I

Studying virtues

This book is a prolegomenon to a study of human virtues in their connections with emotions and related affective / motivational states. Accordingly, it is a contribution to the broad and diverse field of “ethics,” in particular to the subfield sometimes known as “moral psychology.” Initiated (or rather, revived) fifty years ago by Elizabeth Anscombe’s “Modern Moral Philosophy” (1958) and given increased momentum by several works of Alasdair MacIntyre (1981, 1988, 1990, 1999), the study of virtues is now a major part of moral philosophers’ business. In this introductory chapter I will explore the question of how to study the virtues and defend this book’s approach.

To ask how the virtues should be studied immediately raises the more basic question of why study the virtues. Our way of studying them will surely be affected by what we take to be the point of doing so. What is the point of studying the virtues? What do we want to accomplish in studying them? This question, in turn, leads to a broader one: why study moral concepts? Why study the concept of the good, or of obligation, or of supererogation? Maybe we ought to study the virtues for the same reason we ought to study these other concepts. And this question really comes to why we should do moral philosophy – assuming that moral philosophy is in large part the study of moral concepts.

Virtues and theories

Many philosophers these days would say that the point of studying moral concepts, including the concept of a virtue or the concepts of the virtues, is to produce a moral theory. This answer raises a couple of other questions, namely, what are moral theories? and what good are they?

If we look at recent writings on the virtues – say, Michael Slote’s *Morals from Motives* (2001) and Rosalind Hursthouse’s *On Virtue Ethics* (1999) and many papers on virtue that have appeared in professional journals in
recent years, we note frequent recurrence of words like “derive,” “derivative,” “depend on,” “subordinate,” “define in terms of,” “construct out of,” “based,” “grounded,” “grounding,” “independent,” “conceptually prior,” “supreme,” “fundamental,” “foundation,” and the like in connection with moral concepts. That is, we find authors preoccupied with ordering moral concepts in such a way that some of them are subordinated to or derived from other concepts or some single other concept so that some one or small number of moral concepts become the source, the ground, the foundation, the base, the site of derivation, of the others. This ordering will also be an order of justification. The foundational concept presumably does not need justifying, or is self-justifying, and it justifies the other, derivative concepts. Ethical theories will then differ according to which concept or concepts provide the foundation, and which ones are derivative. They will also differ as to how the derivations work and which concepts are closer to the foundation and which more distant from it.

This is one standard model for ethical theory. The pattern is widespread in the modern period. David Hume tends to derive ethical concepts from the human passions, Immanuel Kant from a certain conception of practical reason or obligation, utilitarians from the concept of quantified pleasure, and social contract theorists from the concept of harmonious social functioning. Mark Murphy (2002: 6) comments that “The principle that God is to be obeyed is, on normative versions of divine command theory, supreme due to its status as an independent moral principle that is the source of the correctness of all other moral principles.” Some thinkers have tried to avoid the implausibilities entailed by such a broad divine command theory by deriving only the concept of obligation or duty – not, say, the concept of the good or the concepts of all the virtues – from the idea of a divine command. See Adams 1999: chap. 10 and Evans 2004: chap. 5. On this hierarchizing model of moral theory, in “virtue ethics” the concept of virtue somehow becomes the source or foundation of the other concepts of ethics.

A particularly clear example is Slote’s 2001 book, whose project is to defend an “agent-based approach to virtue ethics,” which “treats the moral or ethical status of acts as entirely derivative from independent and fundamental aretaic (as opposed to deontic) ethical characterizations of motives, character-traits, or individuals” (5). Slote’s proposed moral theory accounts for the moral status of acts (good or bad, right or wrong) by deriving it “entirely” from aretaic construals of motives, etc. Thus if we wish to know whether a particular action is a good one, we will try to ascertain the agent’s motive in performing it. If his motive was virtuous,
then the action was good, and if it was vicious, the action was bad. So, if I give $10 to a beggar out of a morally indifferent motive such as a desire to lighten my billfold a bit, my action is morally indifferent; and if I do so out of a virtuous motive, such as compassion for the beggar, my action is good; and if I perform the action out of a vicious desire to humiliate my walking companions, my action is bad. And this is the whole story about the moral goodness or badness of the action, on the sort of theory that Slote is promoting. These moral qualities of the acts do not derive at all from the social consequences of the acts — say, the benefit or harm I do to the beggar or the damage to my relationship with my walking companions — nor do the acts have an “independent” moral value. Furthermore, the moral status of the motives of compassion or cruelty is not derived from any ethical concepts at all. It is “independent [of other moral concepts] and fundamental.” This arrangement of the concepts may seem rather artificial and even paradoxical, but it has the merit of bringing a simplifying order to the conceptual array, an order, furthermore, that is secured at a certain point. The whole idea of a base or foundation is that of something that needs no securing or deriving, something that will stand fast by itself and give orientation and steadiness to the whole, keeping everything else from slipping. And Slote’s theory is an example of virtue-ethics because of the place the virtue-concepts occupy in the conceptual system: they (or rather, their motivational aspect) form the base or foundation.

