

Hume
on
Motivation and Virtue:
New Essays

Edited

By

Charles R. Pigden

TO ZENA

For putting up with a philosopher for thirty-two years

Contents

Preface and Acknowledgements

Notes on the Contributors

A Note on References to Hume and Locke

Introduction

Charles Pigden

1. Expressivism, Motivation Internalism, and Hume

Richard Joyce

2. Is Hume Inconsistent? — Motivation and Morals

Norva Lo

3. If Not Non-Cognitivism, Then What?

Charles Pigden

4. The Motivation Argument for Non-cognitivism

Michael Smith

5. Experiences of Value

Graham Oddie

6. Hume and the Debate on Motivating Reasons

Constantine Sandis

7. Against all Reason: Scepticism about the Instrumental Norm

Stephen Finlay

8. Why Internalists about Reasons Should be Humeans about Motivation.

Kent Hurtig

9. Humean Sources of Normativity

Herlinde Pauer-Studer

10. Two Kinds of Normativity

Luke Russell.

11. What Kind of Virtue-Theorist is Hume?

Christine Swanton

12. Kinds of Virtue Theorist: A Response to Christine Swanton

Annette Baier

13. Reply to Annette Baier

Christine Swanton

14.. Hume on Justice.

Rosalind Hursthouse.

Consolidated Bibliography

Index

Preface and Acknowledgments

There must be mistakes that it is possible for an editor to make that I haven't made and misfortunes that can befall an editor that haven't befallen me, but if so, I am not anxious make those mistakes or experience those misfortunes. So first of all, I would like to thank my contributors collectively, both for their work, and for their patience in putting up with a project that has taken such a *very* long time to come to fruition, partly because of my blunders. Some who were just starting out in 2003, are now not only rising but *risen* stars, and I apologize to them in particular for occluding their sparkle for such long time. As for my older contributors with plenty of publications to their credit, it is still a bit annoying to have some of that credit deferred, and I would like to apologize to them too. I was forced by reasons of space to omit a couple of papers I would have liked to have included. My apologies to those I have disappointed and inconvenienced. My thanks to Palgrave Macmillan for rescuing the project when it seemed about to fall through, and to my colleagues at Otago - particularly Peter Anstey and James Maclaurin for their judicious combination of nagging and encouragement. I would like to thank my research assistant Rebecca Thompson for her editorial services and the research cluster on Early Modern Thought at the University of Otago for providing the wherewithal to pay for them. Thanks to Annette Baier, Josh Parsons and Helen Beebee for useful discussions. Finally, my family. Even those bachelor philosophers of long ago - such as David Hume himself - were often heavily reliant on family support and I am not one of those bachelor philosophers. So thanks to my wife Zena, my mother Jean and my (now grown-up) children, Guy Jemima and Abigail for their love, help and encouragement.

Notes on the Contributors

Annette C. Baier was educated at the universities of Otago and Oxford. She writes on Hume and on trust-centered ethics. She has taught at the universities of Aberdeen, Auckland, Sydney and Pittsburgh. In retirement she lives in her native New Zealand.

Stephen Finlay is an assistant professor at the University of Southern California, and specialises in meta-ethics. Originally from New Zealand, he completed his Ph.D. at the University of Illinois at Urbana-Champaign in 2001, and now lives in Los Angeles with his wife Sarah and two daughters.

Kent Hurtig was born in Sweden and studied philosophy at Wichita State University in Kansas, before going on to do an M.Litt and a PhD at University of St Andrews, Scotland. Since 2005 he has taught philosophy at University of Stirling, Scotland. He has published on topics in meta-ethics, practical reason and rationality.

Rosalind Hursthouse did her undergraduate degree and MA in philosophy at the University of Auckland, and a B.Phil and D.Phil at Oxford, where she taught for six years before becoming a lecturer at the Open University. She found this so rewarding and inspiring that she stayed there happily for the next twenty-five years, until even more happily, she returned to the Auckland Department as Professor. She is the author of *Beginning Lives* and a series of articles on ethical questions and the virtues, culminating in her book *On Virtue Ethics*.

Richard Joyce was born in England and raised in New Zealand. After studying at the University of Auckland he went on to graduate school at Princeton University. He has been a lecturer at the University of Sheffield and a Research Fellow at the Australian National University. He is presently a Research Fellow in the Philosophy Department at the University of Sydney. His principal areas of research are meta-ethics and moral psychology, though he is also interested in practical reason, philosophy of biology, and the emotions. He is author of *The Myth of Morality* (CUP, 2001) and *The Evolution of Morality* (MIT Press, 2006).

Norva Y. S. Lo was born in Hong-Kong and received her Ph.D. in Philosophy from the University of Western Australia in 2002. After two years as a post-doctoral fellow she became a lecturer in the Department of Philosophy, La Trobe University, Melbourne, Australia, in 2006. She has published in *Inquiry*, *Proceedings of Aristotelian Society*, *Environmental Ethics*, *Environmental Values*, and the *Stanford Encyclopedia of Philosophy*. Her research interests include: Moral Philosophy, Experimental Philosophy, Environmental Philosophy, and David Hume.

Graham Oddie was born and raised in Timaru, in Aotearoa (New Zealand). After stints as a shepherd and a truck driver, he left for the University of Otago where he was entranced by Pavel Tichy's demolition of Popper on truthlikeness. He wrote a PhD on verisimilitude at University of London (LSE) and which morphed into his first book *Likeness to Truth* (Reidel, 1986). He has taught at Otago (as lecturer) Massey (as professor) and is currently Professor of Philosophy (and Associate Dean of Humanities) at the University of Colorado at Boulder. He is interested in metaphysics and value theory with the odd foray into the philosophy of science and philosophical logic. He is the author of *Value Reality and Desire* (Oxford University Press, 2005).

Herlinde Pauer-Studer, is Associate Professor for Philosophy at the University of Vienna. She is the author of *Das Andere der Gerechtigkeit* (Akademie Verlag 1996); *Autonom Leben* (Suhrkamp 2000); *Einführung in die Ethik* (UTB 2003), *Commentary on David Hume, Of Morals* (Suhrkamp 2007).

Charles Pigden was born in England and studied philosophy at Cambridge, before going on to do a PhD at La Trobe University, Melbourne, Australia. Since 1988 he has taught philosophy at Otago University, Dunedin, New Zealand. He has published on a wide range of topics from conspiracy theories to the reality of numbers, but, if pressed, will admit to being a meta-ethicist with special interests in Russell, Moore and Hume. He is one of the very few academics to have published a philosophical dialogue in blank verse ('Complots of Mischief').

Luke Russell completed both a BA and PhD at the University of Sydney, where he is now Lecturer in the Department of Philosophy. He has published papers on various topics in ethics, including normativity, virtue and evil.

Constantine Sandis is a lecturer in Philosophy at Oxford Brookes University and NYU in London. He works primarily on the philosophy of action and its explanation and is the author of numerous related articles and a forthcoming book, *The Things We do and Why We Do Them* (Palgrave Macmillan). He is the editor of *New Essays on the Explanation of Action* (Palgrave Macmillan) and is currently co-editing (with Timothy O' Connor) *A Companion to the Philosophy of Action* (Wiley-Blackwell)

Michael Smith is Professor of Philosophy at Princeton University. His books include *The Moral Problem* (1994), *Mind, Morality, and Explanation: Selected Collaborations* (2004) (with Frank Jackson and Philip Pettit), and *Ethics and the A Priori: Selected Essays on Moral Psychology and Meta-Ethics* (2004).

Christine Swanton is a professor in the philosophy department at the University of Auckland. Her book *Virtue Ethics: A Pluralistic View* is in paperback, Oxford University Press 2005. She now works on virtue ethics (including the ethics of Nietzsche and Hume) role ethics, moral particularism, and right action. She has published widely in these areas.

A Note on References to Hume and Locke

Generally we employ the name/date system for references, but with Hume himself and one or two others we employ a minor variant of the *Hume Studies* system, now widely used by Hume scholars. With Hume's *Treatise of Human Nature*, the letter T is followed by the book, part, section and paragraph numbers with a page reference to the famous Selby-Bigge/Nidditch edition of 1978 following a forward slash. Thus the reference for the famous is/ought paragraph is T, 3.1.1.27/469-70.

References to the *Abstract* will be by paragraph number followed by a forward slash and page reference to the reprint at the back of Selby-Bigge/Nidditch. With the *Enquiry Concerning Human Nature* and the *Enquiry Concerning the Principles of Morals* the letters EHU or EPM will be followed by the section (part) and paragraph numbers, with a page reference to the Selby-Bigge/Nidditch edition of 1975 following a forward slash. The appendices to the EPM will be referred to as EPM, App 1 etc and the *Dialogue* as EPM, *Dialogue*, with paragraph numbers and page references to Selby-Bigge/Nidditch. Thus the reference for the famous 'consign them then to the flames' passage is ECU, 12.3.34/165; the reference for 'Have the gods forbid self-murder?' is EPM, *Dialogue*, 35/335. Since most modern editions of Hume's works have either part, section and paragraph numbers or the Selby-Bigge/Nidditch page numbers in the margins, I hope this will make things easy for readers.

References to the *Dialogues Concerning Natural Religion (Dialogues)* and the *Natural History of Religion (NHR)* follow the same format with page references to Gaskin's Oxford Classics edition of 1993, since Oxford is the only publisher with the wit to serve up *both* of Hume's principle works on religion in the one volume (though unfortunately the reader must put in the paragraph numbers herself).

References to *Letters* are to the two volume 1932 edition of Hume's letters, edited by Greig, by volume, letter and page number (following the forward slash).

References to Hume's *Essays* are to Miller's revised edition of 1987, published by the Liberty Fund, by part, essay and, where appropriate, page number, following the forward slash.

