

# What Would a Deontic Logic of Internal Reasons Look Like?

Rufus Duits

**Abstract:** The so-called ‘central problem’ of internalism has been formulated like this: one cannot concurrently maintain the following three philosophical positions without inconsistency: internalism about practical reason, moral rationalism, and moral absolutism. Since internalism about practical reason is the most controversial of these, the suggestion is that it is the one that is best abandoned. In this paper, I point towards a response to this problem by sketching a deontic logic of internal reasons that deflates moral normativity to the normativity of instrumental rationality, and provides support for the assertion that one can hold fast simultaneously to internalism and at least many of the intuitive commitments of liberal moral thinking. Crucial to the proposal is an account of the enkratic principle – I ought to attempt to realise what I ultimately desire – as the source of obligations we owe to ourselves. I attempt to show how from this, in conjunction with some plausible assumptions, obligations to others might be derived.

**Keywords:** deontic logic, enkratic principle, instrumentalism, internalism, universal prescriptivism, weakness of will.

## 1. Internalism and Instrumentalism

One view about practical reason that is common both in the academy and outside of it is a sort of combination of internalism and instrumentalism. Roughly, internalism about practical reason is the thesis that a necessary condition of something being a reason for action is that it is connected in some straightforward way to the agent’s motivation, and instrumentalism about practical reason is the view that it is not possible to practically reason about ends since practical reasoning is reasoning about means to ends. Internalism is a theory about what counts as a reason, and instrumentalism is a theory about that which makes something a reason.

Now, as has long been recognised, reasons internalism, if right, poses problems to any moral theory that espouses a conjunction of any forms of the theses of moral absolutism and moral rationalism (the view that holds that the wrongness of an action implies that there is a reason not to do it). Call the family of such theories “the morality system” (Williams 1985, Chapter 10), and consider the ‘amoralist’s challenge:’ on the shore of a lake, two moral philosophers who can’t swim are trying to convince an onlooker to save a drowning child. One says, “the child has a right to life! You have a corresponding duty to jump in and save her!” The other says, “think of the distress her death will cause to her parents!

You're a good swimmer; it's a hot day; what have you got to lose by jumping in?!" Suppose the onlooker replies: "I know the child has a right to life; I know her death will cause distress to her parents; I know it would be easy for me; but I just don't feel like it just now. Sorry."

Suppose that internalism is right. Then it appears that the exhortative assertions that the philosophers utter do not express practical reasons for our amoralist, since they are not connected in any way to her motivational state. On the internalist account of this situation, the most rational thing our onlooker could do (given that she knows that she has alternatives that she finds more attractive), *ceteris paribus*, involves not jumping in.

Put generally, if internalism is true, agents may sometimes be in positions in which the most rational thing for them to do is at odds with any of the demands issuing from the 'morality system.' From the perspectives of the theories within, this conclusion is absurd.

Problems are compounded when instrumentalism is introduced. Instrumentalism is the view that agents can only reason about means, not ends. But, arguably, traditional moral theories are precisely attempts to reason about the ultimate ends of moral action. If instrumentalism is true, our philosophers cannot get the amoralist to recognise or adopt their 'reasons' as her fundamental or ultimate reasons by any rational process. She could only come to recognise these as ultimate reasons by an affective process.

Now these problems have been considered to be among the best reasons for rejecting internalism and instrumentalism about practical reason. Since we are invested so fundamentally and in so many ways in a capacity to reason about fundamental moral norms and to take those norms as reasons for action, we take ourselves to have a *prima facie* case against the internalist and the instrumentalist. But, in this paper, I want to outline the structure of one of the approaches by which views that maintain a combination of internalism and instrumentalism about practical reason could go about responding to the amoralist's challenge – this so-called "central problem" of internalism (Finlay and Schroeder 2015). My hypothesis is that a deontic logic of internal reasons can provide a sufficiently fecund account of the normativity that philosophers attribute to the moral domain – its variety, extent, and bindingness – , such that these fears about the implications of internalism and instrumentalism can be pushed back. Whilst this account of logical form omits much – most, perhaps – requisite detail, I hope that I can at least render plausible some steps in this direction, and if these steps are taken with big philosophical boots on, this nevertheless enables me to reach my destination more efficiently.

My conclusions, formal though they are, look a bit like a strengthened and expanded version of universal prescriptivism, and might have practical implications not dissimilar to contractualism. I draw heavily on the work that R.

M. Hare pioneered on the logic of imperatives,<sup>1</sup> and, in undertaking this programme, I also aim to respond to some of the key objections against universal prescriptivism that Hare himself recognised; in particular, I want to indicate how the universal prescriptivist can and should engage in normative debate with Hare's *amoralist*, his *akratic*, and his *fanatic*.<sup>2</sup> The logic is presented in the form of a series of propositions with accompanying commentary. The key innovation of the account is the claim that at the heart of a deontic logic of internal reasons stand *duties to ourselves*.

## 2. Desire

The logic requires a bit of a walk in. Note that I will argue here neither for internalism nor for instrumentalism.<sup>34</sup>

---

<sup>1</sup> In particular, in his 1952, 1963, 1971 and 1981.