Slote’s thought is a particularly clear example of this understanding of moral theory, but we find many examples of it in the literature. Gary Watson (1990) says that an ethics of virtue is “a set of abstract theses about how certain concepts are best fitted together for the purposes of understanding morality . . . [in which] the concept of virtue is in some way theoretically dominant” (451). “An ethics of virtue is . . . the . . . general claim that action appraisal is derivative from the appraisal of character” (452). Watson attributes to John Rawls the view that moral theories differ according to the direction of conceptual derivation. Linda Zagzebski (1996) says that her favored type of moral theory “makes the concept of a right act derivative from the concept of a virtue or some inner state of a person that is a component of virtue” (79), and that “pure virtue theorists deny that virtue is an excellence because it is a means to some external good” (99). Robert Louden’s (1997) critique of virtue ethics assumes that it is a theory analogous to deontology and utilitarianism, thus characterized by “conceptual reductionism.” For a heroic defense of the foundation proposed by “pure” virtue theory, see Kawall 2009.

Virtues and theories
We see this pattern in Rosalind Hursthouse’s (1999) comment that “according to virtue ethics – and this book – what is wrong with lying, when it is wrong, is not that it is unjust (because it violates someone’s ‘right to the truth’ or their ‘right to be treated with respect’) but that it is dishonest, and dishonesty is a vice” (6). (Note the exclusions. But isn’t injustice as much a vice as dishonesty?) Later in the book she affirms vegetarianism “on the grounds that (i) temperance (with respect to the pleasures of food) is a virtue, and (ii) for most of ‘us’, eating meat is intemperate (greedy, self-indulgent)” (227). But her theory is not a virtue ethics in the strong sense of Watson and Slote, since it makes human nature even more fundamental than the virtues (see her chaps. 9–11).

However, Hursthouse appears to me to be inconsistent in her advocacy of any moral theory at all. For she advocates peaceful coexistence among utilitarianism, deontology, and virtue ethics (5) after, however, earlier noting that virtue ethics has now “acquired full status, recognized as a rival to deontological and utilitarian approaches” (2), and she does do the characteristic theoretical thing of deriving other ethical concepts from her privileged one – virtue – to the exclusion of others, even if she does not make that concept foundational. On the purer, Watsonian or Slotean conception, virtue ethics could not peaceably coexist with the other theories because each theory’s defining structure – its exclusive derivation of all other concepts from its foundational concept – is incompatible with each other theory’s defining structure.

Sometimes, of course, the conceptual regimenting move typical of modern theory is directed at the concept of virtue. Thus the concept of virtue is derived from some other single item in the array of ethical concepts. Kant’s reduction of virtue to strength in the sense of duty is one historical example. Another is Hume’s somewhat indecisive reduction of it to traits that we respond to with positive emotions. A contemporary example is Julia Driver’s consequentialist theory (Driver 2001). On Driver’s theory, a trait is a virtue exactly to the extent that it generates good consequences, given a “normal” world for its operation. It may accomplish this by a variety of mechanisms that have been thought by many people to contribute non-instrumentally to a trait’s status as a virtue – say, by way of benevolent desires, or intentions to perform virtuous actions, or sound moral opinions, or skillful practical reasoning, or morally good emotions. In principle, any dispositional mechanism that regularly produced good states of affairs would be a virtue. It could be a disposition to get pleasure, to reason
fallaciously, to wish other people ill, to believe falsehoods, or to intend harm to others (see Driver’s example of the Mutors in 2001: chap. 3).

