A Conflict in Common Sense Moral Psychology

AARON Z. ZIMMERMAN

University of California, Santa Barbara

Abstract: Ordinary moral thinking about morality and rationality is inconsistent. To arrive at a view of morality that is as faithful to common thought as consistency will allow we must admit that it is not always irrational to knowingly act against the weight of reasons.

Common sense tells you that you shouldn't knowingly act immorally, and that when you shouldn't do something, there is a reason for you to refrain from doing it, a reason that outweighs those reasons in favor of its performance. But it is also common sense that when you know that the reasons against some course of action outweigh the reasons in favor of it, you would be irrational to go ahead with the action. Still, though common sense says that you shouldn't knowingly act immorally, it denies that immoral actions always arise from either ignorance or irrationality. Common sense is therefore inconsistent.

Consider an example: Surely you know that you shouldn't rob someone to increase your already quite adequate standard of living, even if you can get away with the crime undetected. And if you know that you shouldn't steal for personal gain, you must know that there is a reason not to execute the action that outweighs those self-interested reasons that speak in its favor. But it seems trivially true that when you know that the reasons for theft are outweighed by the reasons against it, you would be irrational to go

ahead with the crime. And this seems too strong: to be rational one needn't act in a morally acceptable way. Irrationality is one thing, immorality another.

To resolve this conflict in ordinary thought we must abandon at least one of the following four intuitive principles:

(1) One should never knowingly act immorally.

(2) If one knows that one should not act immorally, one must know that there is a reason not to act immorally—a reason that outweighs those reasons in favor of one's acting immorally.

(3) If one knows that there is a reason for one to refrain from acting immorally—a reason that outweighs those reasons in favor of one's so acting—then one would be irrational to act immorally.

(4) It is not always irrational for one to knowingly act immorally.

I will argue in what follows that (3) is the culprit. Common sense should persist in maintaining that the reasons against acting immorally always outweigh the reasons in its favor, but it must allow that knowingly acting against the weight of reasons is not necessarily irrational. Though it does not entirely conform to common thought, the stance I will describe is as faithful to ordinary thinking as consistency will allow. Or so I will argue.

SECTION 1

Before moving on to consider our four incompatible propositions, I want to first address a charge of equivocation to which many readers might feel tempted. Surely, one might argue, premise (1) is trivially true if it is supposed to say that *morally speaking* one should never knowingly act immorally. And premise (2) will prove equally obvious if given a similar reading: When one knows that morally speaking one should not perform some immoral act, one will know that there is some moral reason to refrain from the immoral action, a reason one knows to have more moral weight than any set of opposing reasons that might be offered. But if the terms in (1) and (2) are glossed in this way, (3) will contradict common sense. For we do not think that the mere fact that one knows that there are morally overriding reasons to refrain from an action—reasons that have more moral weight than the reasons in its favor—entails that the action would be irrational to perform. If, for instance, one knows that those reasons that might be given in favor of an immoral action are weightier from a prudential or self-interested point of view, common sense allows that one need not be irrational in acting immorally.

Though, as I have said, I agree that in the final accounting (3) must be rejected, I do not think that it can be dismissed in precisely this way. Moreover, though I intended to advance (1) and (2) as common sense truths, I did not intend to trivialize them in the manner suggested.

We can dispose of one common way of drawing the distinction necessary to trivialize premise (1): 'ought' is not, as many have claimed, ambiguous between the 'ought' of morality and the 'ought' of rationality. In saying this, I take the ambiguity thesis to be descriptive rather than prescriptive in form. The theorist in question is

arguing that as it is actually used, the English word 'ought' really has the two distinct senses he has distinguished. (He is not arguing, as a reformist might, that we should use 'ought' in these two different ways for the sake of conceptual clarity.) Now descriptive claims of ambiguity must be buttressed with linguistic data of two main varieties. First, we can turn to languages other than English to look for two distinct lexical items each one of which best captures exactly one of the two senses of the word proposed. (Here we assume that native speakers of the language in question do not regard the two terms as synonymous.) But when we investigate foreign tongues, the results count against an ambiguity thesis: for as far as I can determine, there is no language in which one word does the work of the supposed 'ought' of rationality, whereas a distinct word fills in for the 'ought' of morality.

There is, however, a second source of data for ambiguity: we can consult our own language and employ a test based on 'verb phrase ellipses'.[1] Consider, as an example, the genuinely ambiguous 'bank'. 'Lewis went to the bank and so did Phil' can be used to say that Lewis and Phil both went to the shores of a river, or that they both visited a financial institution; but it cannot be used to literally say that Lewis visited a financial institution whereas Phil spent time next to a river. Why? Because 'bank' really has two distinct senses, and one is forced to employ the meaning expressed by the explicit occurrence of 'bank' when interpreting the 'implicit' occurrence evoked by 'and so did Phil'. That is, one interprets 'and so did Phil' by looking back at 'Lewis went to the bank' and conceptualizing the destination to which Phil traveled in the very way one has conceptualized Lewis' destination. The mixed reading is ruled out by the fact that the

concept expressed by ‘bank’ when it is used to denote riverbanks is distinct from the concept ‘bank’ expresses when it is used to denote financial institutions.

Now I think that ‘ought’ is quite unlike ‘bank’ in this regard. There is really one concept expressed by ‘ought’ (and ‘should’) in both of the uses at issue even though the relation denoted ties together an extraordinarily diverse range of relata. To see this, suppose that John is Frank’s child and Bill their neighbor. Frank in entirely indifferent to John’s wellbeing, and, despite his moral obligation to care for his children, does nothing to help him. Bill, on the other hand, while not morally obligated to care for the children of others, has decided to help John in any way that he can. When Bill asks your advice on how best to execute his decision, you suggest he buy the books that John needs for school. Still, you can’t help reminding Bill that Frank is morally obligated to do what you have recommended. You therefore utter, ‘You should buy John his books, but that goes double for Frank’, to describe what Bill should do in order to accomplish his settled ends and what Frank must do in order to respect his moral duties. A single univocal occurrence of ‘should’ is used to do what the ambiguity theorist claims it is incapable of doing.