<sup>2</sup> Hare believed that metaethical universal prescriptivism logically implied normative utilitarianism: "It is in the endeavour to find lines of conduct which we can prescribe universally in a given situation that we find ourselves bound to give equal weight to the desires of all parties; and this, in turn, leads to such views as that we should seek to maximise satisfactions." (1963, 123) Anticipating my discussion below, I should indicate why I do not think that utilitarianism results from the deontic logic offered here, nor indeed from Hare's own universal prescriptivism. Plausibly, it is a necessary condition for a moral theory to be classed as utilitarian that it derives moral judgements from the *aggregation of outcomes*. Whilst my account focuses evaluation on the way in which actions respond to the desires of affected parties, it is not in the least aggregative. Where there is a difference between the course of action that aggregatively leads to the best desire-satisfaction overall, and the course of action that leads to the satisfaction of the strongest – i.e., most ultimate – desire, the utilitarian would advocate the former course of action whilst I would advocate the latter. There is no logical implication of aggregation contained within the propositions advocated here. Similarly, despite Hare's profession, aggregation of outcomes is not a logical consequence of his metaethical position. His argument for utilitarianism goes something like this: "Since moral assertions are universal prescriptions, they forbid the making of exceptions; i.e., they must be responsive to the wellbeing of everybody equally. But giving everybody's wellbeing equal weight is the same as acting to produce the best overall outcome." The latter claim is false, however. Take a case in which acting to produce the best overall outcome requires harming one person significantly. An aggregative morality must legitimate so acting. But since I would prefer not being so harmed over any marginal personal benefit attained by so acting, I cannot, on universal prescriptivist grounds, propose acting in this way. Universal prescriptivism does not logically imply aggregation and thus does not lead to utilitarianism. My point here connects with Rawls' so-called 'separateness of persons' objection to utilitarianism.

<sup>3</sup> For a comprehensive bibliography of accounts of the cases for these positions see Finlay and Schroeder 2015, and for a crisp overview of the debates, see Millgram 2001.

<sup>4</sup> There is, however, at least one objection to these positions that it is incumbent on me to forestall: the objection that the truth of internalism implies the falsity of instrumentalism (this idea is perhaps to be found in Korsgaard 1986). The argument goes something like this: Suppose internalism to be true. Then my reasons for action are essentially connected in some appropriate way with the contents of my 'subjective motivational set.' Suppose also that – as part of this set – I hold a commitment to not utilising instrumental rationality. Now it seems that any insistence that I utilise instrumental rationality can only be premised upon the truth of

Nor will I attempt to provide a definitive account of the concept – fundamental to any internalist account of practical reason – of *desire*. Several remarks are in order here, however. I hope that my reasoning will be neutral as regards at least some of the different conceptions of desire in philosophy's marketplace. The essential proviso is that we take a conception compatible with instrumentalism. More specifically, it must be the case that desires and the ends of desires can either count as practical reasons for action themselves, or at least that they can appropriately give rise to practical reasons. This eradicates at least the following desire-conceptions from contention: conceptions according to which desires are solely causes of actions, dispositions to act, mental 'pushes,' or entropic rebalances towards psychic equilibrium; conceptions according to which desires are always the results of deliberation; behaviourist conceptions. But it leaves open choice between at least the following: conceptions according to which desires are attitudes characterised by a 'direction of fit' to the world; cognitivist conceptions, such as the affective conception of desire, according to which desires are 'joyful thoughts,' or the conception according to which desires are representations of welfare, benefit, etc., or the semiotic conception according to which desires are semic hierarchies.

There is a host of vocabulary in English to refer to the multifarious phenomena I gather under the single term desire: commitment, need, value, demand, wish, want, intention, inclination, proclivity, predilection, drive, instinct, ambition, hope, fear, ideal, pro-attitude, etc. These are all possible elements of what Williams calls the "subjective motivational set" [SMS] of an agent (1981, 103). It may, in some contexts, be important to distinguish between these as different sorts, but here such distinctions will be left out of focus for the sake of the simplicity and generality of the account.

---

externalism – i.e., it can't appeal in the appropriate way to my motivations. So it seems that we can't both be internalists and instrumentalists. An initial distinction that might be made in response to this supposed inconsistency is between descriptive and prescriptive instrumentalism. Descriptive instrumentalism would be the psychological thesis that, at least in large part, when persons reason practically they in fact reason instrumentally. Prescriptive instrumentalism would be the view that instrumentalism sets the standards of practical reason, so any purported practical reasoning that is not instrumental is thereby irrational. On the descriptive version, the above objection fails since there is no principled contradiction between the internalism requirement and the fact that persons for the most part reason instrumentally when they reason practically. Such a solution will not do for me, however, since prescriptive instrumentalism is required by the ensuing logic. Rather, the response should be that the establishment of instrumentalism as the correct account of practical reason cannot itself be the result of a process of practical reason, since that would be straightforwardly question begging. Instead, instrumentalism about practical reason must be a result of theoretical reason, if it is to be a result at all. But in this case the problem above is solved since the argument for instrumentalism does not need to be premised on internal reasons. On the basis of theoretical reasons – reasons that I won't here outline – we are entitled to claim that the person who is committed to non-instrumentalism is irrational and involved in paradox without contradicting internalism.

## What Would a Deontic Logic of Internal Reasons Look Like?

Every desire 'aims' at an end, *telos*, goal, target, etc. I take this teleological structure of desire to be uncontroversially the basic structure of desire. Desires are specified by reference to their *teloi*, and are thus individuated one from another by their respective ends. Ends can be means to other ends: desires *nest* within each other. I may leave the house desiring to buy cat food – but only because I wish to feed the cat. Since it is implausible to think that finite agents could have infinitely many distinguishable desires, in order to avoid circularity in justifications, instrumentalists are generally committed to the existence of what we might, with Aristotle (1984, 1097a), call *final ends* – end that are no means to further ends. Final ends would be the ends of what I will call *ultimate desires*, taking the adjective in its etymological sense as last, final, farthest, most distant, extreme. There may be psychological limitations on the scope of possible final ends, but this will have no bearing on the argument below, and I need not claim, in contrast to Kantians, for example, that there are any ultimate desires that are shared necessarily by all agents.