For Hume's *History of England* we employ W.D Todd's excellent Liberty Fund edition of 1983 with the format *History*, 1.3/137 (volume I, chapter 3, page 137.)

For the *Dissertation on the Passions* we use the Clarendon edition of 2007 edited by Beauchamp.

References to Locke's *Essay Concerning Humane Understanding* are to the 1975 Clarendon edition, edited by Nidditch, and follow the rubric *Essay*, 4.3.18/549 (Book 4, Chapter 3, paragraph 18/page 549).

Introduction

1. Prologue

We come to Hume to argue, not to praise him
 For Hume's philosophy lives after him
 And since it's *not* interrèd with his bones
 Like dead men's thought that lives, it lives because
 We think that it contains important truths
 (Truths that illuminate the nature of
 Morality, the mind and reason too,
 The 'sentiments' that make us do our duty,
 The role of passion and the sense of beauty);
 Or think that Hume is wrong, but that his faults
 Are kissing cousins *to* important truths
 (So, loving truth, we love the truths they're like);
 Or think him wrong, but think that when he errs
 They're mighty errors, great and grand mistakes,
 That represent temptations of the mind
 Which must be wrestled with and overcome
 That we may learn to see these things aright.
 But if we seek to learn from David Hume,
 We cannot hope to learn without debate,
 His ideas must be tested in the fire
 Of criticism if we are to find
 The philosophic gold that lies within.
 For philosophic truth is our concern
 Above all else; less so, historic facts
 About a thinker's thought, however great.
 For though we seek to know great David's mind,
 We want to know what David thought so we
 Can find if we agree or disagree
 Or better still, if David's wrong or right.
 But right or wrong, we hope that Hume repays

The homage that *we* pay – attention first,
 ('What does he say? What does he mean by it?
 What was the context of that barbed remark?
 Who were his targets? And who influenced him?
 What larger ends did David have in view
 When arguing for this doctrine or for that?')
 Next comes *analysis* ('Exactly how
 Is this specific inference meant to work?
 Does Hume appeal to missing premises,
 Obvious to him but not so much to us?')
 Next comes *critique* ('Well then, is David right?
 Does this conclusion follow from that claim?
 And what about these claims - are they all true?')
 Sometimes that's it, since Hume, it seems, is right,
 ('So, Hume's correct and reason's passion's slave!')
 Sometimes *amendment* follows ('It is clear,
 That in this case the argument is lame,
 But nonetheless another argument
 Can be constructed which would prove the point
 From something like great David's premises.
 "No Ought From Is" – amended - we can prove,
 Employing logic quite unknown to Hume,
 So he's not right, but yet not wholly wrong.')
 Sometimes what follows is an inference:
 ('It seems that in this instance Hume's correct -
 At least if he's corrected he's correct –
 But what then follows, what does this suggest
 About the status of our moral claims?')
 Sometimes amendment seems impossible
 But still we learn by seeing why that's so
 ('So David's wrong - the argument's a dud,
 And rational reconstruction is in vain,
 But if he's wrong and can't be put to rights
 Still, Hume's missteps can lead us to the light,

For from his errors we can lessons take -
This is what we've learned from his mistake ...’).

History is a necromantic art
 At least when history serves philosophy;
 We aim to resurrect the mighty dead,
 To grapple with their thought and thus to learn,
 Contesting boldly their contentious claims,
 Whether to lose the contest or to win.
 No Faustus we, whose Mephistopheles
 Could only summon silent ghosts, for we
 Must needs have ghosts that speak, since it's their thoughts
 That interest us. We need no Helen dumb
 To make us all immortal with a kiss -
 Immortal thoughts suffice, but since those thoughts
 Still live, although the thinker is long dead,
 Because they live, they move, migrate and change
 Thus undergoing metamorphosis.
 The thoughts which have an influence on men's minds
 (And women's too – let's not be sexist here!)
 Allegedly from some philosopher
 May not be quite the thoughts that stirred his brain.
 Thus 'Humean' doctrines, 'Humean' arguments,
 That dominate the philosophic scene,
 Are sometimes not the thoughts of David Hume -
 That does not mean they should not be discussed,
 For thoughts, mutating, may perhaps become
 More like the truth, or if mistaken still,
 More like the happy errors from the which
 We learn the most by thinking why they're wrong.
 Is Reason passion's slave in David's sense?
 Perhaps not – but perhaps in Michael Smith's -
 As true philosophers, we'd like to know!
 Are moral judgments meant to be truth-apt,

The sorts of things that can be true or false?
 So David thought (or so I think he thought)
 But that is not the view of many now
 Who think Hume's arguments prove otherwise.
 If arguments derived from David Hume
 Can prove what Hume perhaps did not believe,
 That may still be a matter of some pith
 And moment that deserves to be discussed.
 Ideas, mutating, sometimes are improved
 And even if the change is for the worse
 May still acquire a major influence,
 And consequently merit some debate.
 This is a book addressed to '*Humean*' themes;
 Sometimes they're themes direct from David's brain,
His arguments or theses *he* advanced;
 Sometimes they're themes *derived* from David's thought -
 Arguments *based* on arguments of his,
 Theses *quite like* the theses he advanced,
 Sometimes they're *consequences* of his thought,
 Deductions he may not have followed through;
 Sometimes we can't be sure, when David's texts
 Bear two interpretations (if not more!)
 And we cannot be certain which he meant;
 But whether they're his thoughts unmodified
 Or thoughts derived at some removes from his
 They're thoughts which dominate today's debates.
 There few philosophers are cited more.
 At least in ethics; for there's scarce a book
 In ethics, meta-ethics, action theory
 Or what is called 'moral psychology'
 That does not doff its hat to David Hume.
 That's our excuse (if an excuse is needed)
 For turning once again to his ideas.
 There's gold in Humean hills or thereabouts

And these are our attempts to dig it out.

O Hume, great David, thou art mighty yet!
 Thy spirit walks abroad; now thy thoughts thrive
 They live more now than when thou wast alive!

2. Three Themes and a Conference

This book deals with three themes from Hume's moral philosophy which loom large in contemporary ethical thought:

- (1) Hume's famous argument that 'Reason is, and ought only to be, the slave of the passions, and can never pretend to any other office than to serve and obey them' (T, 2.3.3/413-418);
- (2) the Motivation (or Influence) Argument (T, 3.1.1.5-10/456-458) which purports to prove that 'since morals ... have an influence on the actions and affections' and since 'reason alone [as Hume claims to have already proved] can never have any such influence', 'it follows that [moral distinctions] cannot be derived from reason;
- and
- (3) Hume as a virtue theorist.

The various papers deal with Hume himself and with 'matters arising' from the Humean agenda. The book is loosely based on a mini-conference held at the University of Otago in January 2003, *Hume, Motivation, 'Is' and 'Ought'*; 'loosely' because the collection includes several papers *not* presented at the conference and excludes some which *were*. Most of the excluded papers - the No-Ought-From-Is papers - have been hived off in a separate collection, *Hume on 'Is' and 'Ought'* (Pigden, 2009b). This is partly a matter of convenience and partly a matter of conviction. For it is my editorial belief that the Motivation Argument has a lot less to do with the famous No-Ought-From-Is paragraph (T, 3.1.1.27/469-470) than is commonly supposed. Of course they both appear in the same section of Hume's *Treatise* (3.1.1) and are both designed to prove (or to help to prove) that 'moral distinctions are not derived from reason' (whatever precisely *that* means). But they are distinct arguments and the one does not depend on the other. But this is far from being the conventional wisdom amongst

philosophers, so before discussing what this book *is* about I am going to devote a couple of pages to what it is *not* about to get the issue out of the way. I shall have to be dogmatic as I aim to be brief. Readers who want the gory details can turn to the other collection.

3. Clearing the Decks: No-Ought-From-Is

‘In every system of morality which I have hitherto met with’, says Hume, ‘the author proceeds for some time in the ordinary way of reasoning, and establishes the being of a God, or makes observations concerning human affairs; when of a sudden I am surprized to find, that instead of the usual copulations of propositions, *is*, and *is not*, I meet with no proposition that is not connected with an *ought*, or an *ought not*... But as this *ought*, or *ought not*, expresses some new relation or affirmation ...[it] seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it’. And the implication would appear to be that this not only *seems* inconceivable but actually *is* inconceivable, and that you cannot deduce moral conclusions from non-moral premises.

Note that Hume is speaking the language of logic. In every other instance, when Hume talks of ‘copulation’ it is sex that he has in mind. Here alone he is talking about the ‘copula’, a technical term in the logic of the day which was seldom employed in any other connection. This suggests that in the No-Ought-From-Is passage, Hume is making a logical point. Furthermore he is appealing to a commonplace of Eighteenth Century logical theory - that the conclusion of a valid inference is contained in the premises. So if there is new matter in the conclusion that is *not* contained in the premises – namely the copulations of propositions *ought* and *ought not* – then the inference cannot be ‘just and true’, i.e. logically valid. Notice that all that is required for this point to hold (at least so far as Eighteenth Century logical theory is concerned) is a *prima facie* distinction between moral words and others. There need be no deep difference between semantic kinds.