This is not to say that ‘should’ and ‘ought’ do not display either ambiguity or some less drastic form of context-sensitivity along other dimensions. For instance, the relation denoted by ‘should’ in (1) and (2) is arguably different from the one it denotes in ‘We should have enough grain to last the winter’, a sentence one most commonly uses to assert that we will have sufficient grain so long as nothing extraordinary happens. To show this we might suppose that Frank’s family has entrusted him with storing enough of the grain harvested in the fall to sustain them through the winter months. You have entrusted me to do the same, and in the absence of an unexpected calamity we will have

enough food to survive until spring. Imagine, too, though, that I know that Frank's stores have been depleted through a series of disastrous transactions in which he has exchanged his produce for pretty colored beads. Surely, I cannot hope to inform you that we will almost certainly survive until spring and that Frank has the relevant duty (a duty he has failed to fulfill) with a single utterance of, 'We should have enough grain to last the winter and the same goes for Frank'. If you knew Frank's situation, the intended reading would come across as a bad joke—something like 'Bob's at Credit Suisse—so we're at a bank and he is too', when it's lamely offered after too many beers at a riverside picnic.

Thus, while I admit that 'should' and 'ought' display a certain amount of context sensitivity, I disagree with the sort of contextualist position that Ralph Wedgwood endorses, according to which a thief's helper can truly say to her 'You ought to use the Phillips-head screwdriver to open the safe', as a morally indignant onlooker also truly says, 'You ought not use that kind of screwdriver or any other to open the safe. You ought not open this safe at all'.[2] Instead, I would argue, the moralist says something true here whereas the thief's helper says something false—though the false proposition at issue is of great use to the thief given her immoral ends. Note, however, that the falsity of the thief's helper's utterance, 'You ought to use the Phillips-head screwdriver', might be thought of as compatible with the truth of 'You ought to use the Phillips-head, if you are going to open the safe'. For if we join Wedgwood in treating 'ought' as a sentential operator, and the conditional is given its classical truth-functional reading, our 'if…then' sentence will be logically equivalent to 'Ought ((not (you open the safe)) or (you use the Philips-head))', a sentence that is made true by the fact that the thief shouldn't open the safe. Of course, the problems for classical readings of 'if…then' arise here in spades, as

our analysis would also give a true reading to, 'You ought to use a wet banana, if you are going to open the safe', and therein suggest, in a Gricean vein, that a clever moralist might use this absurd (but true) sentence to convey the information that the thief ought not open the safe. But there is no reason to think that the slate of (largely inadequate) non-classical semantics that have been offered for 'if…then' would face any special obstacles in normative contexts like this one.[3]

Still, there are cases in which we explicit modify 'should' and 'ought', as when we cautiously tell someone that, 'morally speaking' she oughtn't steal for financial gain. What are we to make of these uses? First, when someone says that, morally speaking, one ought not x, she is reporting what morality 'says' about one's x-ing. And when someone speaks metaphorically of morality 'saying' that one should not x, she is literally asserting that a prohibition against one's x-ing must be included in any complete set of true moral claims. (If one thinks of morality as theoretical in nature, one will use 'Morally speaking one should not x' to assert that the proposition that one should not x could be derived from a correct moral theory were that theory conjoined with the relevant non-moral facts.) On the other hand, when someone simply says that one should not x, she is saying something about x itself. That is, even if the grounds of her assertion are moral in nature, in saying nothing more than that one should not x she asserts something directly about x rather than morality's 'view' of the action. Moreover, because 'Morally speaking one should not x' can only be used to indirectly report on the status of one's x-ing, it is not typically used to resolve deliberation or to issue advice that is final or conclusive in nature. So, for instance, if I ask you whether I should x, and you tell me that 'morally speaking' I should not x, you haven't yet fulfilled my request. (When I ask my question,

I don't just want to hear what *morality* says about x-ing, I want to hear what *you* have to say.) And if I ask myself whether I should x, and I tell myself that morally speaking, I should not x, I haven't yet concluded my deliberation on the matter; I must still go on to decide whether I should do what I am morally speaking supposed to do. In contrast, (1) is to be understood as saying that if one knows that an action is immoral one *simply* should not perform it; one should not perform it *full stop*. This is not quite the same thing as saying that one should not perform the action 'all things considered', as in simply asserting or judging that someone should not do something, one needn't take oneself to have considered everything potentially relevant to her decision (nor need one imply or pragmatically implicate that one has done so). Nevertheless, the two sorts of judgment are not too far apart; both serve to render advice without qualification.

Note too that, as (1) is meant to be interpreted, to say that an action is immoral is to say more than that is morally suboptimal given that it is not entirely intuitive that one should always act in a morally optimal manner. That is, premise (1) should be read as stating that one simply should not engage in an action when one knows that one is morally obligated to refrain from it, and it is meant to allow the possibility of supererogatory actions that are morally good but not obligatory.[4]

Suppose, for instance, that you are deciding whether or not to use your medical training to help relieve the physical suffering rampant in an impoverished rural area or to instead remain in a culturally diverse, intellectually lively city in which you will have a much more enjoyable existence. You might think to yourself, 'Surely it would be morally better to give medical assistance to those who need it most, but it is not clear if that is what I should do. I really don't want to leave the city'. And it might be that a moral

person who was trying her best to advise you on the matter, who had no doubts as to which of the two actions you are considering would be morally better than the other, might still be uncertain as to what advice to give.

Of course, hesitation on this score might be wrong-headed. It may be that we should always do that which we know to be morally best for us to do, and that our reluctance to choose or advise the morally optimal path always reflects some deficiency in our thinking.[5] But the quite widespread hesitation among otherwise moral people to choose and advise what is known or justifiably believed to be the morally optimal course of action shows that the fact that we should always do what we recognize to be morally best, if indeed it is a fact, is not so central to our ordinary thinking as to warrant assigning it the status of 'intuition'. Thus, if (1) is interpreted as saying that we should simply do what we know to be morally best for us to do, the incompatibility evidenced by (1)-(4) will not run particularly deep. But it will run deep if it is restricted to those particularly loathsome actions and omissions that we commonly forbid without hesitation. Consider the example with which I began: You simply should not steal from someone to increase your already quite adequate fortune even if you can get away with it. It is not just that refraining from theft would constitute kindness on your part; nor is it just that 'morally speaking' you shouldn't go in for this kind of thing. We say without qualification that you should not steal, and we say this because we think that it is morally prohibited.

