

Moral reasoning is individual or collective practical reasoning about what, morally, one ought to do. Philosophical examination of moral reasoning faces both distinctive puzzles — about how we recognize moral considerations and cope with conflicts among them and about how they move us to act — and distinctive opportunities for gleaning insight about what we ought to do from how we reason about what we ought to do.

Part I of this article characterizes moral reasoning more fully, situates it in relation both to first-order accounts of what morality requires of us and to philosophical accounts of the metaphysics of morality, and explains the interest of the topic. Part II then takes up a series of philosophical questions about moral reasoning, so understood and so situated.

1. The Philosophical Importance of Moral Reasoning

1.1 Defining “Moral Reasoning”

This article takes up moral reasoning as a species of practical reasoning — that is, as a type of reasoning directed towards deciding what to do and, when successful, issuing in an intention (see entry on [practical reason](#)). Of course, we also reason theoretically about what morality requires of us; but the nature of purely theoretical reasoning about ethics is adequately addressed in the various articles on [ethics](#). It is also true that, on some understandings, moral reasoning directed towards deciding what to do involves forming judgments about what one ought, morally, to do. On these understandings, asking what one ought (morally) to do can be a practical question, a certain way of asking about what to do. (See [section 1.5](#) on the question of whether this is a distinctive practical question.) In order to do justice to the full range of philosophical views about moral reasoning, we will need to have a capacious understanding of what counts as a moral question. For instance, since a prominent position about moral reasoning is that the relevant considerations are not codifiable, we would beg a central question if we here defined “[morality](#)” as involving codifiable principles or rules. For present purposes, we may understand issues about what is right or wrong, or virtuous or vicious, as raising moral questions.

When we are faced with moral questions in daily life, just as when we are faced with child-rearing, agricultural, and business questions, sometimes we act impulsively or instinctively and sometimes we pause to reason, not just about what to do, but about what we ought to do. Jean-Paul Sartre described a case of one of his students who came to him in occupied Paris during World War II, asking advice about whether to stay by his mother, who otherwise would have been left alone, or rather to go join the forces of the Free French, then massing in England (Sartre 1975). In the capacious sense just described, this is probably a moral question; and the young man paused long enough to ask Sartre's advice. Does that mean that this young man was reasoning about his practical question? Not necessarily. Indeed, Sartre used the case to expound his skepticism about the possibility of addressing such a practical question by reasoning. But what is reasoning?

Explicit reasoning is responsibly conducted thinking, in which the reasoner, guided by her assessments of her reasons (Kolodny 2005) and of any applicable requirements of rationality (Broome 2009), attempts to reach a well-supported answer to a well-defined question. Characterizing “reasoning” as “responsibly conducted thinking” of course does not suffice to analyze the notion. For one thing, it fails to address the fraught question of reasoning's relation to inference (Harman 1986, Broome 2009). Still, it will do for present purposes. It suffices to make clear that the idea of reasoning involves norms of thinking that we often fail to live up to.

1.2 Empirical Challenges to Moral Reasoning

To start with the mildest departure from norms of reasoning: our thinking, including our moral thinking, is often not explicit. We could say that we also reason tacitly, thinking in much the same way as during explicit reasoning, but without any explicit attempt to reach well-supported answers. In some situations, even moral ones, we might be ill-advised to attempt to answer our practical questions by explicit reasoning. In others, it might even be a mistake to reason tacitly — because, say, we face a pressing emergency. Yet even if we are not called upon to think through our

options in all situations, and even if sometimes it would be positively better if we did not, still, if we are called upon to do so, then we should conduct our thinking responsibly: we should reason.

Recent work in empirical ethics has indicated that even when we are called upon to reason morally, we often do so badly. When asked to give reasons for our moral intuitions, we are often “dumbfounded,” finding nothing to say in their defense (Haidt 2001). Our thinking about hypothetical moral scenarios has been shown to be highly sensitive to arbitrary variations, such as in the order of presentation. Even professional philosophers have been found to be prone to such lapses of clear thinking (e.g., Schwitzgebel & Cushman 2012). Some of our dumbfounding and confusion has been laid at the feet of our having both a fast, more emotional way of processing moral stimuli and a slow, more cognitive way (e.g., Greene 2007). An alternative explanation of moral dumbfounding looks to social norms of moral reasoning (Sneddon 2007). And a more optimistic reaction to our confusion sees our established patterns of “moral consistency reasoning” as being well-suited to cope with the clashing input generated by our fast and slow systems (Campbell & Kumar 2012).

Eventually, such empirical work on our moral reasoning may yield revisions in our norms of moral reasoning. This has not yet happened. This article is principally concerned with philosophical issues posed by our current norms of moral reasoning. For example, given those norms and assuming that they are more or less followed, how do moral considerations enter into moral reasoning, get sorted out by it when they clash, and lead to action? And what do those norms indicate about what we ought to do do?

1.3 Situating Moral Reasoning

The topic of moral reasoning lies in between two other commonly addressed topics in moral philosophy. On the one side, there is the first-order question of what moral truths there are, if any. For instance, are there any true general principles of morality, and if so, what are they? At this level utilitarianism competes with Kantianism, for instance, and both compete with anti-theorists of various stripes, who recognize only particular truths about morality (Clarke & Simpson 1989). On the other side, a quite different sort of question arises from seeking to give a metaphysical grounding for moral truths or for the claim that there are none. Supposing there are some moral truths, what *makes* them true? What account can be given of the truth-conditions of moral statements? Here arise familiar questions of [moral skepticism](#) and [moral relativism](#); here, the idea of “a reason” is wielded by many hoping to defend a non-skeptical moral metaphysics. The topic of moral reasoning lies in between these two other familiar topics in the following simple sense: moral reasoners operate with what they take to be morally true but, instead of asking what *makes* their moral beliefs true, they proceed *responsibly to attempt to figure out* what to do in light of those considerations. The philosophical study of moral reasoning concerns itself with the nature of these attempts.

These three topics clearly interrelate. Conceivably, the relations between them would be so tight as to rule out any independent interest in the topic of moral reasoning. For instance, if all that could usefully be said about moral reasoning were that it is a matter of attending to the moral facts, then all interest would devolve upon the question of what those facts are — with some residual focus on the idea of moral attention (McNaughton 1988). Alternatively, it might be thought that moral reasoning is simply a matter of applying the correct moral theory via ordinary modes of deductive and empirical reasoning. Again, if that were true, one’s sufficient goal would be to find that theory and get the non-moral facts right. Neither of these reductive extremes seems plausible, however. Take the potential reduction to getting the facts right, first.

Contemporary advocates of the importance of correctly perceiving the morally relevant facts tend to focus on facts that we can perceive using our ordinary sense faculties and our ordinary capacities of recognition, such as *that this person has an infection* or *that this person needs my medical help*. On such a footing, it is possible to launch powerful arguments against the claim that moral principles undergird every moral truth (Dancy 1993) and for the claim that we can sometimes perfectly well decide what to do by acting on the reasons we perceive instinctively — or as we have been trained — without engaging in any moral reasoning. Yet this is *not* a sound footing for arguing that moral reasoning, beyond simply attending to the moral facts, is always unnecessary. On the contrary, we often find ourselves facing novel perplexities and moral conflicts in which our moral perception is an inadequate guide. In addressing the moral questions surrounding whether society ought to enforce surrogate-motherhood contracts, for instance, the scientific and technological novelties involved make our moral perceptions unreliable and shaky guides. When a medical researcher who has noted an individual’s illness also notes the fact *that diverting resources to caring, clinically, for this individual would inhibit the progress of my research, thus harming the long-term health chances of future sufferers of this illness*, he or she comes face to face with conflicting moral considerations. At this juncture, it is far less plausible or satisfying simply to say that, employing one’s ordinary sensory and recognitional capacities, one sees what is to be done, both things considered. To posit a special faculty of moral intuition that generates such overall judgments in the face of conflicting considerations is to wheel in a *deus ex machina*. It cuts inquiry short in a way that

serves the purposes of fiction better than it serves the purposes of understanding. It is plausible instead to suppose that moral reasoning comes in at this point (Campbell & Kumar 2012).

For present purposes, it is worth noting, David Hume and the moral sense theorists do not count as short-circuiting our understanding of moral reasoning in this way. It is true that Hume presents himself, especially in the *Treatise of Human Nature*, as a disbeliever in any specifically practical or moral reasoning. In doing so, however, he employs an exceedingly narrow definition of “reasoning” (Hume 2000, Book I, Part iii, sect. ii). For present purposes, by contrast, we are using a broader working gloss of “reasoning,” one not controlled by an ambition to parse out the relative contributions of (the faculty of) reason and of the passions. And about moral reasoning in this broader sense, as responsible thinking about what one ought to do, Hume has many interesting things to say, starting with the thought that *moral* reasoning must involve a double correction of perspective (see [section 2.4](#)) adequately to account for the claims of other people and of the farther future, a double correction that is accomplished with the aid of the so-called “calm passions.”