And this is fortunate since Hume probably did not believe in such a difference. After all, he explicitly compares vice and virtue ‘to sounds, colours, heat, and cold’, (T, 3.1.1.26/469) which, suggests that in saying ‘X is virtuous’ we are saying something similar to ‘X is red’. This is confirmed by a famous passage in the *EPM*, App.1.10/289: ‘The hypothesis which we embrace is plain. It maintains, that morality is determined by sentiment. It defines virtue to be *whatever mental action or quality*

*gives to a spectator the pleasing sentiment of approbation; and vice the contrary. We then proceed to examine a plain matter of fact, to wit, what actions have this influence]'. Here Hume gives an *analysis* of 'virtue' as a causal concept, *defining* it via its causal role, with a view to establishing, by empirical research, that *virtues* (mental actions or qualities which give to a [suitably qualified] spectator the pleasing sentiment of approbation) are also qualities '*useful or agreeable to the person himself, or to others*' (EMP, 9.1.1/268). Thus Hume seems to think (to use post-Humean terminology) that it is *analytic* that a virtue is a quality that arouses in a (suitably qualified) spectator the pleasing sentiment of approbation, and furthermore he *needs* it to be analytic since it hard to see how he could establish the point by empirical enquiry. Rather he has to take it as read before the empirical enquiry into 'what actions have this influence' (that is, of exciting approbation) can get off the ground. Indeed, following his sometime philosophical hero Hutcheson [Hutcheson, 2002, p. 146], Hume actually *defines* the word 'obligation', if not the word 'ought': 'All morality depends upon our sentiments; and when any action or quality of the mind pleases us *after a certain manner* we say it is virtuous; and when the neglect or nonperformance of it displeases us *after a like manner*, we say that we lie under an obligation to perform it' (T, 3.2.5.4/517). Thus to say that A ought to do B is to say that if A does not do B, this act of omission would excite the sentiment of disapprobation in a suitably qualified spectator.*

The upshot is that in the Is/Ought passage, Hume is probably *not* trying to argue that non-moral premises do not *entail* moral conclusions, where a set of premises A *entails* a conclusion B if B can be logically derived from A with the aid of (perhaps) unstated analytic bridge principles. Thus No-Ought-From-Is is not a thesis about entailment at all, and neither implies nor presupposes a fundamental distinction between the moral and the non-moral. In particular, it neither implies nor presupposes any kind of non-cognitivism.

So what is the point of it then? Well, Hume's over-arching aim in *Treatise*, 3.1.1, is to argue that moral distinctions are not derived from reason in *any* sense. But to prove this larger claim he has to prove the subsidiary thesis, that the basic principles of morality cannot be demonstrated. That they *could* be so demonstrated was a popular view in the late Seventeenth and early Eighteenth Centuries. Now for moral truths to be demonstrable they would have to be deduced from self-evident truths. By the end of 3.1.1, Hume has already proved, to his own satisfaction at any rate, that there are no

self-evident *moral* truths. This still leaves open the possibility that the truths of ethics might be derived from self-evident truths of some other kind, a possibility suggested by Locke, *Essay*, 4.3.1.18, p. 549. The point of No-Ought-From-Is is to foreclose this option. If there are no self-evident moral truths and if moral conclusions cannot be deduced from non-moral premises, then moral truths are not demonstrable, whatever Locke, Clarke and Spinoza may have to say to the contrary.

4. Reason, the Slave of the Passions

So much for what this book is *not* about, now for the main themes. One of the most talked about sections in Hume's *Treatise* is 2.3.3 'Of the Influencing Motives of the Will'. This is the locus for Hume's famous claim 'that reason alone can never be a motive to any action of the will' (T, 2.3.3.1/413) or (more picturesquely) that 'Reason is, and ought only to be, the slave of the passions, and can never pretend to any other office than to serve and obey them' (T, 2.3.3.1/414). What does he mean by this claim? What are his arguments? What part does this thesis play in his overall polemic against moral rationalism? And what is its relevance to current concerns?

4.1. The Slavery of Reason Thesis: Meaning

Let's start with meaning. What Hume means by 'reason alone can never produce any action, or give rise to volition' depends on what he means by 'reason' and 'alone'. With respect to 'reason' there are three possibilities:

[1] 'Reason' either stands for our belief-forming faculty or is a collective name for our belief-forming faculties taken together (especially when they are properly exercised or are operating without error). Thus the Slavery of Reason Thesis amounts to the claim that '*beliefs* by themselves can never produce any action or volition' or that '*beliefs* alone cannot motivate'.

[2] Reason is a sub-faculty of [1], namely the intellectual organ by which we *infer* new beliefs (whether the inference is inductive or deductive). Thus the Slavery of Reason Thesis amounts to the claim that '*inferred* beliefs by themselves can never produce any action or volition' or that '*inferred* beliefs alone cannot motivate'.

[3] Reason is a sub-faculty of [2], namely the intellectual organ which derives new beliefs from old beliefs, or, at least, that derives new beliefs

from *ideas* rather than *impressions* (where ideas are copies or combinations of copies of prior sensations or passions). Thus the Slavery of Reason Thesis amounts to the claim that ‘*beliefs inferred from ideas* can never produce any action or volition by themselves’ or that ‘*beliefs inferred from ideas* cannot motivate alone’.

There is textual evidence for all three interpretations. For the moment we can leave the matter open.

What about ‘alone’? Here there is a lot less ambiguity. When Hume says that ‘reason alone can never produce any action, or give rise to volition’, what he means is that reason-derived beliefs cannot produce an action or a new desire *without the aid of a pre-existing passion*. Passions - that is pre-existing passions - are necessary if the products of reason are to result in action. As Rawls puts in his *Lectures on Moral Philosophy*, ‘the passions specifying the final ends in deliberation ... must be passions we have and may be moved by *now*, at the time of deliberation and before we carry out the action’. This means that ‘the bare knowledge of our future passions does not move us now, unless that knowledge guides, or connects with, some passion that we have now’ (Rawls, 2000, p. 34). The belief that unless I start saving *now* I am likely to have a bad time in my old age (at which point I will passionately want to have more money) isn’t going to move me unless I currently care about my future self. I must *now* want that my future wants be satisfied, if the knowledge of those wants is to move me to action in the present. Luckily, most of us are in fact endowed with a ‘general appetite to good, and aversion to evil’ (T, 2.3.3.8/417), a desire for our long term survival and satisfaction, which means that reason-derived beliefs about our future wants sometimes translate into immediate action. But this a basic instinct: it is not a want that is derived from reason but an appetite that we must presuppose if reason-based beliefs about our future wants are to have any effect on our current conduct. A reason-based belief, even a belief *about* future passions, can neither prompt us to action or create new desires without the aid of a passion that is already in existence.

But what kind of impossibility are we talking about here? For Hume, I think, the impossibility has to be *causal*. It can hardly be conceptual since it is part of his official theory that it is conceptually possible for anything to cause anything. So the idea is that it is *causally* impossible (given the basic structure of the human mind) for a reason-based belief to produce an action or a volition without the cooperation of a

passion.

But what does Hume mean by a ‘passion’? That is a long and tangled tale. Fortunately for us, the details of Hume’s theory are largely irrelevant to the concerns of this collection. For Hume’s latter-day disciples tend to equate passions with desires, and when Hume himself talks about the passions in this connection, it is often desires that he is thinking of. Reason produces an action or gives rise to a new volition (that is a movement of the will such as a new desire) by pointing the way to realize an end set by passion. Thus passions for Hume are end-setting entities. And the obvious example of an end-setting mental entity is a desire or a want.

There are two basic ways that reason-based beliefs can cause a new desire or action in tandem with a pre-existing desire: 1) The belief can suggest the means to a desired end; or 2) the belief can suggest a constituent or a specification of the desired end. I want to get to work without increasing my carbon footprint; reason tells me that I can do this by going by bike; and the want in tandem with the belief causes a new desire cycle to work. I want to have a good time tonight and reason informs me that seeing *the Dark Knight* conforms to my idea of a good time. But seeing *the Dark Knight* is not a *means* to having a good time tonight – rather my good time *consists* in seeing the *The Dark Knight*.

But although, for Hume, all desires are passions, some passions are not desires. Consider, for example ‘kindness to children’ which is certainly a widespread tendency. (T, 2.3.3.8/417.) Is this a desire? Not if desires are propositional attitudes, that is to say attitudes towards *propositions* (that of wanting them to be so). For being kind to children is not a matter of wanting something to be the case. There are no doubt people (such as *Save the Children* activists) who want everything to be well with the world’s children. But with most of us, our kindness to children, in so far as it exists, is not like that. It is a ‘tendency’ rather than a desire, a disposition to *acquire* desires for the welfare of the individual children, not a generalized desire that children should be happy. Kindness to children is not the only motivational state that falls into this category. If I have a taste for chocolate the desires that I feel for Mars bars and Kitkats are not desires for the means to achieve some larger end (such as keeping up a daily intake of chocolate). Rather I am unusually prone to *acquire* desires for chocolate, especially when it stimulates the senses. An amorous person is not (or need not be) someone with the Valmont-like project or maximizing his or her sexual conquests. Rather he or she is unusually prone to sexual desire, that is, unusually disposed to

acquire desires for sexual relations with the attractive people that they happen to meet. Even ‘curiosity or the love of truth’ would appear to fall into this category. When people complain that George Bush lacks curiosity, they don’t mean that he lacks the desire to approximate omniscience. They mean that when exposed to information about this or that (such as the existence of foreign countries or the workings of the economic system) Bush is not disposed to acquire desires to know more. So although some passions are desires, some are dispositions to acquire desires (or DTADs), a point not often noticed by Hume’s latter-day disciples. Indeed, it is not clear that Hume himself realized that he was postulating DTADs as well as desires, and that the two are distinct. For once we allow for DTADs as well as desires, the Motivation Argument begins to look distinctly shaky (see Essays 3 and 4).

Hume is careful to stress that the passions which move us are often ‘calm’, and that a passion can exercise a powerful causal influence even though it is principally known from its effects and does not excite much ‘emotion’ in the mind. Thus for Hume a strong, or causally influential passion, (such as the desire to lead a long and healthy life) can be almost invisible from a phenomenological point of view, whilst a passion which excites a psychic stir (such as the craving for just one more cigarette) can be relatively weak in its effects. It is this that leads the reformed smoker to believe that when she forgoes the cigarette she is actuated by reason alone, whereas she is in fact motivated by the reason-based belief that smoking can kill *plus* the calm desire to live long and prosper.