This is not to suggest that our moral obligations are always easy to respect. Suppose, for instance, that you would have to kill an unknown innocent person in order to survive. You might know that you are morally obligated not to commit the murder, but plan to do so anyway. Still, I take ordinary moral thought to hold that, however

predictable a misdeed may be, its moral impermissibility entails that we simply should not commit it. In cases of this sort, when a moral obligation is particular onerous, and failure to fulfill it particularly commonplace, givers of sound advice commonly mark the situation by saying something like, 'You shouldn't do it, but I admit that, if I were you, I'd probably do it myself'. Though hedges of this kind inevitably diminish the persuasive force of one's counsel, they do nothing to affect its truth.[6]

Now we sometimes fail to live up to even those moral obligations that are not particularly difficult to respect; and there are those among us who relish transgression and are happy to insist that we shouldn't or needn't do what we are morally obligated to do.[7] Moreover, for all I will argue here, such challenges to our ordinary moral thinking may ultimately prove successful. I am simply insisting here that moral skepticism—though it may turn out to be defensible, and may even emerge from an intricate argument that begins with elements of our ordinary thinking about morality—lacks the *direct* endorsement of common sense. We commonly tell someone that she just shouldn't perform an action that she knows she is morally obligated to forgo. Similarly, it seems that we commonly regard someone as deliberating correctly when she moves from an appreciation of the fact that an action would be morally impermissible to the conclusion that she simply should not perform that action. Thus, while it is trivially true that morally speaking one should not x when one knows that one is morally obligated not to x, and it is no part of intuitive moral thought that everyone simply should always act in a morally perfect fashion, it is indeed intuitive but not at all trivial that, speaking simply and without qualification, one should never perform an action that one knows one is morally obligated to forgo.

Still, it must be admitted that abandoning (1) is one way to resolve the inconsistency we have identified; one might say that, despite the intuitive plausibility of the proposition, it just isn't true. Though philosophers since Kant have used 'categoricity' in somewhat different ways, this reformist maneuver is, I think, best described as abandoning the categoricity or *universal authority* of morality—for it admits that some people are not 'bound by' their moral obligations.

How serious a departure from ordinary thought would result? Some have argued that if we abandon our belief in the purported categoricity of morality, we must then deny the very existence of moral obligations. For example, Richard Joyce (2001) argues that just as the existence of women with supernatural powers is a referential presupposition of 'witch' so too 'morally impermissible' only applies to pairs of people and prospective actions if people simply should not perform these actions because they have decisive 'non-institutional' reasons to avoid them.[8] If Joyce is right, the abandonment of the categoricity of morality would render proposition (1) somewhat bizarre. If there are no moral obligations, none can be known, and no one can knowingly violate them. Premise (1) would then be akin to "Thou shalt not kill dinosaurs" in its lacking any normal conditions of utterance. (It would not, however, wear its deviance on its face, like the analytically impotent 'Though shalt not kill extinct animals'.) If we decide that despite their oddity, some principles lacking normal conditions of utterance are nevertheless true, and we wed this to the Joycean claim that moral obligations would have to be categorical if there were to be any of them at all, rejecting the categoricity of morality would not in fact resolve the conflict arising from (1)-(4). Since it is logically impossible that (1)-(4) all be true, some other premise would have to be rejected as well. In any event, one might

hope to buy consistency in ordinary moral thought with a less drastic revision of its tenants.

SECTION 2

Consider, now, premise (2). The concept of a *reason* for doing one thing rather than another is essential to its interpretation. The sense or use of 'reason' I intend here is that of a consideration for or against someone's adopting an action or attitude.[9] Reasons for performing an action are commonly referred to as 'pros'; reasons against its performance are denoted 'cons'." (In its most common employment, 'costs and benefits' refers to a class—though a substantially restricted class—of these considerations.) Admittedly, the term has other uses. 'Reasons' can be used to pick out the brute causes of a phenomenon as when one sets out the reasons that apples fall from trees without intending to anthropomorphize, justify or recommend the behavior of fruit. And 'reason' may also be used to denote a person's state of mind when acting, as when one cites John's belief that he was drinking gin as the reason why he drank a glass of gasoline.[10] But the interrelation of these seemingly distinct uses is far from obvious. For example, some theorists have argued that when an agent believes that the reasons in favor of an action outweigh the reasons against it, her belief in this proposition either must prove sufficient to cause her to perform that action, or its citation must provide a 'rationalizing' explanation of why she acted as she did. Moreover, it has been argued, when an agent believes that the considerations in favor of an action outweigh the considerations against it and this belief causes or explains her performing that action, the belief in question can always be cited as 'the reason' why she acted. Still, though these theses seem to have some truth to them,

they are best regarded as substantive claims seeking to establish a link between two distinct uses or senses of 'reason'.[11] The categorical difference between, on the one hand, an agent's *state of mind* when performing an action, and, on the other, *facts* about the action that might be cited as considerations in favor of its performance, militates against a straightforward identification of the concepts expressed when we use 'reason' in the two apparently different ways in question.[12]

Thus, 'reason' as it appears in premise (2) is to be understood as denoting pros and cons. So interpreted, (2) entails that if one simply should not x in a case in which one knows that one is morally obligated not to x, there must be something telling against one's x-ing, where the 'con' in question is the kind of thing that might be cited by someone deciding whether or not to x or someone advising someone else whether or not to x.

Which considerations are relevant in this regard? It seems that common thought is so varied on this matter as to contain few easily discernable restrictions. Many religious people think that God will punish those who violate their moral obligations and reward those who do not, or that God ultimately bestows his love upon the virtuous and his hatred upon the vicious. Thus, the purported fact that God will punish you or deprive you of his love for x-ing is commonly regarded as a reason not to x when, it is believed, your x-ing is morally prohibited. Those attracted to secular consequentialist thinking often say that when one is morally obligated not to x, this is because x-ing would cause great pain, unhappiness or harm, or detract in some other way from the flourishing of plants, animals or people. Thus, when it is thought that we are morally obligated not to x, the purported fact that our x-ing would seriously detract from happiness or flourishing is commonly

regarded as a reason for us not to x. Virtue theorists claim that when one is morally obligated not to x, x-ing would be vicious in some determinate way, by, say, being mean, nasty, greedy, dishonest or inhumane. Thus, when x-ing is believed to be morally prohibited, the purported fact that one's x-ing would be vicious is commonly regarded as a reason for one not to x. Those who place particular importance on fairness will argue that when someone is morally obligated not to x it is always true that were she even minimally rational she would not want others to x were they given the opportunity to do so, or that she would 'will' (or legislate) a prohibition against x-ing if she were to regard the matter with the requisite neutrality. Thus, the purported fact that someone would not want others to x were they given the opportunity to x, and the purported fact that someone would vote for a prohibition on x-ing were she to consider the matter from a neutral standpoint are often treated as reasons for that person not to x. Certain religious voluntarists think that the fact that God prohibits x-ing is itself a reason not to x irrespective of any reasons God might have for prohibiting x-ing.[13] Indeed, some secular deontologists—not to mention parents who have grown tired of arguing with their children—go so far as to claim that certain actions just should not be done: that there are actions that shouldn't be performed for no other reason than that they are morally prohibited. Thus, it is not entirely uncommon to regard the mere fact that it would be immoral of one to x as a reason not to x when, one thinks, one is morally obligated to forgo x-ing. For the purposes of assessing the conflict in ordinary moral thought constituted by (1)-(4) above, even 'minimal' reasons of the kind offered by strict deontologists (and frustrated parents) must be considered.[14]

Now if we acknowledge that this admittedly broad understanding of reasons for action is at least a common one—a conception reflected in our practices of deliberating, seeking advice, and giving and asking for reasons to act or refrain from doing so—a rather simple source of support for our attribution of (2) to common sense makes itself available. For it would be utterly bizarre to tell someone that though she knows that she should not x, there is no reason at all for her not to x, unless, that is, one meant by this that the mere fact that she should not x is the only reason for her to refrain from x-ing that can be offered. The absurdity of asserting that someone should not x even though there is *absolutely* no reason for her not to x—when this is not intended as an expression of strict deontology but as something even more stark—provides ample warrant for regarding (2) as a feature of our ordinary thinking.