If we turn from the possibility that perceiving the facts aright will displace moral reasoning to the possibility that applying the correct moral theory will displace — or exhaust — moral reasoning, there are again reasons to be skeptical. One reason is that moral theories do not arise in a vacuum; instead, they develop against a broad backdrop of moral convictions. Insofar as the first potentially reductive strand, emphasizing the importance of perceiving moral facts, has force — and it does have some — it also tends to show that moral theories need to gain support by systematizing or accounting for a wide range of moral facts (Sidgwick 1981). As in most other arenas in which theoretical explanation is called for, the degree of explanatory success will remain partial and open to improvement via revisions in the theory (see [section 2.6](#)). Unlike the natural sciences, however, moral theory is an endeavor that, as John Rawls once put it, is “Socratic” in that it is a subject pertaining to actions “shaped by self-examination” (Rawls 1971, 48f.). Accordingly, that which is to be explained by moral theory is arguably less independent of revisions in provisionally accepted theories than are the data in some other domains. By the same token — and this is the present point — a moral theory is subject to being overturned because it generates concrete implications that do not sit well with us on due reflection. This being so, and granting the great complexity of the moral terrain, it seems highly unlikely that we will ever generate a moral theory on the basis of which we can serenely and confidently proceed in a deductive way to generate answers to what we ought to do in all concrete cases. This conclusion is reinforced by a second consideration, namely that insofar as a moral theory is faithful to the complexity of the moral phenomena, it will contain within it many possibilities for conflicts among its own elements. Even if it does deploy some priority rules, these are unlikely to be able to cover all contingencies. Hence, some moral reasoning that goes beyond the deductive application of the correct theory is bound to be needed.

In short, a sound understanding of moral reasoning will not take the form of reducing it to one of the other two levels of moral philosophy identified above. Neither the demand to attend to the moral facts nor the directive to apply the correct moral theory exhausts or sufficiently describes moral reasoning.

1.4 Gaining Moral Insight from Studying Moral Reasoning

In addition to posing philosophical problems in its own right, moral reasoning is of interest on account of its implications for moral facts and moral theories. Accordingly, attending to moral reasoning will often be useful to those whose real interest is in determining the right answer to some concrete moral problem or in arguing for or against some moral theory. The characteristic ways we attempt to work through a given sort of moral quandary can be just as revealing about our considered approaches to these matters as are any bottom-line judgments we may characteristically come to. Further, we may have firm, reflective convictions about how a given class of problems is best tackled, deliberately, even when we remain in doubt about what should be done. In such cases, attending to the modes of moral reasoning that we characteristically accept can usefully expand the set of moral information from which we start, suggesting ways to structure the competing considerations.

Facts about the nature of moral inference and moral reasoning may have important direct implications for moral theory. For instance, it might be taken to be a condition of adequacy of any moral theory that it play a practically useful role in our efforts at self-understanding and deliberation. If this condition is accepted, then any moral theory that would require agents to engage in abstruse or difficult reasoning may be inadequate for that reason. J.S. Mill and R. M. Hare (1981) attempt to meet this difficulty by, in quite different ways, positing two different levels of moral thinking.

Accordingly, the close relations between moral reasoning, the moral facts, and moral theory do not reductively eliminate moral reasoning as a topic of interest. To the contrary, because moral reasoning has important implications about moral facts and moral theories, these close relations lend additional interest to the topic of moral reasoning.

1.5 How Distinct is Moral Reasoning from Practical Reasoning in General?

The final threshold question is whether moral reasoning is truly distinct from practical reasoning more generally understood. (The question of whether moral reasoning, even if practical, is structurally distinct from theoretical reasoning that simply proceeds from a proper recognition of the moral facts has already been implicitly addressed and tentatively answered in the affirmative.) In addressing this final question, it is difficult to overlook the way different moral theories project quite different models of moral reasoning — again a link that might be pursued by the moral philosopher seeking leverage in either direction. For instance, Aristotle's views might be as follows: a quite general account can be given of practical reasoning, which includes selecting means to ends and determining the constituents of a desired activity. The difference between the reasoning of a vicious person and that of a virtuous person differs not at all in its structure, but only in its content, for the virtuous person pursues true goods, and the vicious person gets side-tracked by apparent ones. To be sure, the virtuous person may be able to achieve a greater integration of his or her ends via practical reasoning (because of the way the various virtues cohere), but this is a difference in the result of practical reasoning and not in its structure. At an opposite extreme, Kant's categorical imperative has been taken to generate an approach to practical reasoning (via a "typic of practical judgment") that is distinctive from other practical reasoning both in the range of considerations it addresses and its structure (Nell 1975). Whereas prudential practical reasoning, on Kant's view, aims to maximize one's happiness, moral reasoning addresses the potential universalizability of the maxims — roughly, the intentions — on which one acts. Views intermediate between Aristotle's and Kant's in this respect include Hare's utilitarian view and Aquinas' natural-law view. On Hare's view, just as an ideal prudential agent applies maximizing rationality to his or her own preferences, an ideal moral agent's reasoning applies maximizing rationality to the set of everyone's preferences that its archangelic capacity for sympathy has enabled it to internalize (Hare 1981). Thomistic, natural-law views share the Aristotelian view about the general unity of practical reasoning in pursuit of the good, rightly or wrongly conceived, but add that practical reason, in addition to demanding that we pursue the fundamental human goods, also, and distinctly, demands that we not attack these goods. In this way, natural-law views incorporate some distinctively moral structuring — such as the distinctions between doing and allowing and the so-called [doctrine of double effect's](#) distinction between intending as a means and accepting as a by-product — within a unified account of practical reasoning (see entry on the [natural law tradition in ethics](#)). In light of this diversity of views about the relation between moral reasoning and practical or prudential reasoning, a general account of moral reasoning that does not want to presume the correctness of a definite moral theory will do well to remain agnostic on the question of how moral reasoning relates to non-moral practical reasoning.

2. General Philosophical Questions about Moral Reasoning

To be sure, most great philosophers who have addressed the nature of moral reasoning were far from agnostic about the content of the correct moral theory, and developed their reflections about moral reasoning in support of or in derivation from their moral theory. Nonetheless, contemporary discussions that are somewhat agnostic about the content of moral theory have arisen around important and controversial aspects of moral reasoning. We may group these around the following seven questions:

1. How do relevant considerations get taken up in moral reasoning?
2. Is it essential to moral reasoning for the considerations it takes up to be crystallized into, or ranged under, principles?
3. How do we sort out which moral considerations are most relevant?
4. In what ways do motivational elements shape moral reasoning?
5. What is the best way to model the kinds of conflicts among considerations that arise in moral reasoning?
6. Does moral reasoning include learning from experience and changing one's mind?
7. How can we reason, morally, with one another?

The remainder of this article takes up these seven questions in turn.

2.1 Moral uptake

One advantage to defining "reasoning" capaciously, as here, is that it helps one recognize that the processes whereby we come to be concretely aware of moral issues are integral to moral reasoning as it might more narrowly be understood. Recognizing moral issues when they arise requires a highly trained set of capacities and a broad range of emotional attunements. Philosophers of the moral sense school of the 17th and 18th centuries stressed innate emotional propensities, such as sympathy with other humans. Classically influenced virtue theorists, by contrast, give more importance to the training of perception and the emotional growth that must accompany it. Among contemporary philosophers working in empirical ethics there is a similar divide, with some arguing that we process situations using

an innate moral grammar (Mikhail 2011) and some emphasizing the role of emotions in that processing (Haidt 2001, Greene 2007). For the moral reasoner, a crucial task for our capacities of moral recognition is to mark out certain features of a situation as being morally salient. Sartre's student, for instance, focused on the competing claims of his mother and the Free French, giving them each an importance to his situation that he did not give to eating French cheese or wearing a uniform. To say that certain features are marked out as morally salient is not to imply that the features thus singled out answer to the terms of some general principle or other: we will come to the question of particularism, below. Rather, it is simply to say that recognitional attention must have a selective focus.

What will be counted as a moral issue or difficulty, in the sense requiring moral agents' recognition, will again vary by moral theory. Not all moral theories would count filial loyalty and patriotism as moral duties. It is only at great cost, however, that any moral theory could claim to do without a layer of moral thinking involving situation-recognition. A calculative sort of utilitarianism, perhaps, might be imagined according to which there is no need to spot a moral issue or difficulty, as every choice node in life presents the agent with the same, utility-maximizing task. Perhaps Jeremy Bentham held a utilitarianism of this sort. For the more plausible utilitarianisms mentioned above, however, such as Mill's and Hare's, the agent does not always calculate afresh, but must instead be alive to the possibility that because the ordinary "landmarks and direction posts" lead one astray in the situation at hand, one must make recourse to a more direct and critical mode of moral reasoning. Recognizing whether one is in one of those situations thus becomes the principal recognitional task for the utilitarian agent. (Whether this task can be suitably confined, of course, has long been one of the crucial questions about whether such indirect forms of utilitarianism, attractive on other grounds, can prevent themselves from collapsing into a more Benthamite, direct form: cf. Brandt 1979.)