Just like other ends, final ends do not need to be specified clearly or consciously in order to conform desires – desires are often significantly non-thetic.<sup>5</sup> Ultimate desires do not need to meet stringent persistence conditions: they can be transient and fleeting. A single agent may have one or many ultimate desires concurrently. If an agent has many concurrently, some of them may conflict, in the sense that the satisfaction of one implies the dissatisfaction of others.<sup>6</sup> Whilst final ends are not subject to rational scrutiny *in vacuo*, they may be subject to rationalisation from the perspective of further ultimate desires. As

---

<sup>5</sup> The specificationist view of practical reason – the view that at least a substantial part of what we do when we reason practically is to specify ends more precisely in a way that enhances the overall coherence of our projects – has been proposed as a genuine alternative to instrumentalism (Kolnai 1961, Richardson 1990). Whilst I think that the specificationist insight is a genuine and important one, I think the instrumentalist can quite easily accommodate an account of the phenomena specificationists lean on. In the first place, coherence amongst our projects, plans and desires is itself a commitment, part of one's SMS, and so reasoning that attempts to better it by trimming one's ends, so to speak, can be subordinated as instrumental to a further end. Second, specification of ends that takes the form of becoming more clear about what it is actually that one wants can be analysed either as specification of the best means to some further end (being clearer about one's ends so as to facilitate their realisation), or as *theoretical* reason used in service of practical ends. Suppose I have decided to clear my debts. One thing I might do is specify more precisely what this will mean practically by summing my debts together. But summation is a function of theoretical, not practical reason. So pointing out that theoretical reason is used – only ever? – in service of practical reason protects the instrumentalist's account from such specificationist worries.

<sup>6</sup> So-called commensurability of ends – that is, the possibility of assigning them all values on a single hierarchical scale – may be instrumentally advantageous to individuals, but it is not an essential commitment of an instrumental account of practical reason. However, the possibility of conflictual final ends certainly introduces into any logic of instrumental reasons considerable complexity – complexity that I here grant myself the favour of avoiding since its analysis is tangential to my aims. See Chang 1998 for an overview of the debate.

Christoph Fehige writes (2001): “We can have reasons to adopt, or to make ourselves adopt, new desires and to dispose of existing ones. These reasons... will bottom out, say, in our desire to have, or not to have, certain desires, or in the fact that certain desires cannot, or cannot jointly, be satisfied.” Having so-called maieutic desires – desires about what desires or ends to have, even what final ends to have – does not take us out of the reach of instrumentalism (Schmidtz 1994).

How should we understand the relationship between desires and practical reasons? The so-called neo-Humean theory is typically held to propose two alternatives: either desires just are practical reasons, so the relationship is one of identity; or desires are necessary conditions of practical reasons, such that something is only the practical reason that it is in the presence of a particular desire – if it is bundled up, so to speak, with a particular desire. On the former view, the desire to get cat food is my reason for leaving the house; on the latter view, my reason for leaving the house can be construed to be the *fact* that the closest cat food available is in the shop around the corner, but only if taken in a bundle with – i.e., taken to be dependent on – the desire to get cat food (as efficiently as possible). Mark Schroeder (2007) has shown that these views have substantially different implications, but none of these implications bears significantly on the logic that follows here, so I will leave this issue unarbitrated. At any rate, in the simplest case (i.e., where we have only to do with a single desire): Desire,  $D$ , (or whatever it is bundled up with) of agent,  $S$ , is a practical reason for  $S$  to  $\varphi$  iff  $\varphi$ -ing is taken by  $S$  to facilitate most effectively – or, at least, effectively enough – of the options practically available the realisation of the target of  $D$ .

Taking this still rather rough but sufficiently determinate concept of desire in conjunction with instrumentalism and internalism about practical reason, one is committed to the following view: for any given agent,  $S$ , her practical reasons bottom out in her ultimate desires and their associated final ends, and thus her ultimate desires circumscribe for her the region from which she can draw practical reasons. On these assumptions, no purported reason that cannot be embedded appropriately within the end-seeking of one or more of  $S$ 's ultimate desires can be a practical reason for  $S$  at all.

### 3.1. Axioms

In the first place, an axiom that is henceforth implicit for every proposition that follows:

0. *Subsumption axiom*: If an agent,  $S$ , can in good faith describe her action,  $\varphi$ , by descriptors constituting any of the following propositions, then  $\varphi$  is subject to the norms expressed in and implied by those propositions.

‘In good faith’ captures two conditions: one, that  $S$  believes that she can accurately so describe her action, and, two, that  $S$  believes that she is sufficiently informed to make such a judgement. If either of these conditions fails, the action is not, by my lights, subsumed under the norms expressed by the following

propositions. Note that there is no requirement for objective standards of description here. This is a result of the fact that, on the internalist view, the sorts of practical reason that *S* has are not functions of the way the world is independently of *S*, but are rather constitutions of her SMS.

I endorse two norm-giving principles that are staples of many moral philosophical theories because they appear to be necessary postulates of any account of practical reason:

1. *Generalisation axiom*: the semantic import of any given non-indexical assertion remains the same across all its utterances.

This principle is found well expressed and developed in Hare's work<sup>7</sup>; he considers it a necessary part of any analysis of the semantics of assertions, not solely of normative assertions. It implies the following: If I genuinely assert "that leaf is green," I thereby normatively commit myself to the assertion that anything relevantly like the leaf indicated, i.e., of a conventionally similar enough colour, in any relevantly similar context, i.e., appropriately similar lighting conditions, is green. I thus commit myself to not asserting this sort of thing: "leaf *X* is green; leaf *Y* is the same colour as leaf *X*; and leaf *Y* is not green."

Distinguishing qualitative from numerical difference is useful here. The generalisation axiom is an implication of the fact that assertions track qualitative differences – differences in properties – across differences in numerical identities. Given two qualitatively identical but numerically distinct objects, the set of assertions that holds for one, will, *ipso facto*, hold for the other. Thus a distinction solely in numerical identity cannot ground a difference in predication.

