequilibrium that generates unending improvement in our knowledge by means of:

- (i) active search for inconsistencies within our currently accepted knowledge;
- (ii) pursuit of new facts and development of new hypotheses which are inconsistent with our current knowledge and that offer the prospect of radical change;
- (iii) achievement of greater explanatory coherence by removing inconsistencies in ways which are not ad hoc and by supplanting an old theory by a new one when, but only when, the new one provides better explanations.

This dynamic conception of fruitful reflective disequilibrium is not fitted to provide an account of truth or of justification. If our aim is the growth of knowledge, and we recognise that new knowledge often contradicts previously accepted theories and observation statements, we should never maintain that our currently accepted propositions, however good a set they may make, are

either true or justified. We should view them always as more or less ephemeral steps in the progress of our knowledge. We should not seek to justify our theories: we should seek to replace them with better ones.

It might seem that advocates of reflective equilibrium could accommodate these criticisms. Few of them, if any, would claim to have reached the ideal of reflective equilibrium; they would regard any equilibrium currently attained as being justified only defeasibly and thus open to revision in the light of new discoveries. Further, given a choice between an ad hoc resolution of an inconsistency and an alternative resolution that teaches us something new, they would be expected to prefer the latter, other things being equal. That, of course, is true. The problem with reflective equilibrium theory is that it does not *encourage* the growth of knowledge: it does not *require*, or even *commend*, the active search for counterexamples, paradoxes, novel theories, and equilibrating adjustments that avoid ad hoc manoeuvres. As we have seen, its advocates discourage change, especially revolutionary change, and when they concede that a change is necessary, they are happy to accept an ad hoc change if it leads them back to a state of rest. Reflective equilibrium is for the shiftless, who are more interested in attempted justification than in improvement through criticism. Fruitful reflective disequilibrium is for those who are restless for the growth of knowledge:

[S]cientific progress is revolutionary. Indeed, its motto could be that of Karl Marx: 'Revolution in permanence'. (Popper 1994, p. 12)

## 5. Moral knowledge

The dynamic conception of fruitful reflective disequilibrium was illustrated above with examples from empirical science, but it should also apply to moral enquiry. If the aim of such enquiry is to improve our understanding of moral matters, to extend our moral knowledge, then we should not want simply to achieve consistency and coherence in our particular moral judgements and general moral principles. We should rather strive for progress in moral enlightenment by insisting that:

- (i) consistency is achieved in ways that are not ad hoc;
- (ii) complacency (the search for justification) is abjured;

- (iii) there is vigorous encouragement of attempts to increase disequilibrium by attempting to generate
  - paradoxes within existing theory,
  - inconsistencies between accepted theory and judgements about particular cases,
  - novel theories which contradict accepted theories and which promise to have greater explanatory merit.

We noted above that reflective equilibrium theory includes none of those demands and is thus unfit to guide intellectual enquiry. Our illustrations showed that it is out of line with good practice in empirical science. How does it compare with the practice of moral enquiry?

Contemporary moral philosophers generally accept reflective equilibrium theory, but while their practice conforms to the theory in ignoring demands (i) and (ii), it diverges from it in partially meeting demand (iii), in two ways. First, moral philosophers often contrive imaginative thought-experiments, describing unusual possibilities, through which our moral theories can be re-shaped, or which yield intuitively acceptable moral judgements about particular cases against which moral principles can be tested. Rawls' original contractual position is such a thought-experiment (cf. Rawls 1999, 11-19, 102-168). Second, some moral philosophers are prepared to propose or consider revisionary moral theories, such as Stoicism, act-utilitarianism, egoism, feminist ethics, vegetarianism and so on, that contradict previously accepted moral principles and which generate moral judgements about particular cases that contradict previously accepted moral judgements about those cases. Thus, what moral philosophers actually do is better than what they say they should do.

Moral theorists would make a stride forward if, adhering to (ii), they insisted on progress in moral knowledge by refusing to accept that a moral theory or judgement can be justified and resolving instead to try to improve upon even the best moral theory that they have so far achieved. They would then reject Donald Davidson's claim that "we should expect people who are enlightened *and fully understand one another* to agree on their basic values" (Davidson 2004, 49). If, as I have argued, enlightenment involves an ongoing process of discovery, rather than being a state that can be (or is already) achieved, then, however enlightened people are, they should seek to become more enlightened, so they should encourage each other to dispute currently accepted moral propositions.

Moral theorists would make a further stride forward if, following (i), they insisted that amendments to theory which remove inconsistencies should not be ad hoc. Unfortunately, it is not easy to see how moral enquiry could generally meet the demand that acceptable amendments to theory should explain something new in addition to solving the problem for which they are proposed. In the empirical sciences, a new amendment to theory may entail novel factual propositions which can then be checked by observation, as in the examples given in section 2. However, that option for avoiding ad hoc manoeuvres is not available to moral theorists given the a priori character of moral enquiry. They would instead need to ensure that an amendment to theory explains a moral judgement or a moral principle that was not previously explained, in addition to resolving the difficulty for which it is proposed. That may seem to make the stricture against ad hoc manoeuvres very demanding, perhaps impossibly so.