Thus the Slavery of Reason Thesis amounts to this: given the constitution of the human mind, beliefs, inferred beliefs or beliefs derived from ideas, are causally incapable of producing either actions or new desires without the aid of pre-existing passions, which passions themselves are usually to be understood *either* as desires *or* as dispositions to acquire desires (DTADs).

That’s the doctrine: what are the arguments?

4.2 The Slavery of Reason Thesis: the Arguments

Hume’s chief argument for the Slavery of Reason Thesis is a simple one. Its argument can be summarized thus:

Slavery of Reason Argument: Non-Committal Version.

- i) Beliefs derived from reason are either analytic (to do with the relations of ideas) or synthetic (to do with causal relations). [‘The understanding <that is, *reason*> exerts itself after two different ways, as it judges from demonstration or probability; as it regards the abstract relations of our ideas, or those relations of objects, of which experience only gives us information’.]
- ii) Analytic beliefs cannot motivate except in so far as they lead to synthetic beliefs. [‘Abstract or demonstrative reasoning, therefore, never influences any of our actions, but only as it directs our judgment concerning causes and effects’.]
- iii) Synthetic/causal beliefs cannot motivate if the causes and effects believed in are indifferent to us. [‘It can never in the least concern us to know, that such objects are causes, and such others effects, if both the causes and effects be indifferent to us.’]
- iv) Hence synthetic/causal beliefs cannot motivate unless they point the way to realize some pre-existing passion (from iii). [‘Tis evident in this case that the impulse arises not from reason, but is only directed by it.’]
- v) Beliefs derived from reason cannot motivate by themselves, that is, without the aid of some pre-existing passion (from i), ii) & iv)). [‘Reason alone can never be a motive to any action of the will.’ ‘Reason is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them.’]

This is a carefully non-committal exposition of the argument. The conclusion is *that beliefs derived from reason* cannot motivate by themselves, but whether this means that *all* beliefs, *inferred* beliefs or beliefs *derived from ideas* cannot motivate is simply left open for further debate.

Hume’s fundamental thought is this. Given the basic structure of the human psyche it is causally possible for a human being to entertain any reason-based belief (and any conjunction of reason-based beliefs) whilst remaining psychologically unmoved. The belief, by itself, is causally inert. So, if a reason-based belief is to cause

an action or volition, it can only do so with the aid of some *other* psychological entity. Another reason-based belief could not do the trick. For it is psychologically possible (given the basic structure of the human mind) to remain unmotivated by the conjunction of the first reason-based belief with the second. So this extra cause must be a psychological entity of some other kind. But any entity that *could* cause either an action or a new desire in conjunction with a reason-based belief would fit the functional profile of a passion. Hence, if a belief is to motivate, a passion of some kind is causally necessary.

So much for the thesis that ‘reason alone can never be a *motive* to any action of the will’. But what about the corresponding claim that reason ‘can never oppose passion in the direction of the will’? Hume insists that the second thesis follows from the first by a necessary consequence. (T. 2.3.3.4/414-415.) But how is this supposed to work? The argument, I take it this:

- 1) The final determinants of the will are desires (which are a kind of passion). [Premise.]
- 2) Nothing can conflict with a desire Φ except a) another desire Ψ or b) a belief Γ negating a belief Δ on which the desire Φ causally depends. [Premise.]
- 3) The products of reason are all beliefs. [Premise.]
- 4) Product-of-reason beliefs cannot give rise to new desires by themselves (that is, without the aid of a pre-existing passion). [Premise: a variant of conclusion v) above.]
- 5) Hence the products of reason by themselves cannot give rise to anything that conflicts with an existing desire Φ , except a belief (or beliefs) Γ negating a belief (or beliefs) Δ on which the desire Φ causally depends [From 2), 3) and 4).]
- 6) Hence reason (by itself) cannot give rise to anything that conflicts with an existing desire except a belief (or beliefs) Γ negating a belief (or beliefs) Δ on which the desire Φ causally depends [From 5.]
- 7) Hence reason (by itself) cannot give rise to anything that conflicts with an existing desire (or passion) Φ in determining (that is, directing) the will *except* a belief (or beliefs) Γ negating a belief (or beliefs) Δ on

which the desire Φ causally depends. [From 1) and 6).]

With these arguments in place, and reason defined as either our belief-forming faculty or a sub-faculty of that faculty, Hume is in a position to state his rather blood-curdling conclusion:

'Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my finger [unless that preference is based on a false belief]. 'Tis not contrary to reason for me to chuse my total ruin, to prevent the least uneasiness of an Indian or person wholly unknown to me [Ditto]. 'Tis as little contrary to reason to prefer even my own acknowledg'd lesser good to my greater, and have a more ardent affection for the former than the latter (T, 2.3.3.6/416.)

4.3. The SRT and the Polemic Against Rationalism

What is the point of the SRT? Well, as we shall see, it provides one the premises for the Motivation (or Influence) Argument which is the chief weapon in Hume's anti-rationalist armory in *Treatise* 3.1.1. (Moral Distinctions not deriv'd from Reason'). But there is a bit more to it than that. For there are remarkable affinities between *Treatise* 2.3.3, and Section I of Hutcheson's *Illustrations on the Moral Sense* (Hutcheson, 2002, pp. 137-155). But there is no trace of the Motivation Argument in Hutcheson, presumably because he did not believe the second premise, that moral beliefs motivate. (See Darwall, 1995, ch. 8.) Thus for Hutcheson at least, and presumably for Hume, the SRT has an independent part to play in the refutation of rationalism. In my view, its role is to counter one of the rationalists' chief arguments *against* sentimentalism.

The sentimentalists - Hutcheson, Hume - construed the moral properties as akin to secondary qualities. Slurring over some important complications, their view was that to say an action is *right* or *virtuous* is to say the it would arouse the approbation of a suitably qualified *human* spectator, someone impartial, relevantly informed and dispassionate (that is, devoid of any distorting hang-ups). The rationalists did not like this view. Here, for instance, is John Balguy: 'if God had not framed our natures with such a propensity [i.e. the moral sense], and given us this benevolent instinct, we should have been altogether incapable of virtue,

notwithstanding intelligence, reason and liberty.’ Balguy finds this incredible. ‘How a notion can be true that labours under such a consequence as this, I cannot understand.’ Surely if we are rational we can recognize and be moved by the right whatever our instinctual inclinations. ‘Supposing us void of natural compassion as well as benevolence; might we not possibly be induced to attempt the relief of some person in distress, merely from the reason of the thing? ... In short, if we made use of our understandings, they would not fail, I think, to discover our duty in such a case. Nay they would prompt us to undertake it, and condemn us if we omitted it.’ (Balguy, 1733, pp. 390-391.) His argument, I take it, goes something like this:

1R) It is a conceptual truth that the moral facts, if any, would have to be both accessible to reason and (in today’s terminology) *objectively prescriptive* (that is, such as to defeasibly motivate any conceivable rational agent that became aware of them of them, whatever that agent’s desires or inclinations).

2R) Facts about would what arouse the approbation or disapprobation of a suitably qualified human observer are *not* accessible to reason (at least not in the right kind of way) and are *not* objectively prescriptive (that is, such as to defeasibly motivate any conceivable rational agent that became aware of them, whatever that agent’s desires or inclinations).

3R) So the moral facts, if any, are not facts about would what arouse the approbation or disapprobation of a suitably qualified human observer.

The sentimentalists replied, or can be reconstructed as replying, as follows:

1S) There are no objectively prescriptive facts (facts that would be defeasibly motivating to any conceivable rational agent that became aware of them, whatever that agent’s desires or inclinations).

2S) So if there are to be any moral facts, it is *not* a conceptual truth that the moral facts, if any, would have to be objectively prescriptive (that is, such as to defeasibly motivate any rational agent that became aware of them). They may be facts that are only motivating to human beings or creatures with similar dispositions (and not necessarily to all of them).

The point of the SRT is that it provides the backing for the sentimentalists' premise (1S). Reason is either our belief-forming faculty or a sub-faculty of that faculty. Beliefs derived from reason cannot motivate without the aid of a pre-existing passion, either a DTAD or a desire. Hence *facts* accessible to reason cannot motivate *human* beings without the aid of a pre-existing desire or DTAD, since to be aware of a fact is, in part, to believe in it. No desires or DTADs are *constitutive* of rationality, except, perhaps, a penchant for true, as opposed to false, beliefs and a taste for the right kind of belief-forming strategies. ('Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my finger' [T, 2.3.3.6/416.] but 'a wise [and hence, presumably, a rational] man ... proportions his belief to the evidence' [EHU, 10.1.4/110].) Indeed, this pretty much follows from the definition of Reason as either our belief-forming faculty or a sub-faculty of that faculty. But if facts accessible to reason cannot even motivate *human* agents without the aid of a pre-existing desire or DTAD, and if no morally salient desires or DTADs are constitutive of rationality, then there are no objectively prescriptive facts (facts that would be defeasibly motivating to *any* conceivable rational agent that became aware of them whatever the agent's desires or inclinations).

The upshot for Hume is that morality is rationally optional. Moral considerations will appeal to those with the right kind of psychological make-up but not to those without. Such people may be repellent but they need not be irrational. The point becomes plain in his famous discussion of the 'sensible knave' a sort of mild and discrete Thrasymachus (perhaps like Mr Elliot in Jane Austen's *Persuasion*) who remains unmoved by considerations of justice, except in so far as they serve his self-interest: 'A sensible knave ... may think, that *honesty is the best policy*, [is] a good general rule; but [that it] is liable to many exceptions: And [that] he conducts himself with most wisdom, who observes the general rule, and takes advantage of all the exceptions.' (EPM, 9.2.22/282-283.) Hume admits that he has no rational answer to the knave: 'If a man think, that this reasoning much requires an answer, it will be a little difficult to find any, which will to him appear satisfactory and convincing. If his heart rebel not against such pernicious maxims, if he feel no reluctance to the thoughts of villainy or baseness, he has indeed lost a considerable motive to virtue; and we may expect, that his practice will be answerable to his speculation.' (EPM, 9.2.23/283.) The only consolation that Hume seems to suggest is that such sensible knaves are relatively rare, and that many of them are not as sensible as they like to suppose, and

get caught out because of over-confidence.