Note that this characteristic of assertions is what enables any reasoning whatsoever. To reason is, at the very least, to extrapolate from a given set of propositions to those implied by them. This is only possible on the basis that the given set of propositions contain general statements or rules – that they refer to universals. This enables implied propositions to be identified as those that can be subsumed under the rules. Without the generalisation of predication, language would lack the rules or general statements that enable reasoning. Practical reasoning utilises the same generality. Thus some form of the generalisation axiom is going to be essential to any account of practical reasoning.<sup>8</sup>

---

<sup>7</sup> See, in particular, his 1963, §2.2 and *ff*. Note that his nomenclature is importantly different on this from mine; he speaks of universalisability (c.f., *ibid*, §3.4), not generalisation. I use the latter term so as to forestall confusion with a Kantian approach to practical reason, and for want of a better term. See note 8 below.

<sup>8</sup> It is important to distinguish generalisation as here outlined from the *universalisation* which is at the heart of Kantian approaches to practical reason. To see the difference, consider the questions that would be required by these procedures respectively in regard to the assertion "this leaf is green." On the generalisation procedure, I ask myself whether I am prepared to assert that anything relevantly like the leaf in any relevant similar context is green. On the universalisation procedure, on the other hand, I would have to ask myself whether everybody could assert that anything relevantly like the leaf in any relevantly similar context is green. The

2. *Impartiality principle*: personal identity *per se* is not a deontically relevant difference.

My criterion for determining whether a difference between two circumstances or actions is deontically relevant is the same as Hare's: some factor (in the most general sense) is deontically relevant iff it forms part of the justification for the duty or norm expressed.<sup>9</sup> Anticipating the discussion below, given the definition of practical reasons herewith outlined, such justification can only be the envisaged means by which a desire is to be fulfilled. Thus neither aspects of an action or circumstance which have no bearing on the fulfillment of relevant desires, nor such aspects as prevent overall such fulfillment can be relevant differences. Thus it is not practically rational for an agent to lean on some wholly arbitrary difference between her and others to explain why it is legitimate for her to steal, but not for others. Nor would it be practically rational for an agent to lean on some difference that on an appropriately comprehensive and balanced consideration of counterfactual situations obstructs rather than facilitates the fulfillment of her desires. So, for example, I cannot claim that those who are especially hungry are permitted to push in at the front of the lunch queue when I am especially hungry, if, when I am only a little hungry, this policy will mean that I get nothing to eat.

Expressed in an alternative way: a difference between two circumstances or actions is deontically relevant to a given asserter iff the asserter is prepared to assert different deontic assertions regarding the two circumstances or actions respectively. For example, suppose a person accompanying small children joins the lunch queue. I may be prepared to assert – upon an appropriately comprehensive and balanced consideration of counterfactual situations – that those accompanying small children should be permitted to go to the front of the lunch queue, whilst those who are not, like me, should not. In this case, I am taking accompanying small children to be a deontically relevant difference.

Personal identity *per se* cannot be taken to be deontically relevant without contradicting the generalisation axiom. Suppose the impartiality principle to be false and thus that personal identity *per se* can constitute a deontically relevant difference. This would imply that not all normative assertions would maintain

---

latter would lead to very different commitments to the former. In particular, on the latter account, I need to be sure that everybody sees the leaf as I do before I make my claim, whereas on the former account, the ways that other people see the leaf are irrelevant to the commitments of my assertion. Since different people no doubt sometimes do see the very same emission of light in different ways, in particular, those with colour vision deficiency might not, for example, distinguish red from green, the universalisation procedure would undermine our ability to make assertions about colours. My analysis, below, in the case of practical assertions will be parallel: what counts will be whether I am prepared to accept relevantly similar actions in relevantly similar circumstances, in particular, in those in which I am an affected party. The test will not be "would I like it if everyone does it, or is it possible for everybody to do it?". The question will rather be "would I like it if it was done to me?".

<sup>9</sup> Hare, 1963 §§2.2-2.5, 6.8.



their truth-value across all situations identical in respect of their properties, and thus that the semantic import of such assertions would not remain the same across all their potential utterances. In this case, the generalisation axiom would not hold. So the truth of the generalisation axiom implies the truth of the impartiality principle. Differences in personal identity *per se* are mere numerical differences between persons, not qualitative differences, so they cannot ground differences in deontic predication. Indeed, on many analyses, it is part of what we mean by a moral norm or duty that it applies, if it applies, regardless of the bare personal identity of the agent. On most understandings, to take bare personal identity as a normatively relevant difference is precisely to abandon the deontic domain – and certainly the domain of rational discourse about norms.

### 3.2. Logic and Commentary

#### 3.2.1. The Enkratic Principle

Ceteris paribus:

3. One cannot desire that one does not realise what one ultimately desires.
4. One cannot be indifferent about whether one realises what one ultimately desires.

I propose these two as analytic. If one desires that one does not realise something, it cannot be the case that one ultimately desires that thing. Likewise for indifference. The *ceteris paribus* clause controls for situations in which an agent has two or more contradictory ultimate desires, say an ultimate desire for *X* and an ultimate desire for not-*X*, and extends across all subsequent propositions. Removing this complexity makes the presentation simpler. The result is that the ensuing propositions are relativised to single ultimate desires and the subsidiary desires nested in them.

5. [From 3, 4] One must desire that one realises what one ultimately desires.

'Must' is used henceforth in its familiar sense as norm or obligation-giving. Two different senses of this sort of use can be distinguished, however, having to do respectively with, on the one hand, norms of logic or reasoning, either epistemic or practical, and, on the other, norms of morality. Here, I want both senses to be heard simultaneously: the reduction of moral norms to norms of practical reasoning that is being attempted here proposes that the use of 'must' (and its cognates) to express moral obligations is to be analysed in terms of its expression of obligations of practical reasoning. 'Ought to' is avoided until a later stage for the sake of distinguishing my account from others that accept, implicitly or otherwise, a fundamental distinction here.