Note that, as we have been describing moral uptake, we have not implied that what is perceived is ever a moral fact. Rather, it might be that what is perceived is some ordinary, descriptive feature of a situation that is, for whatever reason, morally relevant. An account of moral uptake will interestingly impinge upon the metaphysics of moral facts, however, if it holds that moral facts can be perceived. Importantly intermediate, in this respect, is the set of judgments involving so-called "thick" evaluative concepts — for example, that someone is callous, boorish, just, or brave. These do not invoke the supposedly "thinner" terms of overall moral assessment, "good," or "right." Yet they are not innocent of normative content, either. Plainly, we do recognize callousness when we see clear cases of it. Plainly, too — whatever the metaphysical implications of the last fact — our ability to describe our situations in these thick normative terms is crucial to our ability to reason morally.

It is debated how closely our abilities of moral discernment are tied to our moral motivations. For Aristotle and many of his ancient successors, the two are closely linked, in that someone not brought up into virtuous motivations will not see things correctly. For instance, cowards will overestimate dangers, the rash will underestimate them, and the virtuous will perceive them correctly (*Eudemian Ethics* 1229b23–27). By the Stoics, too, having the right motivations was regarded as intimately tied to perceiving the world correctly; but whereas Aristotle saw the emotions as allies to enlist in support of sound moral discernment, the Stoics saw them as inimical to clear perception of the truth (cf. Nussbaum 2001).

2.2 Moral Principles

That one discerns features and qualities of some situation that are relevant to sizing it up morally does not yet imply that one explicitly or even implicitly employs any general claims in describing it. Perhaps all that one perceives are particularly embedded features and qualities, without saliently perceiving them as instantiations of any types. Sartre's student may be focused on his mother and on the particular plights of several of his fellow Frenchmen under Nazi occupation, rather than on any purported requirements of filial duty or patriotism. Having become aware of some moral issue in such relatively particular terms, he might proceed directly to sorting out the conflict between them. Another possibility, however, and one that we frequently seem to exploit, is to formulate the issue in general terms: "An only child should stick by an otherwise isolated parent," for instance, or "one should help those in dire need if one can do so without significant personal sacrifice." Such general statements would be examples of "moral principles," in a broad sense. (We do not here distinguish between principles and rules. Those who do include Dworkin 1978 and Gert 1998.)

We must be careful, here, to distinguish the issue of whether principles commonly play an implicit or explicit role in moral reasoning, including well-conducted moral reasoning, from the issue of whether principles necessarily figure as part of the basis of moral truth. The latter issue is best understood as a metaphysical question about the nature and basis of moral facts. What is currently known as [moral particularism](#) is the view that there are no defensible moral principles and that moral reasons, or well-grounded moral facts, can exist independently of any basis in a general principle. A contrary view holds that moral reasons are necessarily general, whether because the sources of their justification are all general or because a moral claim is ill-formed if it contains particularities. But whether principles play a useful role in moral reasoning is certainly a different question from whether principles play a necessary role in

accounting for the ultimate truth-conditions of moral statements. Moral particularism, as just defined, denies their latter role. Some moral particularists seem also to believe that moral particularism *implies* that moral principles cannot soundly play a useful role in reasoning. This claim is disputable, as it seems a contingent matter whether the relevant particular facts arrange themselves in ways susceptible to general summary and whether our cognitive apparatus can cope with them at all without employing general principles. Although the metaphysical controversy about moral particularism lies largely outside our topic, we will revisit it in [section 2.5](#), in connection with the weighing of conflicting reasons.

With regard to moral reasoning, while there are some self-styled “anti-theorists” who deny that abstract structures of linked generalities are important to moral reasoning (Clarke, et al. 1989), there are few who argue that moral reasoning can be well conducted without any appeal to moral principles. Thus, neo-Aristotelians like Nussbaum who emphasize the importance of “finely tuned and richly aware” particular discernment also regard that discernment as being guided by a set of generally describable virtues whose general descriptions will come into play in at least some kinds of cases (Nussbaum 1990). “Situation ethicists” of the last generation (e.g. Fletcher 1997) emphasized the importance of taking into account a wide range of circumstantial differentiae, but against the background of some general principles whose application the differentiae help sort out. Feminist ethicists influenced by Carol Gilligan’s path breaking work on moral development emphasize the importance of the kind of care and discernment that are salient and well-developed by people immersed in particular relationships (Held 1995); but this emphasis is consistent with such general principles as “one ought to be sensitive to the wishes of one’s friends.” Again, if we distinguish the question of whether principles are useful in responsibly-conducted moral thinking from the question of whether moral reasons ultimately all derive from general principles, and concentrate our attention solely on the former, we will see that some of the opposition to general moral principles melts away.

It should be noted that we have been using a weak notion of generality, here. It is contrasted only with the kind of strict particularity that comes with indexicals and proper names. General statements or claims — ones that contain no such particular references — are not necessarily universal generalizations, making an assertion about *all* cases of the mentioned type. Thus, “one should normally help those in dire need” is a general principle, in this weak sense. Possibly, such logically loose principles would be obfuscatory in the context of a metaphysical attempt to reconstruct the truth-conditions of moral statements. Clearly, such logically loose principles would be useless in any attempt to generate a deductively tight “practical syllogism.” In our day-to-day, non-deductive reasoning, however, such logically loose principles appear to be quite useful. (Recall that we are understanding “reasoning” quite broadly, as responsibly conducted thinking: nothing in this understanding of reasoning suggests any uniquely privileged place for deductive inference: cf. Harman 1986. For more on defeasible or “default” principles, see [section 2.5](#).)

In this terminology, establishing that general principles are essential to moral reasoning leaves open the further question whether logically tight, or exceptionless, principles are also essential to moral reasoning. Certainly, much of our actual moral reasoning seems to be driven by attempts to recast or reinterpret principles so that they can be *taken to be* exceptionless. Adherents and inheritors of the natural-law tradition in ethics (e.g. Donagan 1977) are particularly supple defenders of exceptionless moral principles, as they are able to avail themselves not only of a well-thought-through casuistry but also of a wide array of subtle — some would say overly subtle — distinctions, such as those mentioned above between doing and allowing and between intending as a means and accepting as a byproduct.

A related role for a strong form of generality in moral reasoning comes from the Kantian thought that one’s moral reasoning must counter one’s tendency to make exceptions for oneself. Accordingly, Kant holds, as we have noted, that we must ask whether the maxims of our actions can serve as universal laws. As most contemporary readers understand this demand, it requires that we engage in a kind of hypothetical generalization across agents, and ask about the implications of everybody acting that way in those circumstances. The grounds for developing Kant’s thought in this direction have been well explored (e.g. Nell 1975, Korsgaard 1996, Engstrom 2009). The importance and the difficulties of such a hypothetical generalization test in ethics are thoroughly explored in (Singer 1961).

2.3 Sorting Out Which Considerations Are Most Relevant

Whether or not moral considerations need the backing of general principles, we must expect situations of action to present us with multiple moral considerations. In addition, of course, these situations will also present us with a lot of information that is not morally relevant. On any realistic account, a central task of moral reasoning is to sort out relevant considerations from irrelevant ones, as well as to determine which are especially relevant and which only slightly so. That a certain woman is Sartre’s student’s *mother* seems arguably to be a morally relevant fact; what about the fact (supposing it is one) that she has no other children to take care of her? Addressing the task of sorting what is morally relevant from what is not, some philosophers have offered general accounts of moral relevant features. Others

have given accounts of how we sort out which of the relevant features are *most* relevant, a process of thinking that generally goes by the name of “casuistry.”

Before we look at ways of sorting out which features are morally relevant or *most* morally relevant, it may be useful to note a prior step taken by some casuists, which was to attempt to set out a schema that would capture *all* of the features of an action or proposed action. The Roman Catholic casuists of the middle ages did so by drawing on Aristotle's categories. Accordingly, they asked, where, when, why, how, by what means, to whom, or by whom the action in question is to be done or avoided (see Jonsen and Toulmin 1988). The idea was that complete answers to these questions would contain all of the features of the action, of which the morally relevant ones would be a subset. Although metaphysically uninteresting, the idea of attempting to list all of an action's features in this way represents a distinctive — and extreme — heuristic for moral reasoning.

Turning to the morally relevant features, one of the most developed accounts is given by Gert. He develops a list of features relevant to whether the violation of a moral rule should be generally allowed. Given the designed function of Gert's list, it is natural that most of his morally relevant features make reference to the set of moral rules he defends. Accordingly, some of Gert's distinctions between dimensions of relevant features reflect controversial stances in moral theory. For example, one of the dimensions is whether “the violation [is] done intentionally or only knowingly” (Gert 1998, 234) — a distinction that those who reject the doctrine of double effect would not find relevant.