4.4. The SRT and Modern Moral Philosophy

So much for Hume, but what about us? Why is the SRT important for modern moral philosophy?

1) The SRT provides one of the main premises for the Motivation Argument which is taken by many to be the chief argument for non-cognitivism or expressivism. (Non-cognitivists think that, strictly speaking, moral judgments are neither true nor false since their function is to convey commands or express attitudes. Expressivists think that although the primary function of moral judgments is to convey commands or express attitudes, this is compatible with the capacity for a rather watered down kind of truth. Thus expressivism is non-cognitivism for wimps.)

2) The SRT - or at least certain versions of the SRT - poses a problem for modern moral rationalists for much the same reasons that the original SRT posed a problem for their 18th Century counterparts. If you want to base morality on facts which are objectively prescriptive – either for rational beings generally or for human beings in particular – then the SRT stands in your way since it seems to show that there can be no such facts. Suppose (A) that beliefs cannot motivate without the aid of a pre-existing desire or DTAD. And suppose (B) that there are no morally salient desires or DTADs that are constitutive of rationality. Then for any fact there is a rational being that can remain unmoved by it without prejudice to its rationality (except for the motivation to draw certain inferences in accordance with the agent's belief-forming propensities). Hence there are no facts, either accessible to, or constructible by, reason, such that they would be defeasibly motivating to any conceivable rational agent that became aware of them, whatever that agent's desires or inclinations. Thus a philosopher who wants to formulate an objectively prescriptive ethic has three options (though they are not mutually exclusive): She can deny (A) that beliefs cannot motivate without the aid of a pre-existing desire or DTAD; she can deny (B) that there are no morally salient desires or DTADs that are constitutive of rationality; or she can insist that there is a special kind of awareness of the moral facts that *does* guarantee defeasible motivation but that this awareness transcends mere belief. Thomas Nagel in *The Possibility of Altruism* (1970) adopts the first strategy. For him, if I do X in the belief that X leads to Y, it follows that I have a desire for Y, but this desire need be

nothing more than a conceptual shadow, cast by the fact that I did X in the belief that X leads to Y. (See Pigden, 2009a, for a critique of this argument.) Korsgaard as a Kantian also adopts the first strategy (See Korsgaard, 1986, 1996*b*, 1996*d*). Michael Smith, by contrast, adopts the second, since he is a Humean about motivation though a Kantian in ethics. For him a desire to desire what all fully rational agents would desire, and a disposition to desire what the agent *believes* that all fully rational agents would desire are both constitutive of rationality (Smith, 1994). Korsgaard who is something of a motivational double-dipper, also seems to think that there are motivational dispositions that are constitutive of rationality. McDowell, in so far as I understand him, adopts *both* strategy 1 *and* strategy 3, since he endorses Nagel's view that desires need be nothing more than conceptual shadows cast by deliberate action whilst insisting that seeing things in a certain way can be sufficient for rational motivation (McDowell, 1978). Indeed, it would not be going too far to say that there is a whole philosophical industry of practical reasoners devoted to downplaying or denying the SRT in the interests of moral objectivity.

It might be thought that Hume and his disciples are committed to at least one action-oriented DTAD as constitutive of rationality, namely the disposition to acquire desires for (what you believe to be) the means to your ends. This is, so to speak, a function that takes us from desires and beliefs to new desires and subsequently actions, and it would be hard to recognize a being without it as any kind of rational agent. But if it would be irrational *not* to have this disposition, surely it is a disposition that we rationally *ought* to have? Thus Hume has been widely read as an instrumentalist about practical reason, subscribing to the norm that we rationally *ought* to be disposed to perform those actions which are (or are believed to be) the means to our desired ends. So although Hume believed that it is not contrary to reason to prefer the destruction of the whole world to the scratching of my finger, he is sometimes supposed to have thought that (given such a preference) it *would* be contrary to reason for me not push the red button if I correctly believed that by pushing the button I could I could destroy the world whilst saving the finger. Others, most notably Milgram, have argued that there is not much sign of any such norm in Hume, and that he seems to think that we *automatically* do the things that we believe to be conducive to our ends, in so far as we really desire the ends and in so far as we really believe. (Milgram, 1995.) I think that Milgram is correct about Hume and that this view can be vindicated in the light of a modern 'Humean' conception of beliefs and desires. For the very fact that we have

beliefs and desires in the first place guarantees that we will have the disposition to desire the means to our ends. And since we have this disposition *anyway* simply in virtue of having beliefs and desires, there is no place for a norm of rationality recommending rational agents to realize this disposition, where a norm is, minimally, something you might fail to meet. According to Kant it is analytic that whoever wills the end, wills also (so far as reason decides his conduct) the means' (Kant, 1994, 4:417), and though his reasons for saying this are somewhat obscure, I think he is basically correct. But this does not preclude a *causal* disposition to desire the means to our ends, internal to our natures as believing/desiring beings. Following Michael Smith (1994) I characterize a desire as a complex dispositional state grounding a range of counterfactuals. To desire a state of affairs X is to be defeasibly disposed to desire Y given the belief that Y leads to X, to desire Z given the belief that Z leads to Y and to be averse to W given the belief that W would hinder or prevent X. (The qualification 'defeasibly' is necessary since I may have *other* desires which are incompatible with the desire for the means to something that I genuinely desire. If I believe I cannot win my mistress's love without swimming the Hellespont but know that I am a poor swimmer who would die if I made the attempt, prudence - my 'general appetite to good, and aversion to evil' - is likely to trump my instrumental desire to attempt the swim.) To believe that Z leads to Y is to be in a complex dispositional state, grounding a related range of counterfactuals such as a disposition to desire Z given a desire for Y and to be averse to Z given an aversion to Y. If this is roughly right, it follows automatically that we tend to desire as a means what we believe will facilitate the things we desire as ends. If we are not even *defeasibly* disposed to desire the means to what we profess to desire as an end, that is evidence that we do not really desire it, or perhaps that we do not really believe that the supposed means will foster the end. It is analytic that *if* we have beliefs and desires we will have this disposition, but the disposition itself is grounded in the causal configuration of the human mind. But since we have a tendency to desire (and hence to choose) the means to our ends simply in virtue of having beliefs and desires in the first place, there is no need for a norm of practical reason telling us that we rationally ought to do this (See Finlay, Essay 7 for a detailed and sophisticated defense of this kind of view.)

Even philosophers with the modest ambition to ground morality in reasons for action are likely face difficulties with the SRT. A fact Δ constitutes a reason for an

agent Γ to φ if a belief in the fact Δ would motivate Γ to φ in virtue of Γ 's rationality. But if there are no desires or DTADs, and hence no ends, that are internal to rationality, then there are no reasons for action in this sense. Perhaps we can save the day by treating *being a reason for action* as a four-place rather than a three-place relation. A fact Δ constitutes a reason for an agent Γ to φ given ends Ψ if a belief in the fact Δ would motivate Γ to φ in virtue of Γ 's agency and Γ 's ends Ψ . This gives us a Humean conception of reasons for action but makes them unsuitable as the basis for ethics. If a fact Δ only constitutes a reason for an agent Γ to φ given Γ 's ends Ψ , then Δ 's being a reason for Γ to φ cannot be the basis of Γ 's obligation to φ since (generally speaking) that is supposed to hold irrespective of Γ 's ends. (You can't say that you are not obliged to do the right thing because you happen to have evil preferences.) Perhaps we can solve the problem by appealing to a norm of rationality which judges between ends. There is no doubt that we appeal to a variety of such norms in daily life. Some condemn actions which are self-destructive or which cut across the agent's long-term self-interest. The Common Law was long dominated by a norm which condemned suicide as irrational. If somebody killed himself this meant that 'the balance of his mind' must have been 'seriously disturbed'. Economists, at least when they are on duty, tend to subscribe to a self-interested norm of rationality which would condemn the choice of 'my total ruin' for the sake of 'an Indian or a person wholly unknown to me' as, at best, non-rational or a-rational, if not positively *irrational*. Thus economists tend to build a 'general [and selfish] appetite to good, and aversion to evil' into their conceptions of rationality. But even here there is considerable disagreement about such things as the discount rate for future pleasures and pains. On the other hand, we sometimes condemn as irrational actions which *don't* take other people into account. 'He is too dumb [that is, too irrational] to realize that the Party is more important than his big ego.' 'She's too stupid to realize that other people's feelings count.' So perhaps we can say that a fact Δ constitutes an R-reason for an agent Γ to φ if a belief in the fact Δ would motivate Γ to φ *if Γ were R-rational*, that is, if Γ conformed to some norm of rationality R. Indeed, if we could establish the *correctness* of a norm of rationality, we could simply say that a fact Δ constitutes a reason for an agent Γ to φ if a belief in the fact Δ would motivate Γ to φ if Γ were *really* rational, where to be really rational is to conform to the *correct* norm. The problem is that it is hard to see what could justify the claim that this or that norm of rationality is correct. Since we

can skip gaily from one norm to another in the course of a single conversation, we certainly cannot say that it is *analytic* that it either is or is not rational to be motivated by false but intelligently acquired beliefs, to discount the future at this rate or at that, to be motivated or not by other people's interests, or to be motivated or not by moral considerations. We commonly use 'rational' and its cognates in *all* these senses and more. But if it is not analytic, what is the *truthmaker* for the claim that a given norm of rationality R_1 is correct as opposed to one of its rivals R_2 , R_3 etc? It is hard to see how there could be a fact of the matter here. Hence there is something a bit naive about Nozick's book title *The Nature of Rationality*. There is not a thing, rationality, with a nature waiting to be investigated. Rather practical rationality, once it transcends the dispositions internal to being a believing/desiring agent, is a norm-relative affair. But if this is right, we cannot say that a fact Δ constitutes a reason for an agent Γ to φ but only that Δ constitutes a reason for an agent Γ to φ according to the norm of rationality R when it might *not* constitute a reason for Γ to φ according to some other norm. And that looks a bit too relativistic to afford a basis for ethics.