Julia Annas (2011: 110–11) points to some “intuitions” about virtues that we must give up if virtues’ only value is their good consequences. No doubt one consideration that favors counting a trait as a virtue (Slote notwithstanding) is its tendency to generate good consequences, but we praise people not just for the good results of their traits, but for their motives (as Slote would note); and if the results are very good but the motives stink, we don’t call the agent virtuous. We praise people for their virtuous motives even while lamenting the outcome of their actions. Another consideration that favors judging a trait to be a virtue is that it constitutes an aspect of its bearer’s being an excellent specimen of the human kind. Furthermore, as Driver herself notes, we insist more strongly on appropriate motivation in attributing some virtues (for example, generosity) than others (say, justice). These aspects of our moral practice would have to be revised if we took Driver’s theory seriously. In my view, we are less likely to distort the moral concepts and their associated practices if we follow Annas’s policy of aiming to build upwards from our conceptions of virtue, rather than start on the level of abstract discussion of already developed theory. From this point of view losing such central aspects of virtue [as are lost through theorizing in the modern style] makes it at least unclear what the point is of having a theory of virtue at all. (Annas 2011: 111)

I try to exemplify Annas’s approach in Chapter 9, looking at the virtues individually to see how they actually work and interrelate, rather than trying to develop a monolithic theory that runs the risk of unhappy exclusions and procrustean adjustments.

**Why theory?**

In general, ethical theories differ from one another according to which concept is taken to be the foundational source for the other, derivative, concepts, but they tend to have this common basis-and-derivation structure. Why think ethics ought to have this structure? What is to be gained by trying to order the concepts in this way? One possible answer is “the truth about the concepts.” Maybe the ethical concepts just are ordered in one of the ways that the various ethical theories say they are, and the point of ethical theory is to ascertain and clearly present these properties of the moral concepts, much as a point of any science is to ascertain and clearly
present the properties of the things in its domain. It seems to me that moral concepts do, clearly, bear relations of logical or quasi-logical dependency on one another and that the moral philosopher is the chief investigator of these conceptual relations. An interest in the truth about such relations seems respectable enough. But the effort to find a single, exclusionary ordering of the hierarchical kind that I have been illustrating appears always to generate implausible claims and paradoxes which are then leapt on by theorists with rival agendas. Thus we get the interminable disagreements characteristic of discussions between deontologists, consequentialists, social contract theorists, and most recently virtue theorists. Perhaps the truth about the moral concepts is that they are not orderable in the way aspired to by moral theory. Besides the implausibility of moral theories as truth-claims about concept-order, most ethical theorists, if pushed, would be unwilling to think of philosophical ethics as merely a “pure science” of concepts; they want more than just to know, for intellectual satisfaction’s sake, what the logical ordering of the moral concepts is; instead, they tend to think that moral philosophy—and therefore presumably moral theory—has some moral point.

One possible moral point for ethical theory is “foundationalist”: if we can find a foundation that is genuinely secure, and then derive from it, by equally secure methods, everything else that is crucial in ethics, we may have a defense against the moral skepticism and relativism that otherwise seem implied by the deep disagreements that arise in a morally pluralistic society. And as a solvent of moral disagreements, the theory might also be thought a powerful antidote to the social conflicts that can arise from them. This explanation of moral theory’s allure is initially plausible if applied to the older theories, in their original contexts. If it were true, as Kant thought, that practical reason is a singular authoritative structure that is universally accessible to human minds and that lays down a categorical imperative of duty from which all particular obligations can be derived, then from this foundation of reason it might be possible to resolve all moral disagreements and thus dispel the skepticism or conflict that fundamental disagreement seems to entail. Something similar might be admitted if it were true, as the early utilitarians thought, that there is a single, universally plausible concept of human happiness from which, by an incontestable application of universal logic, all particular attributions of the good and the right can be derived. But at the present day, I take it that not even the most confirmed Kantian or utilitarian thinks that the proposed foundation of his or her theoretical structure has the incontestability and universally accessible obviousness that it needs to
perform the foundationalist function. In fact, the debates about Kantian and utilitarian ethics concern, to a great extent, the viability of the proposed foundations in light of what appear to be unacceptable moral judgments that derive from them by the rules of derivation that they employ.