6. [From 5] One cannot desire that one does not attempt to realise any given ultimate desire.

7. [From 5] One cannot be indifferent as to whether one attempts to realise any ultimate desire.

6 and 7 follow given, on the one hand, a very plausible construal of wanting whereby one who wills the end wills also the means: here's Kant: "Who wills the end, wills (so far as reason has decisive influence on his actions) also the means which are indispensably necessary and in his power. So far as willing is concerned, this proposition is analytic." (2012, Ak 4:417), and, on the other hand, given that the ultimate desires in question are ones, the realisation of which, to a degree at least, depend on actions of the agent herself. Whilst no doubt one could have ultimate desires the realisation of which depends *entirely* upon the actions of others or on environmental factors, such as the political or economic situation, reasoning with regard to them is not practical, since it can provide no reasons for actions, and thus is not subject to the norms of practical reasoning.

8. [From 6, 7] For any ultimate desire, one must desire to attempt to realise it.

9. One cannot accept anything as an overriding reason not to attempt to realise an ultimate desire (except, of course, the protocols of another ultimate desire – but such is controlled for by the *ceteris paribus* clause).

This is a result of internalism, instrumentalism and my definition of ultimate desire: ultimate desires circumscribe possible reasons for practical rationality; thus nothing can trump them in reason-giving.

10. [From 8, 9] *The enkratic principle*: (on pain of irrationality) I ought to attempt to realise what I ultimately desire.

This formulation of the enkratic principle differs from the way it is formulated currently in the literature.<sup>10</sup> John Broome (2013) and Andrew Reisner (2013), *inter alios*, for example, have recently formulated it as so: "Rationality requires you to intend to do what you believe you ought to do." The most obvious shortcomings of this way of expressing the principle are as follows:

In the first place, *enkrasia* is juxtaposed to *akrasia*. *Akrasia* has had all sorts of interpretations, but none of these, to my knowledge, has been expressed *wholly* in terms of failing to intend to do what one believes one ought to do. Rather, *akrasia* is always defined in terms of the necessary condition of a failure of *action*. One would expect, then, that the enkratic principle would concern action rather than mere intention. After all, I might intend to do what I believe I ought to do, but at the moment of action suffer from a failure of nerve: I have not manifest *enkrasia*.

Secondly, the word 'ought' is here in the wrong place, at least for my purposes. For the meaning that should be ascribed to 'ought' is precisely to be determined. Too many questions are begged by this reference to what the agent believes she ought to do. Further, the most obvious explanation of what the 'ought' would amount to would invoke traditional understandings of morality in some

---

<sup>10</sup> See, for example, the collection of papers in *Organon F*, International Journal of Analytic Philosophy, 20(4).

way or other. But surely one can be enkratic when such traditional moral norms are not at stake. Of course, on the theory being developed here, all action in breach of 10 would be construed as 'immoral.' Finally, Richard Holton's recent analyses of weakness of will suggest that at least aspects of the phenomena we refer to as weakness of will concern failures to act on resolutions regardless of whether we think we ought to act on them (see, in particular, his 1999 and 2012 – see also McIntyre 2006).

The enkratic principle, 10, targets certain forms of weakness of will, forms that can be analysed in terms of acting on a maxim that one – upon some balanced consideration – believes one has overriding reason to reject, i.e., not to act upon. This might not cover every aspect of akrasia and weakness of will that is philosophically interesting, nor every meaning associated with the lay conceptions of these phenomena, but my intention here is not to give a comprehensive account. I will offer here, however, a few clarificatory remarks regarding this analysis.

The word 'maxim' here does not support any significant philosophical weight and is used *in lieu* of 'principle.' By 'maxim' I mean just the subjective description(s) with which the agent would describe her action. Thus if she understands herself simply to be quitting smoking, then her maxim is simply "quit smoking;" if she understands herself to be producing illicit graffiti art as a form of protest against her authoritarian government, then her maxim would be "produce illicit graffiti art as a form of protest against the authoritarian government." Note that any given action may support multiple subjective descriptions. "Quit smoking," may perhaps equally be "don't light up another cigarette." Given the subsumption axiom, however, the agent is weak of will in the relevant sense if she can *in good faith* describe her action in terms of any maxim that she takes herself to have overriding reason to reject.

The agent must *believe* that she has such an overriding reason. Whatever our analysis of belief, a necessary condition here is that the agent *takes herself* to have overriding reason to act in a certain way. There is a natural usage of 'having an overriding reason' that fails to satisfy this condition – if I take the wrong keys with me when I leave the house, thus locking myself out, there is a perfectly good sense in which I had an overriding reason to take the set of house keys, not the ones I actually took. Since I am focused on some sort of failure of will, rather than a straightforward cognitive failure, however, the condition of belief must be written in, in order to hold the cognitive variable constant and avoid identifying such cognitive issues as issues of willing.

The agent must believe that she has an *overriding* reason. Taking into account all the reasons for acting that the agent takes herself to have, the overriding reason is the output of the balancing up of all these reasons. In the language of contractualism, the agent cannot reasonably reject the action that the balancing of reasons has justified. Exactly how this balancing of competing reasons for action should proceed, whether the agent really does have an

overriding reason for action according to some objective appraisal, and whether the agent has really taken all relevant factors into account in her balancing, is, to my purposes here, insignificant. All that is important is that the agent *take herself* to have an overriding reason. If her balancing of reasons results in a stalemate – two actions, say, that she takes to be equally justifiable – then, on my account, she cannot exhibit weakness of will by acting in one of those ways rather than the other, but she could exhibit weakness of will by doing some third action that she takes to be unjustifiable in her present circumstances relative to them.