3) There is a third reason why the SRT is important for modern moral philosophy. It provides one of the premises for Mackie's well-known Argument from Queerness, (the AFQ) which is designed to support the 'error theory', the thesis that moral judgments are systematically false and morality a myth (though perhaps a necessary one). As I understand it, the AFQ takes a premise apiece from the arguments of the 18th Century rationalists such as Balguy (and I think Kant) and their sentimentalist opponents, Hutcheson and Hume:

AFQ^(CRP)

1R) It is a conceptual truth that the moral facts, if any, would have to be both accessible to reason and (in today's terminology) *objectively prescriptive* (that is, such as to defeasibly motivate any conceivable rational agent that became aware of them of them, whatever that agent's desires or inclinations).

S1) There are no objectively prescriptive facts (facts that would be motivating to *any* rational agent that became aware of them whatever the agent's desires or inclinations). [As noted above, the SRT provides the basis or this claim in both Hutcheson and Hume.]

M3) So there are no moral facts.

M4) Non-cognitivism is false. Moral judgments are truth-apt, true-or-false, in the true/false game. [Premise]

M5) For a substantive non-negative proposition to be true there must be a fact to make it true. [Premise]

M6) But since there are no moral facts to make them true, there are no *substantive* moral truths (what I elsewhere describe as non-negative atomic moral truths).

M7) Hence all systems of morality (and not just the ones I have ‘hitherto met with’) are systematically false. (Mackie, 1977. ch. 1.)

It is because modern rationalists are inclined to accept premises R1), M4) and M5) whilst rejecting the conclusion of the AFQ, that they are so bitterly opposed to the SRT.

It is, however, possible to accept 1R) and *some* versions of the SRT whilst denying S1). If you define a rational agent as an agent which meets a certain norm of rationality, and if this norm involves being motivated by certain ends or considerations - for example moral considerations - then there can be objectively prescriptive facts even if you accept principle (A), that beliefs cannot motivate without the aid of pre-existing desires or DTADs, and the further thesis that talking and thinking being are conceivable whose DTADs and desires would leave them unmoved by the moral facts. A talking mantis from Mars with a taste for human brains and no interest in the well-being of non-mantises, won’t count as rational whatever her intellectual attainments, since she won’t be motivated by the right kind of considerations. So you get objectively prescriptive facts in a sense. There will be facts that are motivating to any rational being that becomes aware of them, but only because we have restricted rationality to beings that happen to share our preferences. It seem to me that this strategy rescues R1) by trivializing it. The idea behind R1) is what might be called Aragorn’s Principle: ‘Good and ill have not changed since yesteryear; nor are they one thing among Elves and Dwarves and another among Men’. (Tolkien, 1995, III.ii, p. 428.) The idea is that morality is supposed to be binding on all talking, thinking and acting beings because of their capacity for thought, talk and action, not that it is only binding on beings that qualify as rational in some artificially restricted sense of the word.

Thus the SRT is important to modern moral philosophy for (at least!) three reasons: 1) it provides a premise for the Motivation Argument which seems to support non-cognitivism; 2) it poses a problem for philosophers who want to found ethics either on reasons for action or objectively prescriptive facts; and 3) it supports a premise of the AFQ which, in turn, supports the error theory.

4.5. Matters Arising

Essays 6-10 deal with matters arising from the Slavery of Reason Thesis.

Constantine Sandis (Essay 6) observes that motivation theorists are divided into two factions, those who think that we are motivated by reasons and that those reasons are partly constituted by beliefs *and* desires and those who think we are motivated by reasons that those reasons can be constituted by *pure* beliefs. In Sandis' opinion, Hume denied the assumption on which the two factions are agreed, that we are motivated by reasons. Sandis also develops a cognitivist interpretation of the Motivation Argument according to which it proves what Hume says he is trying to prove, namely that moral distinctions are not derived from reason.

Hume has sometimes been read as an instrumentalist about practical reason. Irrespective of our ends what we rationally ought to do is to perform those acts that will achieve our ends. Others have argued that Hume is not even an instrumentalist since he does not really believe that there is anything we rationally ought to do (Milgram 1995, and Korsgaard, 1986). This view is widely thought to be absurd, so absurd indeed, that it is difficult to believe that Hume himself believed it. Stephen Finlay (Essay 7) argues both that this *is* Hume's view and that Hume's view is correct. There is no plausible norm of practical reason commanding us to do what (we think) will achieve our ends that meets the minimum requirement for being a norm, namely that we might fail to meet it. Thus you cannot argue from the plausibility of such a norm, to other more sophisticated norms of practical reason that might be employed to underwrite morality.

In an elegant little paper Kent Hurtig (Essay 8) argues against a latter-day Humean, the late Bernard Williams. Williams is an internalist about practical reason. He thinks that X cannot be a reason for Γ to ϕ unless X relates to something in Γ 's 'subjective motivational set.' But unlike Hume himself, who thought of the end-setting members of a motivational set as desires or DTADs, Williams allows that they might

be beliefs. But this, Hurtig argues, is a mistake. An internalist about reasons cannot allow their subjective motivational sets to be contaminated by beliefs on pain of paradox. Thus an internalist about reasons must be a ‘non-cognitivist’ about the subjective motivational set: its membership must be confined to mental items that do not rise to the dignity of beliefs. If you want to be a Humean about motivation you have to go the whole ‘non-cognitivist’ hog.

Essays 9 and 10 focus on the work of Christine Korsgaard whose ideas loom large in these debates. Though a dedicated Kantian, Korsgaard retains an obvious fondness for Hume, perhaps because she was educated by Rawls (We can tell what a Rawlsian education was like since his lectures are now available. See Rawls, 2000.)

In Essay 9, Herlinde Pauer-Studer develops a complex and subtle critique of Korsgaard’s critique of Hume. The key point, I take it, is this. Korsgaard ascribes to Hume the principle that we have a reason to do X if X will achieve our desires. Then if I want to achieve Y, and it is necessary to do X to achieve Y, it follows that I have a reason to do X. But this is implausible if either my ends Y or my means X are wicked or self-destructive. Pauer-Studer, following Broome, objects that this argument rests on a confusion. For the following inference is not valid:

(1) I have a reason to make it the case that (if I want Y and X is a necessary means to Y, that I do X).

(2) I want Y

(3) X is a necessary means to Y.

Therefore

(4) I have a reason to make it the case that I do X

The conditional does not ‘detach’, governed as it is by an intentional operator. In much the same way it is fallacious to infer ‘Necessarily I will sin tomorrow’ from ‘Necessarily if God knows I will sin tomorrow I will sin tomorrow’ and ‘God knows that I will sin tomorrow’ - a point made by Aquinas long ago. In so far as Humeans

can talk about reasons for action, *being a reason for someone to do something* is actually a four place relation, since a fact Δ constitutes a reason for an agent Γ to ϕ given ends Ψ if a belief in the fact Δ would motivate Γ to ϕ in virtue of Γ 's agency and Γ 's ends Ψ . But if Δ constitutes a reason for Γ to ϕ given ends Ψ , it does not follow that if Γ has ends Ψ , then Δ constitutes a reason for Γ to ϕ PERIOD. We cannot transform a four-place relation into a three-place relation by adopting set of desires.

According to Korsgaard her Kantian ethic is better than Hume's because it meets certain requirements that Hume's fails to meet. 1) An adequate meta-ethic (adequate at the level of sociology) must explain motivational normativity, the fact that 'men are often governed by their duties, and are deterred from some actions by the opinion of injustice, and impelled to others by that of obligation'. But 2) it must also explain why moral behavior is justified, why it is rational to do the right thing. Call this justificatory normativity. The basis of Korsgaard's complaint is that Hume succeeds in explaining 1) but not 2). We tend to do what we think is the right thing because we are motivated by our moral sentiments. But that does not explain why we are right to do so. Hume thinks that on reflection we tend to approve of the practice of doing what we tend to approve of, but for Korsgaard this kind of reflective endorsement is not enough. In Korsgaard's view, moral behaviour is justified because each of us shares a common practical identity – that of a rational human being – and that *therefore* we have reason to treat each other with the dignity due to rational agents. Luke Russell (Essay 10) objects that in arguing for this thesis Korsgaard relies on the kind of premise that she scorned in the case of Hume. She points to various kinds of practical necessity to block the justificatory regress – that is, the endless process of 'Why should I?' questions – despite her earlier suggestion that practical necessity is no substitute for positive reasons. She thinks, for instance, that it is arbitrary to halt the reflective process along with Hume when we reach our reflectively endorsed desires. Why isn't it equally arbitrary to stop at our practical identity? There seems to me a further problem if you think that the *same* theory has to explain both motivational and justificatory normativity. Suppose that Korsgaard's argument is a success and her complicated theory explains why it is rational to do the right thing. Since nobody thought of this until she arrived on the scene this theory cannot explain why existing people do what they believe to be the right thing. I certainly do not do what I believe to be right, for Korsgaard's reasons, since, insofar as I understand her theory, I think that

it is false. Thus a *novel* theory which meets requirement 2) is unlikely to meet requirement 1). You can, of course, solve the problem by positing two kinds of normativity with different explanations. But that leaves us with the uncomfortable consequence that when we do the right thing we do not in general do it for the reasons that we should. It may be rational to do the right thing, but most of the people who do it, don't do it because it is rational.