Hursthouse points out that if our ethical evaluations of human beings are to follow the naturalistic model, we have to have some conception of human nature. And she admits, ever so briefly, that this can be controversial:

There is, of course, room for disagreement over what we are. It might be said, for example, that what human beings are are possessors of an immortal soul through which they can come to know and love God for eternity. But “ethical naturalism” is usually thought of as not only basing ethics in some way on considerations of human nature, but also as taking human beings to be part of the natural, biological order of living things. Its standard first premise is that what human beings are is a species of rational, social animals and thereby a species of living things – which, unlike “persons” or “rational beings”, have a particular biological make-up and a natural life-cycle.

If all the above is basically right, then . . . we would expect the structure of our ethical evaluations of ourselves to resemble that of a sophisticated social animal with some differences necessitated by our being not only social but also rational. (Hursthouse 1999: 206)

Hursthouse eschews the foundationalist project, but here she is trying to reduce the number of possible interpretations of “human nature.” The argument is weak. People may agree that human beings are “part of the natural, biological order” and yet think that the most important part of our nature is something else. It’s perfectly possible to think that we are natural biological beings with an immortal soul. And the doctrine of the immortal soul is not necessary for religious conceptions of human nature. With much of pre-Christian Judaism, one might believe that human beings are created in the image of God, without having any opinion about whether the soul is immortal. And part of that concept of human nature would be that human beings are not just biological entities, distinguished from the other animals by being a bit smarter in some ways, but are created with a disposition to worship God and a need to do so and to obey him if they are to function well qua human beings. To religious people, the idea of leaving out this dimension of human nature is laughable. Believing in a God like the Judeo-Christian God will significantly affect one’s list of virtues, for in that case we need not just virtues that fit us for our human-human social
life, but also virtues for our divine-human social life: faith, hope, love (of God), and this reference will alter all the other virtues in various ways.

And the numerous religious conceptions of human nature are not the only variants that can affect our ethics and our list of virtues. Aristotle’s sexism and belief in natural slaves is not just a result of cultural myopia. His teacher Plato had already called Aristotle’s brand of sexism into question, and in the nineteenth century the very sophisticated Nietzsche revived something like Aristotle’s belief in natural slavery. Some people still take Stoicism seriously as an account of human nature, and it doesn’t fit very well Hursthouse’s biological reduction. And Buddhism seems to have yet another way of construing human nature – or perhaps denying altogether that there is such a thing.

Hursthouse adds to certain biological features of human beings that we share with the other animals the attribute of rationality, but what rationality consists in and what its norms are is highly contested. Inspired by economic theories (Becker 1978), sociologists and political scientists have been proposing “rational choice” explanations of human behavior for several decades, but the concepts of practical rationality in the literature are rife and every one of them is contested by one theorist or another. Studies in empirical social psychology have challenged most of the axioms of rationality proposed by rational choice theorists (see Tversky and Kahneman 1986). In the light of all these things, Hursthouse’s attachment to “naturalism” as a kind of biological philosophy can begin to look a little parochial or simplistic.

As I said, Hursthouse does not in the strictest sense advocate a virtue-based ethical theory. But the real virtue ethicists, the advocates of a “pure virtue theory,” would have an even harder time making plausible the idea that their theory can fulfill foundationalist aspirations. Proposed lists of virtues are as controversial as other kinds of moral claims. Just as there are multiple rival concepts of human nature, practical rationality, and human happiness, so there are multiple and rival concepts of the virtues. Even where two moral traditions agree that, say, courage or justice is a human virtue, it is not so clear that they are agreeing on a virtue. Philosophical analysis will readily show differences in virtue-concepts belonging to different moral traditions, even when the names of the virtues are the same across traditions.