In deliberating about what we ought, morally, do to, we also often attempt to figure out which considerations are *most* relevant. To take an issue mentioned above: Are surrogate motherhood contracts more akin to agreements with babysitters (clearly acceptable) or to agreements with prostitutes (not clearly so)? That is, which feature of surrogate motherhood is more relevant: that it involves a contract for child-care services or that it involves payment for the intimate use of the body? Both in such relatively novel cases and in more familiar ones, reasoning by analogy plays a large role in ordinary moral thinking. When this reasoning by analogy starts to become systematic — a social achievement that requires some historical stability and reflectiveness about what are taken to be moral norms — it begins to exploit comparison to cases that are “paradigmatic,” in the sense of being taken as settled. Within such a stable background, a system of casuistry can develop that lends some order to the appeal to analogous cases. To use an analogy: the availability of a widely accepted and systematic set of analogies and the availability of what are taken to be moral norms may stand to one another as chicken does to egg: each may be an indispensable moment in the genesis of the other.

Casuistry, thus understood, is an indispensable aid to moral reasoning. At least, that it is would follow from conjoining two features of the human moral situation mentioned above: the multifariousness of moral considerations that arise in particular cases and the need and possibility for employing moral principles in sound moral reasoning. We require moral judgment, not simply a deductive application of principles or a particularist bottom-line intuition about what we should do. This judgment must be responsible to moral principles yet cannot be straightforwardly derived from them. Accordingly, our moral judgment is greatly aided if it is able to rest on the sort of heuristic support that casuistry offers. Thinking through which of two analogous cases provides a better key to understanding the case at hand is a useful way of organizing our moral reasoning, and one on which we must continue to depend. If we lack the kind of broad consensus on a set of paradigm cases on which the Renaissance Catholic or Talmudic casuists could draw, our casuistic efforts will necessarily be more controversial and tentative than theirs; but we are not wholly without settled cases from which to work. Indeed, as Jonsen and Toulmin suggest at the outset of their thorough explanation and defense of casuistry, the depth of disagreement about moral theories that characterizes a pluralist society may leave us having to rest comparatively *more* weight on the cases about which we can find agreement than did the classic casuists (Jonsen and Toulmin 1988).

Despite the long history of casuistry, there is little that can usefully be said about how one ought to reason about competing analogies. In the law, where previous cases have precedential importance, more can be said. As Sunstein notes (Sunstein 1996, chap. 3), the law deals with particular cases, which are always “potentially distinguishable” (72); yet the law also imposes “a requirement of practical consistency” (67). This combination of features makes reasoning by analogy particularly influential in the law, for one must decide whether a given case is more like one set of precedents or more like another. Since the law must proceed even within a pluralist society such as ours, Sunstein argues, we see that analogical reasoning can go forward on the basis of “incompletely theorized judgments” or of what Rawls calls an “overlapping consensus” (Rawls 1996). That is, although a robust use of analogous cases depends, as we have noted, on some shared background agreement, this agreement need not extend to all matters or all levels of individuals' moral thinking. Accordingly, although in a pluralist society we may lack the kind of comprehensive normative agreement that made the high casuistry of Renaissance Christianity possible, the path of the law suggests that normatively forceful, case-based, analogical reasoning can still go on.

Reasoning by appeal to cases is also a favorite mode of some recent moral philosophers. Since our focus here is not on the methods of moral theory, we do not need to go into any detail in comparing different ways in which philosophers wield cases for and against alternative moral theories. There is, however, an important and broadly applicable point worth making about ordinary reasoning by reference to cases that emerges most clearly from the philosophical use of such reasoning. Philosophers often feel free to imagine cases, often quite unlikely ones, in order to attempt to isolate relevant differences. An infamous example is a pair of cases offered by James Rachels to cast doubt on the moral significance of the distinction between killing and letting die. In both cases, there is at the outset a boy in a bathtub and a greedy older cousin downstairs who will inherit the family manse if and only if the boy predeceases him (Rachels 1975). In Case A, the cousin hears a thump, runs up to find the boy unconscious in the bath, and reaches out to turn on the tap so that the water will rise up to drown the boy. In Case B, the cousin hears a thump, runs up to find the boy unconscious in the bath with the water running, and decides to sit back and do nothing until the boy drowns. Since there is surely no moral difference between these cases, Rachels argued, the general distinction between killing and letting die is undercut. "Not so fast!" is the well-justified reaction (cf. Beauchamp 1979). Just because a factor is morally relevant in a certain way in some contexts does not mean that it either is or must be relevant in the same way or to the same degree in other contexts. Shelly Kagan has dubbed the failure to take account of this fact of contextual interaction when wielding comparison cases the "additive fallacy" (1988). Kagan concludes from this that the reasoning of moral theorists must depend upon some theory that helps us anticipate and account for ways in which factors will interact in various contexts. A parallel lesson, reinforcing what we have already observed in connection with casuistry proper, would apply for moral reasoning in general: reasoning from cases must at least implicitly rely upon a set of organizing judgments or beliefs, of a kind that would, on some understandings, count as a moral "theory." If this is correct, it provides another kind of reason to think that moral considerations could be crystallized into principles that make manifest the organizing structure involved.

2.4 Moral Reasoning and Moral Psychology

We are concerned here with moral reasoning as a species of practical reasoning — reasoning directed to deciding what to do and, if successful, issuing in an intention. But how can such practical reasoning succeed? How can moral reasoning hook up with motivationally effective psychological states so as to have this kind of causal effect? "Moral psychology," — as the philosophical study of intention and action is called — has a lot to say to such questions, both in its traditional, *a priori* form and its newly popular empirical form. In addition, the conclusions of moral psychology can have substantive moral implications, for it may be reasonable to assume that if there are deep reasons that a given type of moral reasoning *cannot* be practical, then any principles that demand such reasoning are unsound. In this spirit, Samuel Scheffler has explored "the importance for moral philosophy of some tolerably realistic understanding of human motivational psychology" (Scheffler 1992, 8) and Peter Railton has developed the idea that certain moral principles might generate a kind of "alienation" (Railton 1984). In short, we may be interested in what makes practical reasoning psychologically possible both for its own sake and as a way of working out some of the content of moral theory.

The issue of psychological possibility is an important one for all kinds of practical reasoning (cf. Audi 1989). In morality, it is especially pressing, as morality often asks individuals to depart from satisfying their own interests. As a result, it may appear that moral reasoning's practical effect could not be explained by a simple appeal to the initial motivations that shape or constitute someone's interests, in combination with a requirement, like that mentioned above, to will the necessary means to one's ends. Morality, it may seem, instead requires individuals to act on ends that may not be part of their "motivational set," in the terminology of Williams 1981. How can moral reasoning lead people to do that? The question is a traditional one. Plato's *Republic* answered that the appearances are deceiving, and that acting morally is, in fact, in the enlightened self-interest of the agent. Kant, in stark contrast, held that our transcendent capacity to act on our conception of a practical law enables us to set ends and to follow morality even when doing so sharply conflicts with our interests. Many other answers have been given. In recent times, philosophers have defended what has been called "internalism" about morality, which claims that there is a necessary conceptual link between agents' moral judgment and their motivation. Michael Smith, for instance, puts the claim as follows (Smith 1994, 61):

If an agent judges that it is right for her to Φ in circumstances C , then either she is motivated to Φ in C or she is practically irrational.

Even this defeasible version of moral judgment internalism may be too strong; but instead of pursuing this issue further, let us turn to a question more internal to moral reasoning. (For more on the issue of moral judgment internalism, see [moral motivation](#).)

The traditional question we were just glancing at picks up when moral reasoning is done. Supposing that we have some moral conclusion, it asks how agents can be motivated to go along with it. A different question about the

intersection of moral reasoning and moral psychology, one more immanent to the former, concerns how motivational elements shape the reasoning process itself.

A powerful philosophical picture of human psychology, stemming from Hume, insists that beliefs and desires are distinct existences (Hume 2000, Book II, part iii, sect. iii; cf. Smith 1994, 7). This means that there is always a potential problem about how reasoning, which seems to work by concatenating beliefs, links up to the motivations that desire provides. The paradigmatic link is that of instrumental action: the desire to Ψ links with the belief that by Φ ing in circumstances C one will Ψ . Accordingly, philosophers who have examined moral reasoning within an essentially Humean, belief-desire psychology have sometimes accepted a constrained account of moral reasoning. Hume's own account exemplifies the sort of constraint that is involved. As Hume has it, the calm passions support the dual correction of perspective constitutive of morality, alluded to above. Since these calm passions are seen as competing with our other passions in essentially the same motivational coinage, as it were, our passions limit the reach of moral reasoning.

An important step away from a Humean moral psychology is taken if one recognizes the existence of what Rawls has called "principle-dependent desires" (Rawls 1996, 82-83; Rawls 2000, 46-47). These are desires whose objects cannot be characterized without reference to some rational or moral principle. An important special case of these is that of "conception-dependent desires," in which the principle-dependent desire in question is seen by the agent as belonging to a broader conception, and as important on that account (Rawls 1996, 83-84; Rawls 2000, 148-152). For instance, conceiving of oneself as a citizen, one may desire to bear one's fair share of society's burdens. Although it may look like any content, including this, may substitute for Ψ in the Humean conception of desire, and although Hume set out to show how moral sentiments such as pride could be explained in terms of simple psychological mechanisms, his influential empiricism actually tends to restrict the possible content of desires. Introducing principle-dependent desires thus marks a sharp departure from a Humean psychology. As Rawls remarks, if "we may find ourselves drawn to the conceptions and ideals that both the right and the good express ... , [h]ow is one to fix limits on what people might be moved by in thought and deliberation and hence may act from?" (1996, 85). While Rawls developed this point by contrasting Hume's moral psychology with Kant's, the same basic point is also made by neo-Aristotelians (e.g., McDowell 1998).