4. The Motivation Argument

The Motivation Argument poses a puzzle for interpreters of Hume. The textual evidence plus the reactions of contemporaries strongly suggest that Hume was what is nowadays known as a moral realist, with a response-dependent view of the moral properties. He thought (or seemed to think) that moral judgments are in today's jargon *truth-apt*, that they are capable of truth and falsity, and that some of them are true. Morality is determined by the reactions of suitably qualified human observers who are presumed to share a moral 'sense', though for Hume, as opposed to Hutcheson, the moral sense, is, in part, a socially constructed affair, the product of a process of social evolution. Thus whether or not a trait is a virtue is 'a plain matter of fact' though a fact to be determined by consulting the feelings of suitable spectators (or by trying to approximate those spectators and consulting ones own feelings). But it is hard to read the Motivation Argument as anything other than an argument for non-cognitivism. This is one reason why Hume was widely regarded as a proto-emotivist (and hence a proto-non-cognitivist) in the middle years of the 20th Century, an interpretation that even today is far from dead especially in the minds of those with a merely undergraduate acquaintance with Hume. (See Hudson, 1969, Ayer 1980, and Flew 1986, for examples of this kind of thing.) Here are four formulations of the Motivation Argument lifted directly from Hume:

Motivation Argument A

- 1) Morals have an influence on the actions and affections. [Premise.]
- 2) Reason alone, as we have already proved, can never have any such influence [Premise: the SRT.]
- 3) Morals ... cannot be derived from reason. (T, 3.1.1.6/457.)

Motivation Argument B

- 1) Morals excite passions, and produce or prevent actions. [Premise.]
 - 2) Reason of itself is utterly impotent in this particular [Premise:the SRT.]
 - 3) The rules of morality, therefore, are not conclusions of our reason.
- (T, 3.1.1.6/457)

Motivation Argument C

- 1) The merit and demerit of actions frequently contradict, and sometimes control our natural propensities.[Premise.]
 - 2) But reason has no such influence. [Premise: an exaggerated version of the SRT.]
 - 3) Moral distinctions, therefore, are not the offspring of reason.
- (T, 3.1.1.10/458-459.)

Motivation Argument D

- 1) The distinction betwixt moral good and evil ... has an influence upon our actions.[Premise.]
 - 2) Of which reason alone is incapable [Premise: SRT.]
 - 3) It is impossible that the distinction betwixt moral good and evil can be made by reason.
- (T, 3.1.1.17/462-463.)

There are several points to note here.

(I) When Hume talks of ‘morals’ or the ‘distinction betwixt moral good and evil’ or ‘the merit and demerit of actions’ having an influence on our actions and affections, what he means is that our *opinions* or *beliefs* about such things often have an influence. It is ‘confirmed by common experience ... that men are often governed by their duties, and are deterred from some actions by the *opinion* of injustice, and impelled to others by that of obligation’(T, 3.1.1.5/457).

(II) Although Hume thinks that our moral opinions *often* influence our conduct there is no suggestion that they are *necessarily* motivating or that they motivate in every case (absent special circumstances). The possibility of indifference or amorality is not excluded. Indeed, as he subsequently makes plain, Hume thinks that there are ‘sensible knaves who ‘feel no reluctance’ at ‘the thoughts of villainy or

baseness' and are therefore *not* 'deterred from some actions by the opinion of injustice'.

(III) The Motivation Argument does not add up to an argument for non-cognitivism unless we interpret 'reason' as our belief-forming faculty and the SRT as the thesis that *beliefs as such* cannot motivate without the aid of a pre-existing passion. If reason is our belief-forming faculty and our moral 'beliefs and 'opinions' are not the offspring of reason then they are not genuine beliefs or opinions. But if we allow the possibility of beliefs or opinions that are *not* derived from reason then the argument only goes to show that our moral opinions are not derived from reason, which is compatible with their being genuine beliefs based on feeling or sentiment.

(IV) Three out of the four formulations of the Motivation Argument are invalid, and version C only approximates validity by exaggerating the SRT. In each case premise 1) says that our moral beliefs have an influence on our actions and affections, a fact which is indeed 'confirmed by common experience'. But except in the case of C, all that premise 2) says is that reason *alone* has no such influence - that is, that reason-based beliefs cannot motivate *without the aid of a pre-existing passion* - which is all that the SRT purports to prove. But if all that 2) says is that reason-based beliefs cannot motivate *by themselves* this opens up the possibility that moral beliefs are in fact derived from reason but do their motivating business with the aid of a pre-existing passion such as a desire to do the right thing. Whatever exactly Hume was trying to prove, he does not manage to prove it.

Thus anyone who wants to extract a decent argument for non-cognitivism from Hume has a good deal of repair work to do. Particularly influential in this connection is the work of Michael Smith for whom a revised version of the Motivation Argument constitutes 'the Moral Problem' the central issue in meta-ethics (Smith, 1994). For Smith thinks that it *is* possible to extract a powerful argument for non-cognitivism from Hume, and that this is a problem since the premises appear to be true and the conclusion false. This is his argument:

- 1) If someone judges that it is right that she ϕ s then, *ceteris paribus*,
[and without the aid of any pre-existing desire] she is motivated to ϕ .

2) An agent can only be motivated to act by a belief when she has a desire for an end to which the belief points out the means, where belief and desire are, in Hume's terms, distinct existences.

Therefore:

3) If someone judges that it is right that she ϕ s then, her [psychological] 'judgment' that it is right that she ϕ s does not constitute a belief.

4) The best explanation of 3) is [insert non-cognitivist theory of choice].

(Smith, 1994, p. 12, slightly doctored. Note that up to step 4), the argument is deductive.)

Several of the papers in this collection (specifically Essays 1-5) are devoted to 'matters arising', including matters arising from Smith's formulations. Was the Motivation Argument intended as an argument for non-cognitivism (or something like it)? Is it possible to extract a decent argument for non-cognitivism from Hume's text and if so does it ultimately succeed? If Hume was *not* arguing for non-cognitivism, what was he arguing for and how is his argument supposed to work? Does the argument succeed and if not, can it be patched up? Since Hume can be construed as arguing *either* for non-cognitivism *or* for a response dependent theory of ethics (both of which continue to be popular amongst philosophers), the Motivation Argument takes us to the heart of contemporary meta-ethics.

Though famous as the contemporary standard-bearer for the error theory, Richard Joyce (Essay 1) does not argue directly for the error theory in this paper, either as a meta-ethic or as an interpretation of Hume. Instead he challenges the Motivation Argument as developed by Hume's non-cognitivist disciples. Suppose we reinterpret premise 1) as the claim that (absent specific circumstances) it is necessary and *a priori* that moral beliefs excite passions, and produce or prevent actions. Call this motivational internalism. Does it follow that the role of moral judgements (conceived as linguistic expressions) must be to express some kind of conative state? Does some kind of non-cognitivism or expressivism follow? Not so, says Joyce, since

expressivism does not imply motivational internalism (thus blocking an inference to the best explanation) and internalism does not imply expressivism.

According to Norva Lo (Essay 2) Hume construes moral properties as response-dependent properties, geared to a shared moral sense. His conceptual analysis makes an action or character virtuous/vicious if and only if all human beings who meet certain conditions (including impartiality and adequate information) are disposed, to feel the sentiment of approbation/disapprobation towards it. Thus Hume cannot have been arguing for non-cognitivism on pain of inconsistency. Indeed, she thinks that Hume was not really an internalist about motivation, that is she thinks that it is possible, according to Hume's theory, to subscribe to a moral belief whilst remaining altogether unmoved. Indeed, it is theoretically possible to acquire moral knowledge without experiencing the moral sentiments (and hence without being in any morally salient conative state). (Consider a Martian anthropologist who has managed to anatomize the human moral sense.) What then is the purpose of the Motivation Argument? It is meant to *disprove* moral rationalism, the thesis that moral knowledge is *typically* the product of some cognitive or discursive process. And Lo provides three reformulations of the Motivation Argument designed to prove precisely this.

Charles Pigden agrees that Hume conceives the moral properties as akin to secondary qualities. To say that a trait is a virtue is to say that it would excite the moral sense of a suitably qualified human observer (impartial, dispassionate and devoid of the delusive glosses of superstition and false religion), just as to say that a flower is violet is to say that it would excite the visual sense of a suitably qualified human observer (someone *with* functioning eyesight but *without* rose-tinted spectacles). Thus the point of the Motivation Argument cannot be to prove non-cognitivism, since Hume was not a non-cognitivist. Furthermore, the argument for non-cognitivism that has been extracted from Hume's text is, in Pigden's view, a failure. Where he parts company from Lo is in thinking that it is impossible to convert the Motivation Argument it a successful argument *against* moral rationalism *and* for *sentimentalism*. *Plausible*, yes, but *successful*, no, since it is formally respectable but based on false and question-begging premises.