On the view sketched out here, 10, or some variant of it, is the primary norm of practical reason – and thus the cornerstone of normativity. It is primary in several senses: In the first place, it is the *practical* pre-condition of the satisfaction of all other practical norms. Unless, one at least attempts to attain what one has overriding reason to achieve, one cannot satisfy norms that are associated with overriding reasons, no matter what they might be. In the second place, it is the *theoretical* pre-condition of the satisfaction of all other norms. Unless I also accept that I should attempt to act compassionately, for example, my acceptance of the norm of compassion is practically meaningless. Thirdly, it is normatively primary: whilst, for reasons that should already be clear, it is impossible on my account for 10 genuinely to conflict with other moral norms, should a conflict apparently occur, 10 would be overriding, since no practical reason can possibly count against it. Finally, my integrity as an agent would be threatened by any account that failed to normatively prioritise 10.

10 locates the *moral pivot* – the fulcrum between right and wrong, or good and bad – at a logical and psychological position somewhere other than that of the traditional theories of the ‘morality system.’ Rather than as the moment of harm, of autonomy-infringement, of disobedience to the will of god, 10 specifies failing to do what one believes one has an overriding reason to do as the basic form of ‘immorality.’ This proposal liberates moral thinking from an almost exclusive focus on relationships to others, and widens it to include, first of all, one’s relationship with oneself. It implies that included within the normative domain are *duties to oneself*.<sup>11</sup> But the view that there is something morally problematic

---

<sup>11</sup> Marcus Singer (1959) questions the very coherence of maintaining that there are duties to oneself. He rests his argumentation on three assertions: that duties always have correlative rights; that those with rights can waive them, thus releasing those subject to the correlative dutiful obligations; that one cannot release oneself from obligations (1959, 203). Taken together, he argues, these imply that the idea of duties to oneself is nonsensical: if the person who is under the obligation is also the person who has the correlative right, then that person would be able to release herself from the obligation by waiving her right. Since, by the third assertion, the latter is impossible, it cannot be the case that any person can have an obligation to herself, and thus there can be no duties to oneself. All three of these assertions can be challenged, however.

In the first place, it seems that not all duties have correlative rights, or at least that it is not contradictory to deny this. Take the duties that some hold we owe to the environment. Whether or not there are such things, those deep ecologists that advocate such duties do not

about weakness of will both has significant philosophical precedent and is familiar from common parlance and popular moral intuition. Just a smattering of heterogeneous ideas from the history of philosophical ethics that are all situated at least very closely to this insight: Alain Badiou's idea of a truth-procedure (2001, Chapter 4); Jacques Lacan's imperative not to give up on one's desire (2007, Chapter 24); Philippa Foot's conception of virtues as engaging the will (2003); Heidegger's conception of resoluteness (2001, §62); Nietzsche's injunction to become what you are (1999, §270); Kant's duties of self-cultivation (1996, 6:417 ff); the sin of sloth in Aquinas and the Roman Catholic natural law tradition (1948, Pt. II-II, Q. 35); Stoic ethics of self-cultivation (see, for example, Foucault 1998). On the other hand, we are all well aware of the pangs of conscience that accompany and threaten laziness, the moral disapprobation attached to 'giving in too easily,' the heroism of mind over matter, the moral approval of strong characters. Willpower is something that we all wish we had more of, and judge our worth at least partly in terms of.

Hare famously argued that weakness of will as traditionally understood was not possible. Hence, insofar as Hare's influence is all over this paper, it is

---

need to claim that the environment has in any sense rights correlative with them. Kant's duties of self-cultivation too are put forward absent correlative rights, as are our duties towards the dead – to execute their wills, carry out their wishes for burial, cremation, etc. It thus seems that we have no trouble making sense of the concept of dutiful obligations that are not accompanied by correlative rights.

On the other hand, given a suitably broad construal of rights according to which at least one sort of right amounts to nothing more than a sort of justified normative expectation, the duty to oneself that I have attempted to establish is accompanied by such a right: when I am weak willed, my attendant feelings of disapprobation, vice and disappointment in myself are justified. Conscience is clearly activated by being weak willed.

Singer's second proposition asserts that those with rights can waive them. Whether or not it makes sense to speak of a right in regards to weakness of will, this assertion is straightforwardly controversial. It is often thought not to hold in the case of the right to life, the right that corresponds to the duty not to kill. It is not obvious that if you want me to kill you, I am thereby morally permitted to kill you, that I am released from my obligation not to kill you. If intuitions were clear and consistent here, then the physician-assisted dying debate would not be so tricky. Further, by definition, one cannot waive the rights not to be tortured, raped, enslaved, etc. If I consent to you causing me pain and suffering, you are not torturing me; if I consent to you having sex with me, you are not raping me, and so on. Perhaps the right (if indeed there is one) associated with the duty of strength of will is of the sort that cannot straightforwardly be waived.