The introduction of principle-dependent desires bursts any would-be naturalist limit on their content; nonetheless, some philosophers hold that this notion remains too beholden to an essentially Humean picture to be able to capture the idea of a moral commitment. Desires, it may seem, remain motivational items that compete on the basis of strength. Saying that one's desire to be just may be outweighed by one's desire for advancement may seem to fail to capture the thought that one has a commitment — even a non-absolute one — to justice. Sartre designed his example of the student torn between staying with his mother and going to fight with the Free French so as to make it seem implausible that he ought to decide simply by determining which he more strongly wanted to do.

One way to get at the idea of commitment is to emphasize our capacity to reflect about what we want. By this route, one might distinguish, in the fashion of Harry Frankfurt, between the strength of our desires and "the importance of what we care about" (Frankfurt 1988). Although this idea is evocative, it provides relatively little insight into *how* it is that we thus reflect. Another way to model commitment is to take it that our intentions operate at a level distinct from our desires, structuring what we are willing to reconsider at any point in our deliberations (e.g. Bratman 1999). While this two-level approach offers some advantages, it is limited by its concession of a kind of primacy to the unreconstructed desires at the unreflective level. A more integrated approach might model the psychology of commitment in a way that reconceives the nature of desire from the ground up. One attractive possibility is to return to the Aristotelian conception of desire as being for the sake of some good or apparent good (cf. Richardson 2004). On this conception, the end for the sake of which an action is done plays an important regulating role, indicating, in part, what one will *not* do (Richardson 2012). Reasoning about final ends accordingly has a distinctive character (see Richardson 1994, Schmidtz 1995). Whatever the best philosophical account of the notion of a commitment— for another alternative, see (Tiberius 2000)— much of our moral reasoning does seem to involve expressions of and challenges to our commitments.

Recent experimental work, employing both survey instruments and brain imaging technologies, has allowed philosophers to approach questions about the psychological basis of moral reasoning from novel angles. The initial brain data seems to show that individuals with damage to the pre-frontal lobes tend to reason in more straightforwardly consequentialist fashion than those without such damage (Koenigs et al. 2007). Some theorists take this finding as tending to confirm that fully competent human moral reasoning goes beyond a simple weighing of pros and cons to include assessment of moral constraints (e.g., Wellman & Miller 2008, Young & Saxe 2008). Others, however, have argued that the emotional responses of the prefrontal lobes interfere with the more sober and sound, consequentialist-style reasoning of the other parts of the brain (e.g. Greene 2007). The survey data reveals or confirms, among other

things, interesting, normatively loaded asymmetries in our attribution of such concepts as responsibility and causality (Knobe 2006). It also reveals that many of moral theory's most subtle distinctions, such as the distinction between an intended means and a foreseen side-effect, are deeply built into our psychologies, being present cross-culturally and in young children, in a way that suggests to some the possibility of an innate "moral grammar" (Mikhail 2011).

A final question about the connection between moral motivation and moral reasoning is whether someone without the right motivational commitments can reason well, morally? On Hume's official, narrow conception of reasoning, which essentially limits it to tracing empirical and logical connections, the answer would be yes. The vicious person could trace the causal and logical implications of acting in a certain way just as a virtuous person could. The only difference would be practical, not rational: the two would not act in the same way. Note, however, that the Humean's affirmative answer depends on departing from the working definition of "moral reasoning" used in this article, which casts it as a species of practical reasoning. Interestingly, Kant can answer "yes" while still casting moral reasoning as practical. On his view in the *Groundwork* and the *Critique of Practical Reason*, reasoning well, morally, does not depend on any prior motivational commitment, yet remains practical reasoning. That is because he thinks the moral law can itself generate motivation. (Kant's *Metaphysics of Morals* and *Religion* offer a more complex psychology.) For Aristotle, by contrast, an agent whose motivations are not virtuously constituted will systematically misperceive what is good and what is bad, and hence will be unable to reason excellently. The best reasoning that a vicious person is capable of, according to Aristotle, is a defective simulacrum of practical wisdom that he calls "cleverness" (*Nicomachean Ethics* 1144a25).

2.5 Modeling Conflicting Moral Considerations

Moral considerations often conflict with one another. So do moral principles and moral commitments. Assuming that filial loyalty and patriotism are moral considerations, then Sartre's student faces a moral conflict. Recall that it is one thing to model the metaphysics of morality or the truth conditions of moral statements and another to give an account of moral reasoning. In now looking at conflicting considerations, our interest here remains with the latter and not the former. Our principal interest is in ways that we need to structure or think about conflicting considerations in order to negotiate well our reasoning involving them.

One influential building-block for thinking about moral conflicts is W. D. Ross's notion of a "*prima facie* duty" (Ross 1988). Although Ross gave various conflicting glosses of this notion, it entered the literature of moral theory, which now generally interprets "*prima facie*" in contrast to "all things considered." One has a *prima facie* duty to do some act just in case there is a sufficient basis for concluding that it is one's duty to do it. A fuller consideration of the act's features might rebut this conclusion by showing that the features providing reason to ascribe the duty are overridden by other features of the act. Ross described each *prima facie* duty as a "parti-resultant" attribute, obtained by looking at one morally relevant aspect of an act, whereas "being one's [actual] duty" is a "toti-resultant" attribute, obtained by looking together at all of the relevant aspects (28). This suggests that in each case there is, in principle, some function that maps the partial contributions of each *prima facie* duty to the conclusion of actual duty. What might that function be? To Ross's credit, he writes that "for the estimation of the comparative stringency of these *prima facie* obligations no general rules can, so far as I can see, be laid down" (41). Accordingly, a second strand in Ross simply emphasizes, following Aristotle, the need for practical judgment by those who have been brought up into virtue (42).

How might considerations of the sort constituted by *prima facie* duties enter our moral reasoning? They might do so explicitly, or only implicitly. There is also a third, still weaker possibility (Scheffler 1992, 32): it might simply be the case that if the agent *had* recognized a *prima facie* duty, he would have acted on it unless he considered it to be overridden. This is a fact about how he *would* have reasoned.

Despite Ross's denial that there is any general method for estimating the comparative stringency of *prima facie* duties, there is a further strand in his exposition that many find irresistible and that tends to undercut this denial. In the very same paragraph in which he states that he sees no general rules for dealing with conflicts, he speaks in terms of "the greatest balance of *prima facie* rightness." This language, together with the idea of "comparative stringency," ineluctably suggests the idea that the mapping function might be the same in each case of conflict and that it might be a quantitative one. On this conception, if there is a conflict between *prima facie* duties, the one that is strongest in the circumstances should be taken to win. Duly cautioned about the additive fallacy (see [section 2.3](#)), we might recognize that the strength of a moral consideration in one set of circumstances cannot be inferred from its strength in other circumstances. Hence, this approach will need still to rely on intuitive judgments in many cases. But this intuitive judgment will be about which *prima facie* consideration is stronger in the circumstances, not simply about what ought to be done. On this approach, a *prima facie* duty becomes what has been called a "*pro tanto*" one — a duty that has a certain force, but may not prevail (see Hurley 1989).

The thought that our moral reasoning either requires or is benefited by a virtual quantitative crutch of this kind has a long pedigree. Philosophical support for this thought involves an idea of practical commensurability. We need to distinguish, here, two kinds of practical commensurability or incommensurability, one defined in metaphysical terms and one in deliberative terms. Each of these forms might be stated evaluatively or deontically. The first, metaphysical sort of value incommensurability is defined directly in terms of what is the case. Thus, to state an evaluative version: two values are metaphysically incommensurable just in case neither is better than the other nor are they equally good (see Chang 1998). Now, the metaphysical incommensurability of values, or its absence, is only loosely linked to how it would be reasonable to deliberate. If all values or moral considerations are metaphysically (that is, in fact) commensurable, still it might well be the case that our access to the ultimate commensurating function is so limited that we would fare ill by proceeding in our deliberations to try to think about which outcomes are “better” or which considerations are “stronger.” We might have no clue about how to measure the relevant “strength.” Conversely, even if metaphysical value incommensurability is common, we might do well, deliberatively, to proceed as if this were not the case, just as we proceed in thermodynamics as if the gas laws obtained in their idealized form. Hence, in thinking about the deliberative implications of incommensurable values, we would do well to think in terms of a definition tailored to the deliberative context. Start with a local, pairwise form. We may say that two options, A and B, are deliberatively commensurable just in case there is some one dimension of value in terms of which, prior to — or logically independently of — choosing between them, it is possible adequately to represent the force of the considerations bearing on the choice.