Of course, to say that all of the purported reasons described above must be 'considered' is not to say that we actually have these reasons to act, much less that we have them in cases in which they fail to move us. That is, because premise (2) is conditional in form it merely asserts a logical relation between cases in which we know that we should not x because we are morally obligated not to x and the existence in these cases of considerations of some sort or other that might be cited against our x-ing.

With this in mind, suppose, as dedicated instrumentalists claim, that one wouldn't have a reason to refrain from a morally prohibited action if one neither cared about one's moral obligations nor cared about anything that one could get by respecting those obligations. If we marry this form of instrumentalism to premise (2) we are forced to say that were I to find myself without moral concerns it just would not be true to say that I should forgo morally prohibited actions. Thus, if she accepted the existence of any actual

conflicts of duty and self-interest, the instrumentalist would be forced to abandon the common-sense proposition (1). Instrumentalism might therefore be seen to provide a resolution to the conflict in ordinary moral thought that we have identified—a resolution that would require a fairly substantial revision in the common way of thinking.[15]

There are, of course, versions of what is often called 'reasons internalism' that argue for a weakened version of the conceptual tie instrumentalists posit between an agent's concerns and the reasons she has for acting in one way rather than another. For instance, some philosophers who regard themselves as internalists allow that a person can have a reason to be moral even if she doesn't actually care about morality (and what it can get her) so long as she would develop moral concerns were she exposed to *cognitive therapy*. Here cognitive therapy is often understood as a process that induces changes in our preferences by forcing us to confront all the relevant (non-moral) information; imaginatively project ourselves into the perspectives of those affected by our actions; consider our ends in light of one another; and exercise thought and imagination in similar ways.[16] Still, I think we can assume—contra Smith (1994)—that morality is 'autonomous' in at least the following sense: unless cognitive therapy is described in morally loaded terms—by, for example, requiring those who have projected themselves into the perspectives of others to feel sympathy in the 'thick' (normatively loaded) sense of this emotion described by Nagel (1970)—we cannot rule out the possibility of someone's emerging from the sort of therapy described without having adopted largely moral concerns. If this is right, then—as David Brink (1986) has argued—standard forms of internalism join straightforwardly instrumentalist views in requiring us to revise our ordinary conception of moral obligation.[17] Pre-theoretically, we think that those who not

only lack concern for morality but would remain unconcerned even after cognitive therapy must have a reason to refrain from doing what they know to be morally impermissible (given that they simply should not act in morally impermissible ways). Thus, the apparently categorical nature of morality is not something that even the most sophisticated internalists can easily save.

Suppose, though, that an instrumentalist (or perhaps even a reasons internalist of a weaker sort) were to try to retain the commitments of ordinary moral thought enshrined in premise (1). To retain (2) she would then have to argue that we can in fact always be given instrumental (or 'internal') reasons for doing what we know we are morally obligated to do. It would be implausibly Panglossian to think that in our transient world, filled as it is with morally imperfect beings, a person's desires are never best served by her engaging in activities that she knows she is morally obligated to forgo. (And the same might go for those 'improved' preferences a person would have if we monkeyed with her in the ways the sophisticated internalist describes.) But if an instrumentalist were to acknowledge that conflicts of this kind occur, she could only retain common sense morality by accepting a substantially religious metaphysics: When I must now forgo achieving what I want to respect my moral obligations, I have a reason to be moral provided by supernatural incentives. I should respect morality by forgoing those desirable activities it prohibits because God will compensate me for my sacrifice in the world to come.

In fact, I take it that common thought is actually divided over whether our reasons to act as we are morally obligated to act depend for their existence on divine incentives in the way imagined. For instance, Dostoyevsky's character Ivan Karamazov concludes (if

not in exactly these terms) that if God does not exist, everything is morally lawful and nothing morally obligatory. Surely some (non-fictional) people accept Ivan's view precisely because they assume an instrumentalist view of reasons for action; and, of course, if enough people assume such a view it must be regarded as a feature of common sense. It is important to note, however, that there are other reasons why one might accept Ivan's conclusion. Suppose, for instance, that one is a voluntarist who thinks there are cases in which the fact that God says that one should not x is the only reason not to x, and that one has this reason not to x regardless of whether one wants to follow God's commands. Or suppose that one is a prudentialist who thinks that the fact that x-ing will diminish one's overall wellbeing is the only reason not to x; that God insures that one's wellbeing is always best served by one's respecting the moral prohibition on x-ing; and that one has this reason to preserve one's overall wellbeing even when one does not care about oneself or one's overall good. Both of these accounts posit non-instrumental reasons for action; and both insist that if God does not exist, there are cases in which one has no reason to do what morality requires. When we add into the equation those non-skeptical secular thinkers who deride as hopelessly optimistic the claim that our desires are always best served when we respect our moral obligations, the assumption of instrumentalism looks much less widespread.