Singer's third proposition asserts that one cannot release oneself from obligations. As regards, weakness and strength of will, it is certainly not the case that one can release oneself from the obligation as here outlined as and when one wants: rather, one is released from one's obligation only when one takes oneself to have an overriding reason not to continue with one's resolution, project, etc. One is released from such an obligation by one or more *reasons*. Thus the account of obligation offered here accords with Singer's third assertion, but this does not make obligations in this sense any less obligations that one owes to oneself: one can have obligations to oneself from which one cannot release oneself at will.

incumbent on me to explain my disagreement with him over this issue. Hare takes the impossibility of genuine weakness of will to be a straightforward implication of his analysis of moral language in terms of universal prescriptivism: "It is a tautology to say that we cannot sincerely assent to a command addressed to ourselves, and at the same time not perform it, if now is the occasion for performing it, and it is in our (physiological and psychological) power to do so." (1952, §2.2, 20) Phenomena that appear to be instances of *akrasia* are thence explained away in one of two ways: either the 'oughts' involved are understood as universal prescriptions but the agent is psychologically or physiologically unable to act on them, or the universally prescriptive nature of the 'ought' is looked away from at the crucial moment of action, due to selfishness, in favour of a mere universal *description*, and is subsequently replaced by feelings of guilty conscience – in other words, we have a case of 'special pleading:' the agent applies the prescription to everyone but herself (1963, §§5.6-5.8, 76-80). The standard objection to Hare's account here is phenomenological: do we not find ourselves sometimes in situations in which we fail to do what we know we ought without ceasing to apply the 'ought' to our own case, and in which it really cannot be argued that we *could not* have performed the 'ought?' My proposal is that the philosophical analysis should give way to the phenomenological evidence here. Why does Hare maintain such a denial of the reality of weakness of will? No doubt, it is because he believes that his prescriptivist analysis of moral language would otherwise be undermined. But exactly why that should be the case is considerably under-argued. In the passage quoted above, he claims that it is simply a matter of *tautology*. It doesn't seem to me to be so. I can imagine sincerely assenting to a command addressed to myself and failing to act on it, despite the fact that the moment of action is at hand and it is within my power to act on it, just because I am too frightened – say, going over the top of the trench. Hare would reply that if I am too frightened, it is not within my psychological power to act; I would disagree, or at least require him to provide argument to this effect. To support my contention: suppose that the person who uttered the command, noticing my reluctance to obey, ups the stakes of my disobedience by threatening some sanction if I don't act as commanded – perhaps that I were to be executed. Now I act. What has happened here? On Hare's analysis, the commander has changed the psychology of the situation sufficiently such that I am now capable of acting; on my analysis, the commander has just strengthened my motivation to act by giving me a more powerfully overriding reason. I don't see any reason to accept Hare's view rather than my alternative. And so I don't see any reason to accept the claim that the truth of prescriptivism necessarily implies the illusoriness of weakness of will.

Several subsidiary principles of practical reason should be inserted at this point with regard to attempts to realise the final ends of ultimate desires. I owe a significant debt here to Onora O'Neill's analysis in her 1985:

## What Would a Deontic Logic of Internal Reasons Look Like?

10.1. One ought to intend all means that one believes are necessary and at least some means that one believes are sufficient to the final end.

10.2. One ought to seek to make such means available when they are not.

10.3. One ought to attempt to ensure that the various means that one pursues in order to realise an ultimate desire are mutually consistent.

10.4. One ought to attempt to ensure that the foreseeable results of actions undertaken in order to realise an ultimate desire are consistent with the realisation of that desire.

These principles amount to a set of *norms of prudence*: norms of practical reason that govern the selection of means appropriate to given final ends.

### 3.2.2. Tolerance Principles

To make this section manageable, it is important to have at hand a distinction between two sorts of practical reasoning which I call first- and second-order. Second-order practical reasoning concerns what can most broadly be called 'punishment:' the norms surrounding the consequences of contraventions of other norms. 'Adulterers should go to prison' might be an example of this sort of norm: adultery is considered to be some form of infraction; going to prison is a form of punishment. First-order practical reasoning, on the other hand, is reasoning about norms concerning actions prior to any relevant contravention: for example, the norm "there is nothing intrinsically wrong with extramarital sex." I believe that the vast majority of our explicit moral reasoning is in fact second-order, but the postliminary discussion is expressly limited to first-order practical reasoning.

11. [From 1, 2] For any action  $\varphi$  in circumstances  $C$  that one accepts the permissibility of, one must accept for everyone the permissibility of actions relevantly similar to  $\varphi$  in circumstances relevantly similar to  $C$ .

I am obliged to accept the permissibility of at least some actions, since I am obliged by 10 to act, and any action that I perform under 10 is an action that I cannot accept the impermissibility of (because of 9). On the account developed here, obligation can be reduced to permissibility. In any given circumstances, I am obliged to  $\varphi$  iff  $\varphi$  is the only permissible action.

12. [From 10, 11] One ought to attempt to realise one's ultimate desire in action, but only in those actions that one accepts as permissible for any in relevantly similar circumstances.

We are now in a position to make quick work of showing what is irrational in general about the amoralist's stance.<sup>12</sup> The amoralist is one who refuses to acknowledge normative constraints upon her actions. Given 10, however, and the provision that she has some desires at the relevant moments, she is rationally

---

<sup>12</sup> C.f., Hare, 1963, §6.6.

Rufus Duits

obliged to act, but, given 2, her actions are subject to 12, and thus she is irrational not to acknowledge some normative constraints on her action. A simplified case: Suppose, walking alongside a river, she feels

$D_1$ : a whim to push a passerby into the turbulent current.

Suppose also that she has

$D_2$ : a commitment to staying alive.

Taken alone,  $D_1$  provides a reason to  $\varphi_1$  (push the passerby into the water). Given the impartiality principle, there is no normative difference between the amoralist being pushed into the water by the passerby and the passerby being pushed into the water by the amoralist, unless there is at least one deontically relevant difference between the two scenarios or agents. So  $D_2$  provides, *ceteris paribus* (i.e., absent the identification of such a difference by the amoralist), a reason to  $\varphi_2$  (refrain from pushing the passerby into the water). But  $\varphi_1 \neq \varphi_2$ : the amoralist cannot do both. The action she has most reason to do is the one that functions as best means to the overriding desire. Suppose this to be  $D_2$ . In this case, *ceteris paribus*, the passerby acts irrationally by doing  $\varphi_1$ .

13. [From 3, 10] One cannot desire that others prevent one from attempting to realise what one ultimately desires.

14. [From 4, 10] One cannot be indifferent as to whether others prevent one from attempting to realise what one ultimately desires.