Philosophers as diverse as Immanuel Kant and John Stuart Mill have argued that unless two options are deliberatively commensurable, in this sense, it is impossible to choose rationally between them. Interestingly, Kant limited this claim to the domain of prudential considerations, recognizing moral reasoning as invoking considerations incommensurable with those of prudence. For Mill, this claim formed an important part of his argument that there must be some one, ultimate “umpire” principle — namely, on his view, the principle of utility. Henry Sidgwick elaborated Mill's argument and helpfully made explicit its crucial assumption, which he called the “principle of superior validity” (Sidgwick 1981; cf. Schneewind 1977). This is the principle that conflict between distinct moral or practical considerations can be rationally resolved only on the basis of some third principle or consideration that is both more general and more firmly warranted than the two initial competitors. From this assumption, one can readily build an argument for the rational necessity not merely of local deliberative commensurability, but of a global deliberative commensurability that, like Mill and Sidgwick, accepts just one ultimate umpire principle (cf. Richardson 1994, chap. 6).

Sidgwick's explicitness, here, is valuable also in helping one see how to resist the demand for deliberative commensurability. Deliberative commensurability is not necessary for proceeding rationally if conflicting considerations can be rationally dealt with in a holistic way that does not involve the appeal to a principle of “superior validity.” That our moral reasoning can proceed holistically is strongly affirmed by Rawls. Rawls's characterizations of the influential ideal of [reflective equilibrium](#) and his related ideas about the nature of justification imply that we can deal with conflicting considerations in less hierarchical ways than imagined by Mill or Sidgwick. Instead of proceeding up a ladder of appeal to some highest court or supreme umpire, Rawls suggests, when we face conflicting considerations “we work from both ends” (Rawls 1999, 18). Sometimes indeed we revise our more particular judgments in light of some general principle to which we adhere; but we are also free to revise more general principles in light of some relatively concrete considered judgment. On this picture, there is no necessary correlation between degree of generality and strength of authority or warrant. That this holistic way of proceeding (whether in building moral theory or in deliberating: cf. Hurley 1989) can be rational is confirmed by the possibility of a form of justification that is similarly holistic: “justification is a matter of the mutual support of many considerations, of everything fitting together into one coherent view” (Rawls 1999, 19, 507). (Note that this statement, which expresses a necessary aspect of moral or practical justification, should not be taken as a definition or analysis thereof.) So there is an alternative to depending, deliberatively, on finding a dimension in terms of which considerations can be ranked as “stronger” or “better” or “more stringent”: one can instead “prune and adjust” with an eye to building more mutual support among the considerations that one endorses on due reflection. If even the desideratum of practical coherence is subject to such re-specification, then this holistic possibility really does represent an alternative to commensuration, as the deliberator, and not some coherence standard, retains reflective sovereignty (Richardson 1994, sec. 26).

Suppose that moral considerations are all commensurable as a matter of ultimate, metaphysical fact, but that our grasp of the actual strength of these considerations is quite poor and subject to systematic distortions. Perhaps some people are much better placed than others to appreciate certain considerations, and perhaps our strategic interactions would cause us to reach suboptimal outcomes if we each pursued our own unfettered judgment of how the overall set of considerations plays out. In such circumstances, there is a strong case for departing from maximizing reasoning without swinging all the way to the holist alternative. This case has been influentially articulated by Joseph Raz, who develops the notion of an “exclusionary reason” to occupy this middle position (Raz 1990).

“An exclusionary reason,” in Raz’s terminology, “is a second order reason to refrain from acting for some reason” (39). A simple example is that of Ann, who is tired after a long and stressful day, and hence has reason not to act on her best assessment of the reasons bearing on a particularly important investment decision that she immediately faces (37). This notion of an exclusionary reason allowed Raz to capture many of the complexities of our moral reasoning, especially as it involves principled commitments, while conceding that, at the first order, all practical reasons might be commensurable. Raz’s early strategy for reconciling commensurability with complexity of structure was to limit the claim that reasons are comparable with regard to strength to reasons of a given order. First-order reasons compete on the basis of strength; but conflicts between first- and second-order reasons “are resolved not by the strength of the competing reasons but by a general principle of practical reasoning which determines that exclusionary reasons always prevail” (40).

If we take for granted this “general principle of practical reasoning,” why should we recognize the existence of any exclusionary reasons, which by definition prevail independently of any contest of strength? Raz’s principal answer to this question shifts from the metaphysical domain of the strengths that various reasons “have” to the epistemically limited viewpoint of the deliberator. As in Ann’s case, we can see in certain contexts that a deliberator is likely to get things wrong if he or she acts on his or her perception of the first-order reasons. Second-order reasons indicate, with respect to a certain range of first-order reasons, that the agent “must not act *for* those reasons” (185). The broader justification of an exclusionary reason, then, can consistently be put in terms of the commensurable first-order reasons. Such a justification can have the following form: “Given this agent’s deliberative limitations, the balance of first-order reasons will likely be better conformed with if he or she refrains from acting for certain of those reasons.”

Raz’s account of exclusionary reasons might be used to reconcile ultimate commensurability with the structured complexity of our moral reasoning. Whether such an attempt could succeed would depend, in part, on the extent to which we have an actual grasp of first-order reasons, conflict among which can be settled solely on the basis of their comparative strength. Our consideration, above, of casuistry, the additive fallacy, and deliberative incommensurability may combine to make it seem that only in rare pockets of our practice do we have a good grasp of first-order reasons, if these are defined, à la Raz, as competing only in terms of strength. If that is right, then we will almost always have good exclusionary reasons to reason on some other basis than in terms of the relative strength of first-order reasons. Under those assumptions, the middle way that Raz’s idea of exclusionary reasons seems to open up would more closely approach the holist’s.

The notion of a moral consideration’s “strength,” whether put forward as part of a metaphysical picture of how first-order considerations interact in fact or as a suggestion about how to go about resolving a moral conflict, should not be confused with the bottom-line determination of whether one consideration, and specifically one duty, overrides another. For example, in Ross’s oft-cited case of conflicting *prima facie* duties, someone who has promised to be somewhere at a certain time passes a boy drowning in a pond, and must choose between saving a life and keeping a promise to meet someone for lunch. (Ross chose the case to illustrate that an “imperfect” duty, or a duty of commission, can override a strict, prohibitive duty.) Ross’s assumption is that all well brought-up people would agree, in this case, that the duty to save a life overrides the duty to keep a promise. We may take it, if we like, that this judgment implies that we consider the duty to save a life, here, to be stronger than the duty to keep the promise; but in fact this claim about relative strength adds nothing to our understanding of the situation. We do not reach our practical conclusion in this case *by* determining that the duty to save the boy’s life is stronger. The statement that this duty is here stronger is simply a way to embellish the conclusion that of the two *prima facie* duties that here conflict, it is the one that states the all-things-considered duty. To be “overridden” is just to be a *prima facie* duty that fails to state an all-things-considered duty because another *prima facie* duty that conflicts with it does do so. Hence, the judgment that one duty overrides another can be understood just in terms of its deontic implications and without reference to considerations of strength. To confirm this, note that we can say, “As a matter of fidelity, we ought to keep the promise; as a matter of beneficence, we ought to save the life; we cannot do both; and both categories considered we ought to save the life.”

Understanding the notion of one duty overriding another in this way puts us in a position to take up the topic of [moral dilemmas](#). Since this topic is covered in a separate article, here we may simply note one attractive definition of a moral dilemma (see Sinnott-Armstrong 1988) as a situation in which the following are true of a single agent:

1. He ought to do *A*.
2. He ought to do *B*.
3. He cannot do both *A* and *B*.
4. (1) does not override (2) and (2) does not override (1).

This way of defining moral dilemmas distinguishes them from the kind of moral conflict, such as Ross’s promises/drowning case, in which one of the duties is overridden by the other. Arguably, Sartre’s student faces a moral

dilemma. Making sense of a situation in which neither of two duties overrides the other is easier if deliberative commensurability is denied. Whether moral dilemmas are possible will depend crucially on whether “ought” implies “can” and whether any pair of duties such as those comprised by (1) and (2) implies a single, “agglomerated” duty that the agent do both *A* and *B*. If either of these purported principles of the logic of duties is false, then moral dilemmas are possible.

Jonathan Dancy has well highlighted a kind of contextual variability in moral reasons that has come to be known as “reasons holism”: “a feature that is a reason in one case may be no reason at all, or an opposite reason, in another” (Dancy 2004). To adapt one of his examples: while there is often moral reason not to lie, when playing liar’s poker one generally ought to lie; otherwise, one will spoil the game (cf. Dancy 1993, 61). Dancy argues that reasons holism supports moral particularism of the kind discussed in [section 2.2](#), according to which there are no defensible moral principles. Taking this conclusion seriously would radically affect how we conducted our moral reasoning. The argument’s premise of holism has been challenged (e.g., Audi 2004, McKeever & Ridge 2006). Philosophers have also challenged the inference from reasons holism to particularism in various ways. Mark Lance and Margaret Olivia Little (2007) have done so by exhibiting how defeasible generalizations, in ethics and elsewhere, depend *systematically* on context. We can work with them, they suggest, by utilizing a skill that is similar to the skill of discerning morally salient considerations, namely the skill of discerning relevant similarities among possible worlds. More generally, John F. Horty has developed a logical and semantic account according to which reasons are defaults and so behave holistically, but there are nonetheless general principles that explain how they behave (Horty 2012). And Mark Schroeder has argued that our holistic views about reasons are actually better explained by supposing that there are general principles (Schroeder 2011).