Indeed, this somewhat uneasy relationship between instrumentalism and ordinary moral thought often leads to a predictable revision in the meta-ethical beliefs of people who lose their religious beliefs. I consider my own thinking on these matters to be not entirely unrepresentative. As a child I was taught to believe that God punishes the wicked and rewards the righteous, and I was prone to thinking of this as *the* reason to do what is

right in cases in which I might profit from doing otherwise. Then, at some point I decided, for somewhat complicated reasons, that even if there is a God, and there is some intelligence responsible for the existence of the universe, there is no reason to think that it plays a role in enforcing human morality. I then concluded, on that basis, that I had no instrumental reason to be moral when I could get away with disobedience, and that since I had no instrumental reason to act in accord with my moral obligations when I could profit from their transgression, I had no reason whatsoever to respect morality in such cases. (I was reading Dostoyevsky at the time.) I therefore decided that it was okay to lie and cheat when it served my purposes, and even okay to find quiet ways to hurt those people who impeded my aims. Nevertheless, after living in this way for a short time—by telling myself that cheating, lying, and harming others is 'permitted'—I gradually returned to a way of thinking more in accord with common sense morality. Feelings of guilt or shame didn't *ground* the change; which is to say that whatever causal role such feelings may have had in my conversion, I don't now regard myself as having instrumental reasons to be moral derived from these feelings. (I don't, for instance, now think that the fact that cheating would make me feel guilty is the reason—or even a very good reason—not to cheat.) Instead, where I before assigned an instrumental basis to moral reasons, I now simply treat them as 'foundational'. I now find myself concluding that when an act would be greedy, or would cause harm, or could only be sanctioned by unjust or inequitable rules, this is itself a sufficient reason not to perform it.

Foundationalism needn't yield dogmatism. For I will, if pressed, try to explain to others why one shouldn't be unfair, greedy or cause harm; and I can show how prohibitions against such things cohere or fit with other parts of my moral framework.

But I don't really feel the need for such explanations, and it would be a stretch to see the explanations I come up with when pressed (or the mere fact that they largely cohere with the rest of my evaluative beliefs) as capturing the justification or evidential basis upon which I actually hold these extraordinarily basic moral beliefs. Indeed, as far as I am concerned, the moral concepts I use when defending my basic beliefs are less fundamental than those of selfishness, harm, and injustice; and I am fairly sure that I could not satisfactorily defend my basic moral beliefs were I to appeal only to non-moral premises. Still, despite this acknowledged inability, I regard myself as rational in holding to my fundamental precepts. I don't think that their argumentative defense is necessary.[18]

Of course, I don't always do what is moral; but I would be surprised to discover that my actions are now less moral than they were when I was convinced that God (or my parents) would punish my indiscretions. When tempted to act in a way I believe to be morally prohibited I am sometimes (but not invariably) held in check by my belief that the act would be selfish, destructive or unjust, just as before my conversion I was sometimes (but not invariably) held in check by my belief that God (or my parents) would make me suffer for my misdeeds because those misdeeds are selfish, destructive and unjust.[19]

Now I do not pretend that everyone's moral thinking resembles my own in this regard. But there are some suggestive data. For instance, in a fairly recent Pew Research Council poll, 50 percent of respondents said that one need not believe in God to be a moral person, 47 percent thought belief in God necessary for morality and 3 percent were undecided.[20] Of course, it may be that respondents classify as "moral" even those instrumentalist atheists and agnostics who believe people should only respect their moral

obligations when this is necessary to accomplish their ends; and this may be due to the assumption that atheists and agnostics never settle on largely immoral ends. (Though I regard this hypothesis as highly unlikely.) At any rate, more precise measures of common meta-ethical thinking are required if we are to both explain these results and acquire a better sense of the variation of meta-ethical thought across times and places. Nevertheless, prior to further investigation, it is not, I think, wildly implausible to regard the instrumentalist account of reasons for action as very gradually, agent by agent, losing whatever grip it once held on ordinary moral thought. While common sense currently has it that we should always do what we are morally required to do, and that we always have reason to do what we ought, it seems decidedly undecided on the nature of these reasons, and, in particular, their relation to such things as self-interest, prudence, and the will of God.[21]

SECTION 3

An adequate understanding of premise (2) depends on more than just the concept of reasons; we also need some grip on what it is for one reason to 'outweigh' another. Does our common sense notion of the weight of reasons dictate that knowing what one simply should not do is always accompanied by reasons weighty enough to 'tip the scales' against that course of action?

Suppose S knows that she simply should not x. We have already concluded that (so far as common thought is concerned) she must then know that there is some reason for her not to x—a consideration that might be cited against her x-ing. But we might also suppose that she is also aware of a reason in favor of her x-ing. Are we to suppose that it

is consistent with her knowing that she simply should not x that the 'pros' in question outweigh the 'cons'? Or, if this is indeed impossible, are we to suppose that S's knowing what she simply should not do is consistent with her ignorance of the fact that the cons outweigh the pros? Even if we cannot provide anything like an analysis of what is commonly meant by the 'weight of reasons', I think we can rule out positive answers to these questions.

I begin by noting that in paradigm cases, S will judge that she ought not x precisely *because* she has compared the reasons for and against x-ing and decided that the cons outweigh the pros. If things go well and she thereby comes to know that she simply should not x, this knowledge will not only be accompanied by, but will in fact be grounded in*,* her knowledge of the relevant reasons and their respective weights. (For instance, her judgment that she should not x will be counterfactually—and perhaps causally—dependent on her judgment that she has better reasons not to x than x, and when questioned, she will typically appeal to this fact about the weight of reasons to defend her judgment.) Indeed, on a 'bottom up' picture according to which reasons for action are prior to facts about we ought and ought not do, the fact that S ought not x will be *made true* by the fact that the cons outweigh the pros, and S will therefore come to be justified in believing the former fact by coming to know the very fact that makes it true. For instance, we might say that the fact that S ought not steal is made true by: (a) the fact that the theft will yield undeserved harm to another person, (b) the fact that the theft will bring (admittedly undeserved) financial benefit to S, and (c) the fact that (a) is a 'stronger' or 'weightier' reason against S's performing the act than (b) is in its favor. In the paradigm case, S will come to know that she ought not steal by grounding her belief

in this proposition in her belief in (c). In consequence, (2) will hold in such cases. S won't know that she should not perform a given immoral action unless she has determined that the reasons against the action outweigh the reasons in its favor.