Michael Smith (Essay 4) gallantly goes bat for a non-cognitivist, or at least an expressivist, interpretation of Hume, and a rational reconstruction of the Motivation Argument as designed to prove non-cognitivism. He freely admits that he may be reading more into Hume than is really there, and maybe more than Hume could have

conceived of (specifically he suggests a concern with what is practically rational which may have been alien to Hume). But it seems to me that in defending an expressivist interpretation of Hume, Smith is retreating several paces. The expressivism that he is worrying about now is a degenerate descendant of the non-cognitivism that preoccupied him in the nineties. It is okay to say that we have moral beliefs and that these beliefs are true (or false). It is just that in the normative case (but not otherwise) the beliefs are somehow constituted by desires. Even the Motivation Argument that Smith defends in Essay 4 is a feeble abductive variant of the deductive argument that he tried to defuse in *The Moral Problem*. As to the claim that Hume cannot have been arguing for non-cognitivism since he was not a non-cognitivist, Smith has two counters. 1) According to Lo, a trait is virtue for Hume if and only if it meets a certain factual condition. But if ‘X if and only if Y’ is in some sense acceptable, and if ‘Y’ abbreviates a factual statement, then ‘X’ must be factual too. Smith objects that if this principle were correct a non-cognitivist could not coherently suppose that an action *A* is right if and only if *A* maximizes utility, since for him ‘*A* maximizes utility’ is factual whereas ‘*A* is right’ is not. But I think that Smith misconstrues Lo’s argument. Perhaps it is sometimes possible to accept ‘X if and only if Y’ whilst interpreting ‘Y’ as factual and ‘X’ as non-cognitive. But this is hardly possible if the biconditional ‘X if and only if Y’ is intended as a *definition* or a *conceptual analysis* of ‘X’. For a definition or a conceptual analysis can only be correct if ‘X’ and ‘Y’ are in some sense synonymous, which they could not be if one were non-cognitive and the other factual. And when Hume actually ‘defines’ virtue as ‘*whatever mental action or quality gives to a spectator the pleasing sentiment of approbation*’ (EPM, App.1.10/289) it does indeed seem that conceptual analysis (or what we would nowadays *call* conceptual analysis) is what he has in mind. 2) Smith claims that on Pigden’s showing, Hume comes out as a non-cognitivist in the end since moral beliefs, like Pigden’s belief that Zena is desirable, are constituted by desires. But that is not the thesis. The claim is rather that moral beliefs are typically *caused* and *justified* by feelings of approval and disapproval not that they are *constituted* by such feelings, since they are beliefs about what a certain kind of spectator *would* feel. To believe that Zena is desirable is not necessarily to desire her, even if such beliefs are typically caused by feelings of sexual desire, and even if Hume himself identifies *some* beliefs with impressions.

In Essay 5 (based on his book *Value, Reality and Desire*) Graham Oddie takes up the idea that to be good, or to be valuable, is to be worthy of desire and that the

evidence for value-judgments consists in our desires and aversions. Such evidence is defeasible, of course, since our desires may be due to some pathological condition. Still, to experience something *as* valuable is to feel the relevant desires. Hence if I *know* that something is valuable (at least if my knowledge has been derived from experience) I will in fact be motivated to pursue or promote it. Thus moral knowledge (and indeed moral belief) has a tendency to ‘excite passions, and produce or prevent actions’. Oddie seems to subscribe to a theory rather like Hume’s (at least as analyzed by Lo and Pigden) but this is not quite correct. For Oddie, unlike Hume, does not cash out *worthiness to be desired* in terms of the dispositions of actual human agents (whether idealized or not). Nonetheless, the fact that something is desired is *evidence* of its desirability, and it is because we form value-judgments on the basis of desires that value-judgments - or rather the desires on which they are based - tend to motivate. Thus Oddie’s theory is able to reconcile the premises of the Motivation Argument - that morals ‘excite passions, and produce or prevent actions’ and that ‘reason itself is utterly impotent in this particular’ - with a realist, indeed, a non-naturalist, conception of value.

5. Hume as a Virtue Theorist

In the 19th Century and in the early part of the 20th, Hume was widely regarded as an ancestor of utilitarianism. He was censured in his own day (and by one of his best friends) for reducing human goodness to ‘that for which we commend a chest of drawers’, since he regarded the one, like the other, as a matter of being conducive to useful or agreeable effects (Smith, 1976 [1759], p. 188). Bentham, more enthusiastically, claimed Hume as an inspiration. ‘But, after all retrenchments, there will still remains enough [in Book 3 of the *Treatise*] to have laid mankind under indelible obligations. That the foundations of all virtue are laid in utility, is there demonstrated, after a few exceptions made, with the strongest force of evidence’ (Bentham, 1988 [1776], p. 51n). Sidgwick concurred, though with a characteristic caveat: ‘When Hume presented utilitarianism as a mode of explaining current morality, it was seen or suspected to have a partially destructive tendency.’ (Sidgwick, 1907, p. 86). But though Hume was indeed a precursor of utilitarianism, he was not really a utilitarian himself. Utilitarianism is a theory of the right and the good, the right being what we ought to do and the good being what we ought to bring about, and it shares the 19th and early 20th Century obsession with the ‘thin’ moral concepts: rightness, duty,

goodness and badness. These are not the concepts that exercised Hume. It was the virtues in general and some specific virtues in particular, such as justice, benevolence and chastity that interested him. But from about 1850 to about 1950 the big questions in ethics could be summed up in the chapter headings of W.D Ross's *The Right and the Good* (1930): 'The Meaning of "Right"', 'What makes Right Acts Right?', 'The Meaning of "Good"', 'The Nature of Goodness', etc, with the proviso that as the 20th Century wore on, the semantic status of claims involving these concepts - were they factual statements about non-natural properties or expressions of emotion? - began to be a major preoccupation. But in the 1950s, beginning with the work of Philippa Foot and G.E.M Anscombe, a reaction set in and the virtues returned to the fore. (See the essays collected in Foot, 1978 and 2002, and in Anscombe, 1981.) The focus was less on what we ought to do and more on what sort of people we ought to be, less on the nature of predicative goodness (a concept regarded with increasing suspicion) and more on what it is to be a good *human being*. Although the original inspiration of this movement was Aristotelian, and although Foot and Anscombe themselves display a marked hostility to Hume, later virtue ethicists began to see Hume as a kindred spirit. After all, he *was* preoccupied with the virtues, rather than the 'thin' moral concepts, and his section headings - 'Of vice and virtue', 'Of Virtue and Vice in General', 'Justice, whether a natural or an artificial virtue?' 'Of chastity and modesty' - look much more like paper titles by Foot than chapters headings by Ross. Hume also shares the virtue ethicists' obsession with human character and human nature. and their interest in social institutions such as promising. And for some virtue ethicists the determinedly modernist and secular cast of his thought was an added attraction. Hume shows that you can put the virtues at the centre of the moral life without succumbing to the creeping Thomism that characterizes Anscombe and Foot. The last four essays in this collection deal with Hume as a virtue theorist.

'What sort of a virtue theorist is Hume?' asks Christine Swanton (Essay 11). Well, to begin with, he is a virtue theorist who *defines* a virtue as a trait which arouses approbation in suitably qualified spectators. But this still leaves a wide arena for empirical investigation. What traits *do* we approve of and why? Do they have anything interesting in common apart from the fact that we are disposed to approve of them? In Swanton's view, Hume has a complex and pluralistic theory of the virtues according to which we are disposed to approve of traits because of a wide variety of features related to human flourishing and bound up with the exigencies of human life. Traits excite

approbation because they are expressive of the bonds of a (non-pathological) love, because they are the objects of a proper pride, because they express joy, because they display a justice-based respect for persons and their property and because they manifest a proper appreciation of status. Many of these traits are productive of human welfare broadly conceived (so that Hume's utilitarian disciples were not wholly wrong). Nonetheless, for Hume there is a lot more to welfare than pleasure and the absence of pain.

For Annette Baier (Essay 12) all this is a bit too pat. She is not sure that Hume is the gung-ho normative virtue theorist that Swanton takes him to be. Sometimes, like Locke, he seems to be using 'virtue' in a sociological sense to refer to what is considered as such in this society or in that. The martial virtues of the Scythians, parading their towels made from human hides, are a far cry from the polite virtues of an Eighteenth Century Scottish gentleman such as Hume himself. Sometimes Hume hints that the cross-cultural virtues - qualities that would be approved by *any* non-deluded and impartial human being - may be in short supply. He often seems to be more interested in anatomizing what are regarded as virtues than in arguing that a particular set of virtues are the real cross-cultural thing. Sometimes he is a cataloguing virtue theorist, sometimes an offensive virtue theorist, out to raise puritan hackles. But though not devoid of normative concerns he is often less of a moral legislator and more of a moral sociologist than Swanton seems to think. Swanton gives us Hume without the ambiguities. Baier wants to put the ambiguities back in.

Swanton replies (Essay 13) by clarifying the concept of a normative virtue ethics, distinguishing Aristotelian virtue ethics in particular from normative virtue ethics in general, and emphasizing that it is possible to be a virtue ethicist without being unduly tainted with Aristotelianism. Which means, of course, that it is not inappropriate to construe an anti-scholastic like Hume as a normative virtue-ethicist.

Though Thomas Jefferson notoriously detested David Hume ('this degenerate son of science, this traitor to his fellow-men') this was primarily on account of his history, not because of his philosophy (Jefferson, 1999, pp 262, 274, 283-5, 383-4). But if Rosalind Hursthouse is right, (Essay 14) Jefferson would have detested Hume even more if he had properly understood the third book of the *Treatise*. 'We hold these truths to be self-evident, That all men are created equal, that they are endowed by their creator with certain unalienable rights; that among these are life, liberty & the pursuit of happiness, that to secure these rights governments are instituted among men.' Thus

for Jefferson and the drafting committee of the *Declaration of Independence*, rights are primary and virtues secondary. Justice, for instance, consists, in part, in giving each person their due, that is in respecting their pre-existing rights. But according to Hume - or, at least, according to Hume as reconstructed by Hursthouse - this gets things the wrong way round. Justice does not consist in respecting pre-existing rights - rather rights are created by the conventions of justice which, for Hume, is an artificial virtue that has evolved because it serves the interests of the members of society. And this applies as much to *human* rights as to the rather mundane property-rights, that preoccupy Hume in Book 3 of the *Treatise*. It is not quite that that talk of human rights is 'nonsense on stilts' (in Bentham's famous phrase). It is just that it is a mistake to suppose that human rights are 'there' prior to the evolution of the the artificial virtue of treating people as if they had them. Not for nothing, it seems, was Hume the buddy of Edmund Burke and the inspiration of Jeremy Bentham, those two notorious critics of the rights of man and of the citizen! (Waldron, 1987.)

Thus Hume has something interesting to say about reason, motivation, morals and virtue and maybe even human rights. Now, read on ...