There is another worry. Suppose that moral theorists could meet that demand. It might be that all the rival moral theories could, given time, be modified in non-ad-hoc ways to overcome internal inconsistencies. In that case, we face the prospect of numerous alternative moral theories, each in reflective equilibrium, each contradicting each of the others, and none having greater explanatory merit than any other, since each is satisfactorily explanatory in its own terms. Lacking an empirical test, we seem to have no way of rating the rival moral theories as epistemically better or worse. Yet, from a commonsensical point of view, it may seem clear that some moral theories are epistemically better than others. For example, suppose that a Buddhist moral theory and an Islamic Fundamentalist moral theory are each self-consistent and that each is modified in non-ad-hoc ways in response to judgements about unusual actual cases or imagined possible cases. On what has been said so far, we have no argument available for preferring one over the other. Yet, if we compare how people flourish or suffer in communities which adhere largely to the one moral theory vis-à-vis those which adhere largely to the other, we are presented with a stark contrast. That contrast seems highly relevant morally, which suggests that the point or function of morality is to facilitate the fulfilment of persons. Indeed, that idea is explicit in theistic accounts of morality that invoke God's plan for His creation. As Robert Young put it: "For many, including Judaeo-Christians, promoting the well-being of humans (and perhaps all sentient creatures) *is* the whole point, or a large part of the point, of having moral principles at all" (Young 1981, 162). The idea is not peculiar to theists: rule-consequentialist accounts of morality make the point explicitly, and act-consequentialist

accounts make it crudely; contractarian and contractualist accounts attribute morality to agreement, either for selfish benefit or for mutual benefit or for the benefit of all; and some evolutionary accounts explain morality as an adaptation that benefits the species. That suggests that we may be able, with some ingenuity, to render moral theories empirically testable, if we accept something like the following meta-ethical postulate:

- (m) The true moral theory is the one such that the best prospects for the fulfilment of persons would be realised if that theory were to be adhered to universally by people as they actually are.

Different moral theories assign different rights and duties to people, and different assignments of rights and duties imply different social structures. For example, unequal rights between races or sexes imply institutions of slavery or the subordination of women (or men), and duties to avoid some specific types of sexual activity imply the social exclusion of homosexuals. An understanding of the effects on human fulfilment of people universally doing their duty as defined by a particular moral theory therefore requires a social-scientific investigation of the consequences of types of action in types of social circumstances. So, ascertaining the consequences for human fulfilment if a particular moral theory were universally acted upon is a matter for the social sciences.

The elaboration of that idea for rendering moral theories empirically testable, and thus better able to meet the challenge of non-ad-hoc development, would require a separate (book-length) discussion. Here, though, we can consider the objection that (m) is an ethical theory, rather than a meta-ethical postulate, because it is tantamount to the claim that the fulfilment of persons is the ultimate moral value. The objection is misplaced for a number of reasons. First, (m) does not imply that there is an ultimate moral value. It may be, for example, that the best prospects for the fulfilment of persons would be realised if people adhered to a moral theory which incorporated a plurality of moral values none of which is overriding or ultimate. In particular, the postulate (m) does not reduce to ‘maximise human fulfilment,’ any more than rule-consequentialism reduces to act-consequentialism (on the latter, see Frederick 2016, 25-26). Second, while it is true that (m) is contestable and may be rejected by some, it should elicit broad assent because it appears to be at least tacitly accepted by so many different moral theorists. An aspiration for universal assent

would plainly be unrealistic. Third, it may be instructive to compare (m) with the following ‘meta-descriptive’ postulate:

- (d) Rival descriptions of reality can be evaluated epistemically by testing them against our sensory experiences.

Accepting (d) does not commit us to the claim that sensory experiences are the ultimate reality; and it does not reduce to the claim that descriptions of reality must use only observational terms. It does involve the descriptive claim that observations generally bring us into contact with reality; but that claim is assumed by the great majority of inquirers, so (d) should elicit broad assent, even though it is denied by some mystics and even by some rationalist philosophers, such as Parmenides.

## 6. Conclusion

It is a standard view in contemporary philosophy that our intellectual enquiries are, or should be, an endeavour to achieve a state of reflective equilibrium in which the propositions to which we subscribe are rendered consistent and explanatorily coherent. I have argued that that view is unacceptable if the aim of our enquiries is to extend our knowledge, because:

- (i) it does not take sufficiently serious account of the fact that equilibrating moves may be ad hoc and thus fail to extend our knowledge;
- (ii) it does not encourage the disequilibrating moves that are necessary for the growth of knowledge;
- (iii) it regards a settled equilibrium, particularly one which is supposed to confer ‘justification’ on the propositions we accept, or on our acceptance of them, as an ideal, rather than as a deplorable state of stagnation.