Of course, one might ask how S comes to know the normative fact (c). Mustn't she infer it 'top-down' from the quite general fact that one ought not steal except in the direst of circumstances? Perhaps not. In the simplest case, where S is deciding 'in a vacuum' between more and less of some morally neutral good that she values, the weight of reasons will simply be determined by the balance of value. For example, suppose S must choose between remaining in the fire and leaping back into the frying pan. If S jumps, she will indeed experience pain and suffer harm, but less pain and harm than she will endure if she remains in the fire. Here S should jump, and that fact is made true by: (a') the fact that jumping will yield n degrees of pain and harm for S, (b') the fact that not jumping will yield n+1 degrees of pain and harm for S, and (c') the fact that (b) is a weightier reason for S to jump than (a) is for S not to jump. In this highly idealized case, S's knowledge that she should jump will be grounded in her knowledge of (a')-(c'), and her knowledge of the normative proposition (c') will be grounded in her knowledge of the evaluative proposition that more pain and harm is worse than less pain and harm, where this last bit of knowledge is arguably *non-inferential*: i.e. introspective, perceptual, or purely conceptual in nature. In the simplest cases, then, knowledge of the right may be grounded in non-inferential knowledge of the good.[22] In any event, so long as facts like (c) and (c') are commonly assumed to be known in some way or other, the case can be made for a limited version of premise (2): At least in paradigm cases, it is commonly held that one's knowledge that one simply should not x is grounded in (and so accompanied

by) one's knowledge that the weight of reasons tells against one's x-ing. And this suggests that in non-paradigmatic cases when a subject knows that she should not x but is not acquainted with the reasons that make this so, she will still know (*de dicto* as it were) that reasons of the relevant sort must exist. Thus, in every case in which we credit a subject with knowing that she should not x we must allow that she is in a position to conclude that there are on-balance reasons against x-ing. Since it is fairly difficult to see what could prevent someone from drawing the requisite inference, (2) should hold in full generality.

Of course, unless we assume a naïve form of utilitarianism, decision-making contexts won't ever be as simple as they are in our toy 'frying-pan or fire' scenario. We may even want to allow situations in which the fact that x-ing would be F is a weightier reason for S not to x than the fact that x-ing would be G is a reason for S to x, even though, generally speaking, being G bestows more value on something than is diminished by its being F. Consider one of our previous examples. You might think: the fact (a'') that you will be bored and unhappy is a reason against your setting up your medical practice in an impoverished rural area; that, on the other hand, (b'') alleviating greater suffering than you would otherwise alleviate is a reason in favor of the move; and that in the final analysis fact (a") is a weightier reason against your making the move than fact (b") is in its favor. Nevertheless, you might judge that an act's alleviating a great deal of suffering adds more value to that act than is diminished by its detracting from your own happiness to a lesser extent, and that it would therefore be *better* if you were to move from the city than stay put.[23] Still, in cases of this kind, where our simple evaluative judgments come apart from our judgments about the weight of reasons, they also come

apart from our simple normative judgments. You don't judge that, because moving is better than staying put, you simply should move, while at the same time judging that the reasons against moving are weightier than the reasons in its favor. The incoherence of this stance bears witness to the truth of (2).

Should we say that our judgments about the weight of reasons come entirely apart from our evaluative conclusions in cases of this kind? Perhaps not. For you would be conceptualizing your situation in an intuitively compelling way were you to judge that though it would be *better* if you were to move to the country, your staying in the city would be *better for you* than your going would be. If you then went on to conclude that you should stay in the city, and your judgment to that effect were grounded in your belief that the reasons for staying are weightier than the reasons for leaving, this judgment about the relative weight of reasons would reflect not the *objective* value of these alternatives, but their putative *subjective* value—their value for you. And if common sense allows that one can truly or correctly judge that subjective value has come apart from objective value in this way, it commits itself to the substantive independence of prudence and morality.

Note, though, that those who think morality and prudence never conflict needn't reject the very existence of facts of subjective value, as there are plenty of non-moral scenarios in which these facts play a central role. You want to watch the ballet, I want to watch the ballgame, and we both prefer watching one of these things together to spending our time apart. On any given occasion in which we are faced with such a choice one of us will give way to the other without resentment, for we know that the contrary course will be chosen sometime in the future, and that the distribution will be just. (Mutual respect is

therefore never in question.) Still, there is no doubt, when we watch the ballet, that what we are doing is better for you than the other activity we contemplated, an activity that would have been better for me but not for you. Thus, when the non-moral value of an action is at issue there is no pressure to think that what is best for the agent must be equated with what is best for the agent to do, which must in turn be equated with what is simply best, which must then be regarded as what the agent has most reason to do. Instead, in many such cases common sense would say that none of the relevant alternatives is *simply* better than any other; that one of the alternatives is still better *for the agent* than any other; that there are equally compelling reasons for choosing that action and another incompatible alternative; and that the agent is rationally permitted to choose either course. Indeed, I find myself in a situation of precisely this sort when with remote control in hand I decide, with your blessing, whether tonight we will watch the game or, instead, the ballet.

SECTION 4

I think the case I've just described allows us to see how one might try to distinguish a class of rational but yet knowledgeable transgressions of the moral law. Suppose a person can know that what she is doing is better for her than inaction would be, but that the action is (simply) worse than the alternatives. Mightn't such an action be so bad that it is morally impermissible? And mightn't the agent tie her judgments about the reasons for and against the action to its objective value, while choosing to act in accord with her judgments about the action's value for her? Moreover, mightn't an action of this sort be correctly described as *prudent* and hence—in one important sense or use of the term—a

rational action to perform? If there are actions of this kind, and common sense can be made to acknowledge their existence, the conflict we've been discussing can be resolved with the rejection of thesis (3): A person might know that there is a reason for her to refrain from acting immorally—a reason that outweighs those reasons in favor of her acting—without displaying irrationality in pursuing that immoral course of action.

Let us then suppose that after waiting for some time to buy a pack of chewing gum as the shopkeeper dawdles in the back room, I grow impatient, and leave the store with the gum in hand, without paying. I'm telling you that I know that people should not steal. I know that I am a person, and that (therefore) I should not steal. I know that there are a slew of admissible exceptions to this rule, but that none of them apply in this case. I know that I should not steal this gum, here and now; that I should wait patiently for the shopkeeper's return or leave the store empty-handed. Moreover, my knowledge is not purely 'top down' as I know some of the reasons why I should not steal the item: reasons that go beyond the mere fact that it would be wrong or that I ought not do it. My breaking the law is unfair. I wouldn't want it done to me were I the dawdling shopkeeper. Perhaps I've in some sense tacitly endorsed rules detailing the just exchange of property by benefiting from large-scale compliance with those rules. Moreover, though things needn't work out this way, to simplify the case we might suppose that I know that the harm to the shopkeeper in losing the gum without compensation is greater (on any reasonable measure) than the benefit to myself of having the gum for free. I therefore conclude, *on all these grounds*, that the reasons to leave the gum or remain patiently in line outweigh the reasons to take it without paying. Nevertheless, I take the gum because I correctly judge this action to be best *for me* in the scenario on hand. I have acted selfishly,

impatiently, and unjustly. I have failed to show the shopkeeper the respect he deserves. Nevertheless, it seems, because I perform the action that is best for me, common thought insists that there is an important sense in which I act *rationally* in performing it.