2.6 Moral Learning and the Revision of Moral Views

If we have any moral knowledge, whether concerning general moral principles or concrete moral conclusions, it is surely very imperfect. What moral knowledge we are capable of will depend, in part, on what sorts of moral reasoning we are capable of. Although some moral learning may result from the theoretical work of moral philosophers and theorists, much of what we learn with regard to morality surely arises in the practical context of deliberation about new and difficult cases. This deliberation might be merely instrumental, concerned only with settling on means to moral ends, or it might be concerned with settling those ends. There is no special problem about learning what conduces to morally obligatory ends: that is an ordinary matter of empirical learning. But by what sorts of process can we learn which ends are morally obligatory, or which norms morally required? That is, how is strictly moral learning possible via moral reasoning?

Much of what was said above with regard to moral uptake applies again in this context, with approximately the same degree of dubiousness or persuasiveness. If there is a role for moral perception or for emotions in agents’ becoming aware of moral considerations, these may function also to guide agents to new conclusions. For instance, it is conceivable that our capacity for outrage is a relatively reliable detector of wrong actions, even novel ones, or that our capacity for pleasure is a reliable detector of actions worth doing, even novel ones. (For a thorough defense of the latter possibility, which intriguingly interprets pleasure as a judgment of value, see Millgram 1997.) Perhaps these capacities for emotional judgment enable strictly moral learning in roughly the same way that chess-players’ trained sensibilities enable them to recognize the threat in a previously unencountered situation on the chessboard (Lance and Tanesini 2004). That is to say, perhaps our moral emotions play a crucial role in the exercise of a skill whereby we come to be able to articulate moral insights that we have never before attained. Perhaps competing moral considerations interact in contextually specific and complex ways much as competing chess considerations do. If so, it would make sense to rely on our emotionally-guided capacities of judgment to cope with complexities that we cannot model explicitly, but also to hope that, once having been so guided, we might in retrospect be able to articulate something about the lesson of a well-navigated situation.

A different model of strictly moral learning puts the emphasis on our after-the-fact reactions rather than on any prior, tacit emotional or judgmental guidance: the model of “experiments in living,” to use John Stuart Mill’s phrase (see Anderson 1991). Here, the basic thought is that we can try something and see if “it works.” For this to be an alternative to empirical learning about what causally conduces to what, it must be the case that we remain open as to what we mean by things “working.” In Mill’s terminology, for instance, we need to remain open as to what are the important “parts” of happiness. If we are, then perhaps we can learn by experience what some of them are — that is, what are some of the constitutive means of happiness. These paired thoughts, that our practical life is experimental and that we have no firmly fixed conception of what it is for something to “work,” come to the fore in Dewey’s pragmatist ethics (see esp. Dewey 1967 [1922]). This experimentalist conception of strictly moral learning is brought to bear on moral reasoning in Dewey’s eloquent characterizations of “practical intelligence” as involving a creative and flexible approach to figuring out “what works” in a way that is thoroughly open to rethinking our ultimate aims.

Once we recognize that moral learning is a possibility for us, we can recognize a broader range of ways of coping with moral conflicts than was canvassed in the last section. There, moral conflicts were described in a way that assumed that the set of moral considerations, among which conflicts were arising, was to be taken as fixed. If we can learn, morally, however, then we probably can and should revise the set of moral considerations that we recognize. Often, we do this by re-interpreting some moral principle that we had started with, whether by making it more specific, making it more abstract, or in some other way (cf. Richardson 2000).

2.7 How Can We Reason, Morally, With One Another?

So far, we have mainly been discussing moral reasoning as if it were a solitary endeavor. This is, at best, a convenient simplification. At worst, it is, as Jürgen Habermas has long argued, deeply distorting of reasoning's essentially "dialogical" character (e.g., Habermas, 1984). In any case, it is clear that we often do need to reason morally with one another.

Even if the simplified presentation of the foregoing parts of this article is pernicious, as Habermas might think, it seems nonetheless that the issues and distinctions they develop pertain to collective moral reasoning as well as to individual moral reasoning. We can ask what we take to be morally salient, whether we appeal to moral principles, and how we deal with conflicting moral considerations. The principal new issues raised by thinking of moral reasoning as collective concern the obstacles that arise from disagreement (this section) and the possibility that differentiated social structures are of intrinsic importance to moral reasoning (the following section).

There is also a threshold question of social structure that is applicable to all cases of collective moral reasoning. Because there is no natural boundary to the "we," collective moral deliberations are often explicitly faced with issues about how the collective should define itself. For instance, consider Canadian citizens who ask themselves, "What, morally, do we owe the indigenous tribes whose ancestors were here before the European settlers arrived?" It will make a crucial difference to how this question is addressed whether "we" includes members of the indigenous tribes or rather (as secessionists among the tribespeople might prefer) does not include them. It may simply be an uncomfortable fact that the reference of the "we" here is unavoidably unstable. It will often be the case that any specification of who "we" are will have morally controversial implications. This fact represents an obvious potential pitfall; but it also represents fruitful possibilities (cf. Vogler 1998). Perhaps the boundaries of the relevant "we" could be defended on substantive moral grounds. For example, an account of democratic deliberation might defend a specification of "the public" by reference to fundamental democratic ideals (cf. Richardson 2002). Habermas argues for extending the deliberating "we" to include all those affected by the matter under discussion (Habermas 2001).

The disagreements that threaten to make collective moral reasoning impossible are generally of two kinds. First, there are simple clashes of claims within a shared moral framework. For instance, suppose a manager had promised to meet with one subordinate who wants to discuss a grievance, but suddenly is confronted by another who is sobbing over some personal crisis and is demanding immediate attention. While all concerned might take both promise-keeping and beneficence to be *prima facie* duties, each subordinate may feel that his or her own claim is more important, and ought to override. In such cases, the effect of self-interest on one's moral beliefs may explain most of the moral disagreement. Famously, Hobbes argued that clashing self-interests yield a generalized [prisoner's dilemma](#) such that each person will have reason to defect from moral arrangements unless a "we" is constructed that has coercive authority. Hobbes's moral psychology was perhaps too narrow in not allowing for the kind of principle-dependent desires noted above, and hence too narrow in the range of routes to moral socialization that it recognized. Still, his views raise an intriguing and very broad set of questions about the relationship between moral reasoning and political power.

In a second type of case, the moral disagreement that threatens the possibility of collective moral reasoning consists in divergent beliefs about what is morally salient and what should be counted as a moral principle. Dramatic examples of this second type of case are familiar from anthropology and from teenagers who suddenly develop moral sensibilities divergent from those of their parents. And of course, the factors of clashing self-interest and divergent moral frameworks can operate together in a given case, as in the controversy between Jews and Moslems in Jerusalem over who has a right to control the Temple Mount/Haram Al-Sharif.

Such deep divergences of world-view need not make moral reasoning between the parties impossible. (In saying that, we must recognize that even simple clashes of self-interest in the absence of any deeper disagreement *can* preclude any effort at joint moral reasoning.) Two elements of an argument for moral reasoning being possible despite such disagreement — both Rawlsian — are already before us: the holism built into the ideal of reflective equilibrium and the notion of overlapping consensus (cf. Richardson 1994, Part V). Insofar as holism is plausible in the deliberation or

reasoning of each of the parties, the moral beliefs of each will not function as rigidly axiomatized systems. Rather, there will be room for each of them potentially to revise most any aspect of his or her view, on the basis of what he or she takes to be good reasons. The idea of overlapping consensus suggests how, against the background of such holism, their joint effort at working towards moral agreement can proceed on the basis of any initial agreement whatsoever. Specifically, there is no need that the initial agreement pertain to what either takes to be foundational or basic.

This appeal to holism and overlapping consensus makes a schematic case that joint moral reasoning is possible even in the face of deep moral disagreement. It reflects at least a bare possibility. When moral disagreement is deep, we want to know more about how the parties can reasonably approach agreement. Each will need to be willing to compromise: to revise his or her view in a way that he or she would not have been willing to do, but for some modicum of concern or respect that he or she has for the other party. And this compromise must go deep, in that it must extend to what each counts as right or wrong, or as worth seeking or avoiding for its own sake. Such deep compromise among deeply disagreeing citizens of one democratic country seems to some both to be possible and to be subject to norms of rationality (e.g., Richardson 2002, chap. 11). Others are more pessimistic, seeing robust possibilities for morally reasoning with one another only on the basis of deeply shared “identities” or moral commitments (e.g., Sandel 1998).