**The vices of moral theory**

The derivations that are motivated by the theoretical mind-set are notoriously problematic. Earlier I quoted Hursthouse to the effect that “what is wrong with lying, when it is wrong, is not that it is unjust (because it
violates someone’s ‘right to the truth’ or their ‘right to be treated with respect’) but that it is dishonest, and dishonesty is a vice” (1999: 6). It seems clear that one thing wrong with lying is that it expresses a vice (when it does express a vice). But it seems equally obvious that other things are or can be wrong with lying: its wrongness can derive from its being a violation of somebody’s right, either to have the truth or be respected or both. It can also be a violation of God’s will. And surely its wrongness derives also from the nasty consequences that so often follow, including human suffering, waste of resources, and the disruptions of social life and degradation of human relationships.

The theorist’s exclusivist mind-set closes off promising avenues of reflection and insight. Hursthouse tells us that vegetarianism is right because temperance is a virtue and meat eating is intemperate (1999: 227), but it seems more plausible to ground the value of vegetarianism in the supposition that animals have a right not to be killed for food, or that meat eating causes animal suffering, or that meat eating cuts down on the overall food supply in the world by feeding one kind of food to another kind of food. If we don’t appeal to any of these consequences or rights, it is not clear how meat eating, in itself, is intemperate. On the dependency of temperance on other kinds of moral considerations, see Roberts 2013.

Slote notes an objection to his agent-based virtue ethics that follows the pattern of my criticisms of Hursthouse’s theoretical exclusivism. The objection is that if to do the right thing is to act with virtuous motivation, as agent-based virtue ethics holds, then it’s not possible to do the right thing for the wrong reason. But this happens all the time, and the philosopher’s ethical reflections ought to allow a place for this, and even illuminate how it’s possible. An example from Sidgwick is a prosecutor who does his duty in prosecuting a defendant, but does it from malice. Slote’s answer to this objection seems to be that if the prosecutor prosecutes, he does it from some sense of duty; because if, horrified by his malice, he decided not to prosecute at all, his inaction would be motivated by “insufficient concern for the public . . . good” (2001: 14). If he does his duty at all, then, he must be motivated by a sense of duty – so this is not a case of doing the right thing for the wrong reason. This defense depends on the fallacy of inferring that if somebody has a duty and acts in conformity with it, the act must be motivated by a sense of duty. It’s bad philosophical practice to construct the moral concepts in a way that rules out distinguishing the value of an action from the value of its motive. In Chapter 6 I’ll show how emotions can give various moral identities to an action, but nothing that I say rules out evaluating an action independently
of its motivation. Many actions are complex enough to be evaluated in a variety of ways – in terms of what they express, how well-thought-out they are, and their human relational character, in addition to the two ways we've been discussing.

Much of moral theorists’ energy is expended in trying to make plausible the implausibilities created by their conceptual reductions. Slote is aware how implausible it is to regiment the moral concepts in a way that rules out evaluating actions in terms of their consequences. To avoid this conclusion, he points out that people who are virtuously motivated just do consider carefully the likely consequences of their actions (2001: 17, 34). He thinks this observation protects the priority of motivation while fully acknowledging the importance of consequences, but he seems not to notice how arbitrary this construction is. Rather than say that consequences of a certain kind are good because they are the kind virtuous people aim at, why not say that virtuous people’s aims are virtuous because they take good consequences as their objects? If certain kinds of consequences, such as disease and disability and despair and loneliness, were not bad independently of the motivations of people who bring them about, it is hard to see how the desire to avoid such consequences and promote their contraries would be virtuous. Virtues seem to be, in part, patterns of appropriate response to the way the world is, dispositions to avoid bad states of affairs, to correct defects in the world and to make the world approximate better the way it ought to be. Again, it’s desirable to be able to acknowledge this kind of plausible explanatory conceptual connection (between motivation and consequences), and Slote’s derivational theoretical mind-set rules it out. Our understanding of moral concepts would be better served by having no theory at all, if every theory must prevent our acknowledging some of the ways the moral concepts work.

An argument for moral theory

Slote offers one argument in favor of moral theory in general: the concepts of commonsense morality are disordered, and moral theory can clean up the mess. He gives two examples. Of two equally negligent acts, non-theoretical morality has us blame much more severely the agent whose act results in a disaster than we blame an agent when, by good luck, no disaster occurs. Slote regards this as an incoherency in commonsense morality. The other example is the asymmetry between our judgments of obligation, as regards others and ourselves. We think we sometimes have an obligation to do good for others and always an obligation not to harm