Is this description of the case the best that we can come up with? I can imagine at least two objections. First, a strong internalist might challenge the very possibility of my judging that the reasons against stealing are weightier than the reasons in its favor while at the same time deliberately choosing to steal. Isn't the fact that I steal sufficient evidence that I have not made this determination? Second, a moderate internalist might admit that it is possible for me to act against the known weight of reasons in the way imagined, but nevertheless insist that I would act irrationally in doing so. We could only conclude that I had appreciated the weight of reasons if I regretted my action, experienced some loss of control while acting, or exhibited some other symptoms from which we might diagnose a breakdown in my 'reason'. And this would provide us with evidence sufficient to attribute the theft to akrasia.

Though I have considerable sympathy for these intuitions, in the end I think that following through on them would lead us to a greater revision in our common sense moral psychology than is necessary. To consider the first point: Someone who can actually give us the reasons why he shouldn't be doing what he is doing; who knows that he cannot defend his action from criticism because that action is really much worse than the alternatives; and therefore regards his behavior as entirely indefensible—such a person is not typically accused of ignorance. Admittedly, the greatest philosophers of antiquity may be right in arguing that, in the scenario I've described, I don't 'really' know that what I'm doing is wrong and that I produce the arguments against stealing as

an actor might, without genuine conviction. It may even turn out that the very existence of my temptation to steal is incompatible with 'full' moral knowledge, which can only be bought through virtue induced through habituation or acquaintance with the form of justice. But, at least as things currently stand, this is not the view of common sense. Indeed, the behavior I've described will be judged in a particularly harsh manner precisely because, as a mature adult who has thought a great deal about what makes stealing wrong, I know as well as anyone can that the cons outweigh the pros, and that my action is therefore indefensible.

The second (more moderate) internalist response is much more difficult to resist. Indeed, as I've said, I think propositions (1)-(4) present us with a genuine conflict in common sense moral psychology precisely because each seems independently quite plausible, and proposition (3) is no different in this regard. It captures the enormous intuitive pull to the claim that I cannot be acting *rationally* in this case because I know that I have more reason to refrain from the theft than to go ahead with it. Still, as I've said, I think that the alternative amendments in ordinary thinking that would have to be made to rescue consistency would be even more disruptive than abandoning this intuition would be. As I have already explained my reasons for thinking that neither (1) nor (2) should be rejected, I will here provide my reasons for thinking that, in the absence of extraordinary arguments to the contrary, (4) should be retained.

First, are we really to suppose that the irrationality of knowingly acting in an immoral way is to be deduced from common sense in so direct a fashion? We have said nothing here about freedom or autonomy, nothing about overestimating the differences between people, and nothing about the unique kind of practical inconsistency that arises

from enacting a piece of moral legislation while at the same time planning to achieve one's ends through its violation. (The arguments of Plato, Hobbes, Aristotle, Kant, and Nagel don't even play a walk-on part in (1)-(3).) Thus, we cannot conclude, as Scanlon does, that 'irrationality in the clearest sense occurs…when a person fails to form and act on an intention to do something even though he or she judges there to be overwhelmingly good reason to do it', and that, 'These are clear cases of irrationality because the…action they involve is in an obvious sense "contrary (to the person's own) reason"' (1998, p. 25).[24] Even if all immorality can be assigned to irrationality or ignorance, the path to that conclusion should not be as easy as this. The rationalist needs to go on and argue that knowingly acting against the weight of reasons is enough like knowingly acting imprudently; or failing to adopt known means to ends one is unwilling to revise; or holding contradictory or incoherent beliefs; to really warrant the epithet 'irrational'.

Second, our best empirical evidence suggests that (4) really is a rather deep feature of ordinary moral thinking. For example, when Shaun Nichols described a mathematician who has 'no emotional reaction to hurting other people' but never hurts people because he believes 'it is irrational to hurt others and that any rational person would be like him and not hurt other people', a majority of subjects claimed that the mathematician didn't 'understand' that hurting others is morally wrong. But a majority of the same people said that a psychopath who felt no moral concern but didn't have the mathematician's theoretical beliefs would understand the immorality of his actions.[25] This pair of responses seems to indicate that most people think that believing in the irrationality of immorality is actually *incompatible* with truly understanding moral

claims. Might someone who thinks that having this meta-ethical belief is inconsistent with moral understanding nevertheless think of the belief as true? It is hard to see how.

Again, I am not arguing that common resistance to the irrationality of selfish theft cannot be overcome. Nor am I cutting off the attempt to sustain historically important arguments against (4). I am only pointing out the ambitious nature of the enterprise.

SECTION 5

If we reject (3) and retain (4), we allow that a person needn't be irrational in performing an immoral action even when she knows that she has weightier reasons to avoid it than commit it. I have suggested that cases of this sort can arise when a person knows that the action she is considering is immoral because it is significantly worse than its non-performance, so long as she also knows that performing the action would be better for her. That is, I have claimed that our conception of the *reasons* for and against a prospective action (and the respective weights of these reasons) can correctly track our judgments about its objective value, while our judgments about the *rationality* of the action instead track our judgments about its value for the agent. S judges that because x-ing is much worse than not x-ing, she has more or better reasons not to x than x. She would therefore never dream of defending her decision to x in a public arena in which arguments are supposed to track the weight of reasons, which are in turn supposed to (at least loosely) match the balance of objective value. But S also judges that x-ing is better for her, and that it is therefore rationally permissible for her to x. She considers her action immoral, but prudent.

Cases in which we know that what is simply best isn't the best thing for us provide us with constant opportunities for moral struggle. Indeed, if we didn't make judgments of this kind there would be no deliberative pull away from the supererogatory. A thankfully less common but still troubling range of cases arises when we know or believe that what is simply bad or impermissible is nevertheless the best thing for us to do. What I have suggested here is that common sense is best modified to view situations in which prudence and morality seemingly conflict, and we seem to knowingly side with prudence, as cases in which we rationally act against the known weight of reasons. If this is right, the conceptual gulf between reasons and rationality cannot be avoided without an even more substantial alteration of ordinary thought than is otherwise required.[26]

---

[1] See Jay Atlas, *Philosophy Without Ambiguity: A Logico-Linguistic Essay* (Oxford, 1989).

[2] See Wedgwood's, 'The Meaning of "Ought"', *Oxford Studies in Metaethics*, Vol. I, ed. R. Shafer-Landau (Oxford, 2006), pp. 127-160. Cf. James Drier, 'Internalism and Speaker Relativism', *Ethics*, 101 (1990), pp. 6-26.

[3] For a recent discussion of the options see Jonathan Bennett, *A Philosophical Guide to Conditionals* (Oxford, 2003).