Bibliography

- Anderson, E. S., 1991. “John Stuart Mill and experiments in living,” *Ethics*, 102: 4–26.
- Audi, R., 1989. *Practical reasoning*, London: Routledge.
- —. 2004. *The good in the right: A theory of good and intrinsic value*, Princeton: Princeton University Press.
- Beauchamp, T. L., 1979. “A reply to Rachels on active and passive euthanasia,” in *Medical responsibility*, ed. W. L. Robinson, Clifton, N.J.: Humana Press, 182–95.
- Brandt, R. B., 1979. *A theory of the good and the right*, Oxford: Oxford University Press.
- Bratman, M., 1999. *Faces of intention: Selected essays on intention and agency*, Cambridge, England: Cambridge University Press.
- Broome, J., 2009. “The unity of reasoning?” in *Spheres of reason*, ed. S. Robertson, Oxford: Oxford University Press.
- Campbell, R. and Kumar, V. “Moral reasoning on the ground,” *Ethics*, 122: 273–312.
- Chang, R. (ed.), 1998. *Incommensurability, incomparability, and practical reason*, Cambridge, Mass.: Harvard University Press.
- Clarke, S. G., and E. Simpson, 1989. *Anti-theory in ethics and moral conservatism*, Albany: SUNY Press.
- Dancy, J., 1993. *Moral reasons*, Oxford: Blackwell.
- —, 2004. *Ethics without principles*, Oxford: Oxford University Press.
- Dewey, J., 1967. *The middle works, 1899–1924*, Vol. 14, *Human nature and conduct*, ed. J. A. Boydston, Carbondale: Southern Illinois University Press.
- Donagan, A., 1977. *The theory of morality*, Chicago: University of Chicago Press.
- Dworkin, R., 1978. *Taking rights seriously*, Cambridge: Harvard University Press.
- Engstrom, S., 2009. *The form of practical knowledge: A study of the categorical imperative*, Cambridge, Mass.: Harvard University Press.
- Fletcher, J., 1997. *Situation ethics: The new morality*, Louisville: Westminster John Knox Press.
- Frankfurt, H. G., 1988. *The importance of what we care about: Philosophical essays*, Cambridge: Cambridge University Press.
- Gert, B., 1998. *Morality: Its nature and justification*, New York: Oxford University Press.
- Greene, J., 2004. “The secret joke of Kant’s soul,” in W. Sinnott-Armstrong (ed.), *Moral psychology* (Vol. 3: The neuroscience of morality: Emotion, disease, and development), Cambridge, Mass.: MIT Press.
- Habermas, J., 1984. *The theory of communicative action: Vol. I, Reason and the rationalization of society*, Boston: Beacon Press.
- Haidt, J., 2001. “The emotional dog and its rational tail: A social intuitionist approach to moral judgment,” *Psychological Review*, 108: 814–34.
- Hare, R. M., 1981. *Moral thinking: Its levels, method, and point*, Oxford: Oxford University Press.
- Harman, G., 1986. *Change in view: principles of reasoning*, Cambridge, Mass.: MIT Press.
- Hauser, M.D., 2005. *Moral minds: how nature designed our universal sense of right and wrong*, New York: Harper Collins.
- Held, V., 1995. *Justice and care: Essential readings in feminist ethics*, Boulder, Colo.: Westview Press.
- Horta, John F., 2012. *Reasons as defaults*, Oxford: Oxford University Press.

- Hume, David, 2000 [1739–40]. *A treatise of human nature*, ed. D. F. Norton and M. J. Norton, Oxford: Oxford University Press.
- Hurley, S. L., 1989. *Natural reasons: Personality and polity*, New York: Oxford University Press.
- Jonsen, A. R., and S. Toulmin, 1988. *The abuse of casuistry: A history of moral reasoning*, Berkeley: University of California Press.
- Kagan, S., 1988. “The additive fallacy,” *Ethics*, 90: 5–31.
- Knobe, J., 2006. “The concept of individual action: A case study in the uses of folk psychology,” *Philosophical Studies*, 130: 203–231.
- Koenigs, M., 2007. “Damage to the prefrontal cortex increases utilitarian moral judgments,” *Nature*, 446: 908–911.
- Kolodny, N., 2005. “Why be rational?” *Mind*, 114: 509–63.
- Korsgaard, C. M., 1996. *Creating the kingdom of ends*, Cambridge: Cambridge University Press.
- Lance, M. and Little, M., 2007. “Where the Laws Are,” in R. Shafer-Landau (ed.), *Oxford Studies in Metaethics* (Volume 2), Oxford: Oxford University Press.
- McDowell, John, 1998. *Mind, value, and reality*, Cambridge, Mass.: Harvard University Press.
- McKeever, S. and Ridge, M. 2006., *Principled Ethics: Generalism as a Regulative Idea*, Oxford: Oxford University Press.
- McNaughton, D., 1988. *Moral vision: An introduction to ethics*, Oxford: Blackwell.
- Millgram, E., 1997. *Practical induction*, Cambridge, Mass.: Harvard University Press.
- Mikhail, J., 2011. *Elements of moral cognition: Rawls's linguistic analogy and the cognitive science of moral and legal judgment*, Cambridge: Cambridge University Press.
- Nell, O., 1975. *Acting on principle: An essay on Kantian ethics*, New York: Columbia University Press.
- Nussbaum, M. C., 1990. *Love's knowledge: Essays on philosophy and literature*, New York: Oxford University Press.
- —, 2001. *Upheavals of thought: The intelligence of emotions*, Cambridge, England: Cambridge University Press.
- Rachels, J., 1975. “Active and passive euthanasia,” *New England Journal of Medicine*, 292: 78–80.
- Railton, Peter, 1984. “Alienation, consequentialism, and the demands of morality,” *Philosophy and Public Affairs*, 13: 134–71.
- Rawls, J., 1971. *A theory of justice*, Cambridge, Mass.: Harvard University Press.
- —, 1996. *Political liberalism*, New York: Columbia University Press.
- —, 1999. *A theory of justice*, revised edition, Cambridge, Mass.: Harvard University Press.
- —, 2000. *Lectures on the history of moral philosophy*, Cambridge, Mass.: Harvard University Press.
- Raz, J., 1990. *Practical reason and norms*, Princeton: Princeton University Press.
- Richardson, H. S., 1994. *Practical reasoning about final ends*, Cambridge: Cambridge University Press.
- —, 2000. “Specifying, balancing, and interpreting bioethical principles,” *Journal of Medicine and Philosophy*, 25: 285–307.
- —, 2002. *Democratic autonomy: Public reasoning about the ends of policy*, New York: Oxford University Press.
- —, 2004. “Thinking about conflicts of desires,” in *Practical conflicts: New philosophical essays*, eds. P. Baumann and M. Betzler, Cambridge: Cambridge University Press, 96–117.
- —, 2012. “Mapping out Improvements in Justice: Comparing vs. Aiming,” *Rutgers University Law Review*, 43(2): 211–242
- Ross, W. D., 1988. *The right and the good*, Indianapolis: Hackett.
- Sandel, M., 1998. *Liberalism and the limits of justice*, Cambridge: Cambridge University Press.
- Sartre, J. P., 1975. “Existentialism is a Humanism,” in *Existentialism from Dostoyevsky to Sartre*, ed. W. Kaufmann, New York: Meridian-New American, 345-69.
- Scheffler, Samuel, 1992. *Human morality*, New York: Oxford University Press.
- Schmitz, D., 1995. *Rational choice and moral agency*, Princeton: Princeton University Press.
- Schneewind, J.B., 1977. *Sidgwick's ethics and Victorian moral philosophy*, Oxford: Oxford University Press.
- Schroeder, M., 2011. “Holism, weight, and undercutting.” *Noûs*, 45: 328–44.
- Schwitzgebel, E. and Cushman, F., 2012. “Expertise in moral reasoning? Order effects on moral judgment in professional philosophers and non-philosophers,” *Mind and Language*, 27: 135–53.
- Sidgwick, H., 1981. *The methods of ethics*, reprinted, 7th edition, Indianapolis: Hackett.
- Singer, M. G., 1961. *Generalization in ethics*, New York: Knopf.
- Sinnott-Armstrong, W., 1988. *Moral dilemmas*, Oxford: Basil Blackwell.
- Smith, M., 1994. *The moral problem*, Oxford: Blackwell.

- Sneddon, A., 2007. "A social model of moral dumbfounding: Implications for studying moral reasoning and moral judgment," *Philosophical Psychology*, 20: 731–48.
- Sunstein, C. R., 1996. *Legal reasoning and political conflict*, New York: Oxford University Press.
- Tiberius, V., 2000. "Humean heroism: Value commitments and the source of normativity," *Pacific Philosophical Quarterly*, 81: 426–446.
- Vogler, C., 1998. "Sex and talk," *Critical Inquiry*, 24: 328–65.
- Wellman, H. and Miller, J., 2008. "Including deontic reasoning as fundamental to theory of mind," *Human Development*, 51: 105–35
- Williams, B., 1981. *Moral luck: Philosophical papers 1973–1980*, Cambridge: Cambridge University Press.
- Young, L. and Saxe, R., 2008. "The neural basis of belief encoding and integration in moral judgment," *NeuroImage*, 40: 1912–20.