Consequently, I have proposed that the method of reflective equilibrium should be replaced by a method of fruitful reflective disequilibrium which:

- (i) demands moves toward reflective equilibrium which eschew ad hoc manoeuvres;

- (ii) encourages moves toward reflective disequilibrium which generate inconsistencies either within accepted theories, or between accepted theories and accepted observation statements, or between accepted theories and potential new rivals;
- (iii) abhors the idea of a static equilibrium and embraces the ideal of unending improvement.

The method of fruitful reflective disequilibrium is implicit in the progress of scientific knowledge, as I have indicated with some illustrations. That is not to say that all scientists practise it, let alone advocate it. It is to say only that, insofar as science has made genuine progress, which it does seem to have done, the method of fruitful reflective disequilibrium can be seen at work. In areas where science has stagnated, that might be explained by adherence to the method of reflective equilibrium.

Application of the method of fruitful reflective disequilibrium to moral theory should stimulate the pursuit of moral enlightenment; but there are some doubts as to how effectively that can be done insofar as moral enquiry is pursued a priori. I have suggested that better progress might be made if we rendered moral theories empirically testable by adopting a plausible meta-ethical postulate linking the truth of a moral theory with the consequences for human fulfilment that would follow upon universal adherence to the theory. We might then hope to see advances in moral knowledge akin to the advances in the empirical sciences.

The champions of reflective equilibrium could deflect my critique by saying that the aim of their enquiries is not to extend our knowledge. One would then be left to wonder what the value of their enquiries might be.<sup>1</sup>

## References

- DANIELS, N. (2011): Reflective Equilibrium. In: Zalta, E. (ed.): *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), available at: <http://plato.stanford.edu/archives/spr2011/entries/reflective-equilibrium/>
- DAVIDSON, D. (2004): The Objectivity of Values. Reprinted in: Davidson, D.: *Problems of Rationality*. Oxford: Clarendon, 39-57.

---

<sup>1</sup> I thank an anonymous referee for this journal for objections and queries that helped me to clarify the argument of this paper.

- FREDERICK, D. (2016): The Good Bishop and the Explanation of Political Authority. *De Ethica* 3, No. 2, 23-35.
- GOODMAN, N. (1983): *Fact, Fiction, and Forecast*. 4<sup>th</sup> ed. Cambridge: Harvard University Press.
- KUHN, T. (1957): *The Copernican Revolution*. Cambridge: Harvard University Press.
- KUHN, T. (1977): A Function for Thought Experiments. Reprinted in: Kuhn, T.: *The Essential Tension*. Chicago: University of Chicago Press, 240-265.
- LAKATOS, I. (1978): Newton's Effect on Scientific Standards. In: *Philosophical Papers, Vol. 1*. Worrall, J. & Currie, G. (eds.), Cambridge: Cambridge University Press, 193-222.
- LEWIS, D. (1983): *Philosophical Papers, Vol. 1*. Oxford: Oxford University Press.
- NEURATH, O. (1983): Protocol Statements. Reprinted in: Neurath, O.: *Philosophical Papers 1913–1946*. Cohen, R. S. & Neurath, M. (eds.), Dordrecht: Reidel, 91-99.
- POPPER, K. (1959): *The Logic of Scientific Discovery*. London: Hutchinson.
- POPPER, K. (1994): The Rationality of Scientific Revolutions. Reprinted in: Popper, K.: *The Myth of the Framework*. London: Routledge, 1-32.
- POPPER, K. (1983): *Realism and the Aim of Science*. London: Routledge.
- QUINE, W. V. (1951): Two Dogmas of Empiricism. *Philosophical Review* 60, No. 1, 20-43.
- QUINE, W. V. & ULLIAN, J. S. (1978): *The Web of Belief*. 2<sup>nd</sup> ed. New York: McGraw-Hill.
- RAWLS, J. (1999): *A Theory of Justice*. Revised ed. Cambridge: Belknap Press.
- SAYRE-MCCORD, G. (1996): Coherentist Epistemology and Moral Theory. In: Sinnott-Armstrong, W. & Timmons, M. (eds.): *Moral Knowledge?* Oxford: Oxford University Press, 137-189.
- STICH, S. (1998): Reflective Equilibrium, Analytic Epistemology and the Problem of Cognitive Diversity. In: DePaul, M. & Ramsey, W. (eds.): *Rethinking Intuition*. Lanham: Rowman and Littlefield, 95-112.
- YOUNG, R. (1981): Theism and Morality. Reprinted in: Helm, P. (ed.): *Divine Commands and Morality*. Oxford: Oxford University Press, 154-164.