[4] For discussion see J. Urmson, 'Saints and Heroes', *Essays in Moral Philosophy*, ed. A. Melden (ed.) (Washington, 1958), pp. 198-216; and, more recently, Michael Zimmerman, *The Concept of Moral Obligation* (Cambridge, 1996).

[5] Such is the effect normally attributed to the very strong principle advocated in Peter Singer's, 'Famine Affluence and Morality', *Philosophy and Public Affairs*, 1 (Spring 1972), pp. 228-43; 'If it is in our power to prevent something bad from happening, without thereby sacrificing anything of comparable moral importance, we ought, morally to do it' (p. 231). Of course, Singer's principle will only have this consequence if we assume that we simply ought to do what (in Singer's words) 'we ought, morally to do'. For related arguments some of which are based on weaker (more easily defensible) principles see Peter Unger, *Living High and Letting Die: Our Illusion of Innocence* (Oxford, 1996).

[6] I thank Walter Sinnott-Armstrong for pressing me here.

[7] There is something of this in D. A. F. De Sade, *Justine, Philosophy in the Bedroom, and Other Writings*, trans. R. Seaver and A. Wainhouse (New York, 1965).

[8] *The Myth of Morality* (Cambridge, 2001).

[9] See Derek Parfit, 'Rationality and Reasons', *Exploring Practical Philosophy: from Action to Values*, eds. D. Egonsson et al (Aldershot, 2001), pp. 17-39; Thomas Scanlon,

What We Owe to Each Other (Harvard, 1998), and John Skorupski, 'Reasons and Reason', *Ethics and Practical Reason*, eds. G. Cullity and B. Gaut (Oxford, 1997), pp. 345-68.

[10] The example comes from Bernard Williams, *Moral Luck* (Cambridge, 1981). For two quite different attempts to reduce certain uses of 'reason' to others see Jonathan Dancy, *Practical Reality* (Oxford, 2000) and Donald Davidson, *Essays on Actions and Events* (Oxford, 1980). The first philosopher I know of to draw the relevant distinction is Francis Hutcheson, who separates 'justifying' and 'exciting' reasons in his *Illustrations on the Moral Sense*, ed. Bernard Peach (Cambridge, 1971).

[11] For endorsement of these claims see, among many others, Thomas Nagel, *The Possibility of Altruism* (Princeton, 1970) and Scanlon (1998). Humean theorists of psychological explanation deny the first claim by arguing that beliefs cannot cause or explain action without the aid of desires. Humean theorists of motivating reasons deny the second claim by arguing that every such reason is a composite of beliefs and desires. (As far as I can tell, these two 'Humean' views are always embraced or rejected as a pair.)

[12] For arguments to this effect, see Michael Smith, *The Moral Problem* (Cambridge, 1994), p. 96. I leave confirming this hypothesis with verb-phrase ellipses as an exercise for the reader.

[13] This view of reasons to act as one should may be losing ground even in traditional religious communities. For example, Larry Nucci found that Amish youngsters reported that if God had not set aside Sunday for rest it would not be immoral to work on that day whereas 80 percent said that if God hadn't prohibited hitting it would still be immoral to hit people; 'Children's Conceptions of Morality, Social Conventions and Religious Prescription', *Moral Dilemmas: Philosophical and Psychological Reconsiderations of the Development of Moral Reasoning*, ed. C. Harding (Chicago, 1986). See Shaun Nichols, *Sentimental Rules* (Oxford, 2004) for discussion of this and related studies.

[14] We might also allow a robustly non-naturalist view that has recently come under fire: that the mere fact that x-ing would be bad (where this is regarded as distinct in nature from x-ing's being bad in some particular way) is a reason not to x. For criticism see Scanlon (1998, chapter 2).

[15] The revisionist view is advocated by Phillipa Foot in 'Morality as a System of Hypothetical Imperatives', *Philosophical Review*, 81 (July 1972), pp. 305-16, though she has since abandoned it.

[16] See, e.g. Richard Brandt, 'Moral Valuation', *Ethics*, 56 (1946), pp. 106-21; and *A Theory of the Good and the Right*, (Oxford, 1979); and Bernard Williams 'Internal and External Reasons', in his *Moral Luck* (1981).

[17] David Brink, 'Externalist Moral Realism', *The Southern Journal of Philosophy*, 24, Supplement (1986), pp. 23-41.

[18]Though I cannot defend the claim here I think that moral sense theories (like those defended by Hutcheson and Hume) offer what is perhaps the best hope for explaining why this stance toward basic moral reasons is a rational one.

[19] What's the psychological difference between the two cases? Can we say that I fail to exercise rationality in those cases in which I don't do what I know I should? Insofar as

common sense holds to (4) our intuitions should suggest a negative answer. I will discuss some of the relevant issues in what follows.

[20] Pew Research Center For the People and the Press and Pew Forum on Religion and Public Life, *2002 Religion and Public Life Survey* (February 25-March 10 2002); http://people-press.org.

[21] I want to reiterate here that I don't mean to suggest that instrumental considerations were once universally thought to ground the authority of morality through appeals to God's executive power, and that this line of thinking has dissipated (in some uniform way) with the growth of secular ethics. It seems that there never has been any ideological consensus on the link between God, instrumentalism, and the reality of our moral obligations. See J. B. Schneewind, *The Invention of Autonomy: A History of Modern Moral Philosophy* (Cambridge, 1998) for support.

[22] I take it that J. S. Mill endorses roughly this picture as a general epistemology of normative and evaluative facts. Our knowledge that we are obligated to x is grounded in our knowledge that an emotional or physical penalty for x-ing would be better to instill than not; *Utilitarianism*, ed. R. Crisp (Oxford, 1998) at 5.14, p. 93. Our knowledge of this last fact is grounded in: (i) observational knowledge of the overall positive effects on happiness that the penalty yields, and (ii) knowledge of the analytic, axiomatic fact that equal amounts of happiness are equally good, so that more happiness is better than less (5.36.25.fn, p. 105).

[23] In saying that someone might make these judgments I am not saying that both might be true. The relation between the good and the right is a substantive issue not to be decided here.

[24] See too Scanlon, 'There certainly are cases in which our failure to give weight to considerations of well-being is irrational. These are cases in which…we judge that that these considerations are good reasons for acting a certain way but then fail to act accordingly' (1998, p. 129).

[25] S. Nicholas, 'Is it Irrational to be Immoral?' *The Monist*, 85 (2002), pp. 285-304.

[26] Much of this paper was written while visiting at the Research School of Social Sciences at the Australian National University during a sabbatical generously granted by the University of California, Santa Barbara. Thanks go to scholars at both ANU and UCSB for interesting conversations. I would also like to thank Jonathan Way for helpful written comments.