15. [From 13, 14] One must desire that others do not prevent one from attempting to realise what one ultimately desires.

16. [From 2, 10, 11, 15] One ought not prevent others from attempting to realise what they ultimately desire – unless their pursual of their ultimate desires conflicts with one's pursual of one's own, and, upon appropriate comparison, one takes oneself to have reasons overriding to pursue one's own ultimate desires, and one adheres to proper procedure: *refraining from coercion; negotiating; seeking consent*.

12 and 16 amount to *principles of tolerance*. They concern the regulation of situations in which the attainment of the ends of one person's desire conflict with the attainment of the ends of another person's desire. Hare gives an example of a trumpeter, which could be useful here. Suppose the trumpeter's desire to play her trumpet conflicts with my desire to get an early night. In order to think about how such a conflict of interests can rationally be resolved, it is illuminating first to consider how conflicts of different desires belonging to the same person can rationally be resolved. Suppose I am tired and would like to go to bed early, but suppose also that I have a trumpet concert tomorrow for which, I judge, I ought to spend a little more time practising. I need to judge which piece of behaviour – retiring or playing – best serves my ultimate desires. This will require consideration of the relative importance to me of my respective ends, and of the efficacy of the means at my disposal to realise them. I may feel, for example, that,



although I need a little more practice, I am very tired, and I might therefore judge that my time would now be better spent resting.

Importing this procedure into the original scenario: I knock on the trumpeter's front door and ask her to stop playing – the noise is interfering with my ability to sleep, and thus preventing me from getting the rest I require in order to continue to fulfil my ultimate desires tomorrow. Perhaps I say to her: "Imagine if I was making a loud noise late into the night such that you couldn't sleep; how would you like it then?". Given 16 and the impartiality principle and the generalisation axiom, to respond rationally, the trumpeter needs to take account of my perspective. But suppose she weighs it against her own and concludes that her need to practice the instrument this evening for her important concert tomorrow is more significant relative to her ultimate desires than my need to get to sleep is significant relative to mine. Perhaps, she reasons, her concert is a one-off event that will make or break her career, whereas if I get a little less sleep the knock on effect that that will have on my ability to realise my ultimate desires will be negligible. In such an instance, the trumpeter may rationally conclude that she ought to go on practising despite my complaint.

This might seem counter-intuitive: isn't the right thing to do always to respect the goals and interests of others? Isn't the definition of immorality precisely the prioritising of one's own requirements over those of others – as our trumpeter is doing?

I want to respond to this query in two stages, the second of which I leave implicit in proposition 18 below. One important feature of the trumpeter case is that some sort of negotiation takes place. As a neighbour who needs some sleep, I want the trumpeter to shut up, but I must also realise that were I in her situation, or in a situation relevantly similar, I would also want to be able to practice my instrument. In neither position, would I wish to be *coerced* or forcibly prevented from pursuing my interests. In which case, I cannot, without being irrational, attempt to coerce or force the trumpeter to refrain from playing – at least in the first instance. What I can do rationally, however, is *negotiate*. Perhaps I could encourage her to use a mute or a different room so that both of our desires can be fulfilled. Perhaps I could encourage her to play just for another ten minutes, as I might decide myself to do were I the tired trumpeter above. Perhaps I could get her to make her practice worth my while by insisting on a favour in return as a way of re-balancing the importance of the realisation of my desire against the realisation of hers. Perhaps, if I knew that half an hour's less sleep would result in a box of my favourite chocolates, I would reevaluate the importance to me of that half an hour of sleep.

*First-order practical reasoning* – reasoning prior to the perception of a contravention by another – , then, appears to carry, in cases of interpersonal conflicts of interests, the following four commitments: In the first place, I must be actively considerate of the interests of others, since I cannot but desire them to be actively considerate of my interests. Secondly, I must never resort in the first

instance to coercion or force. I would never in the first instance consent to the trumpeter torturing me with her trumpet; she would never consent to me confiscating her trumpet. Thirdly, instead, I must abide by an overriding commitment to negotiation to rationally resolve conflicts of interest. I would never rationally consent to my interests being down trodden before I had had a chance to negotiate them. Finally, consent is to be sort from all parties concerned before I can prioritise my interests over those of others. Any alternatives to these conditions would leave my own interests or ultimate desires too vulnerable, such that, due to 10, I could not rationally endorse them.

I recognise the fact that there is mountainous philosophical work to do on the analyses of consent, coercion, and what counts as legitimate negotiation that I am unable to pursue here.<sup>13</sup> In order to keep my lights trained on my current aims, I will assume for the sake of argument that such accounts would not in principle render my appeal to these concepts unworkable, leaving them rough but ready.<sup>14</sup>

Now, this response provides us with better ammunition against the so-called fanatic than Hare has. Hare's fanatic as presented in *Freedom and Reason* is someone who holds onto an ideal even when pursuing it conflicts with all other interests they may have. He gives the example of the Nazi who discovers that both of his parents are Jewish, and, instead of abandoning his racist Aryan ideals upon the clear realisation of their implications, buys his own one-way ticket to Auschwitz. In this early text, Hare despairs of using moral argument against such a Nazi – he is consistent and not irrational, even if mad. He claims, however, that

---

<sup>13</sup> I can offer, however, a few pointers as to the shapes accounts of coercion and consent appropriate to my aims here might take. Broadly in the Nozickian tradition (Nozick 1969), but contrary to a couple of aspects of his analysis: Firstly, to capture the sorts of phenomena that I want, I would not restrict analysis of coercion to conditional threats, including also applications of direct force – “occurrent” coercion (Bayles, 1972). Secondly, there would be no reason to posit as a necessary condition of coercion that it be explicitly communicated to the coercee. If I hide the trumpeter's trumpet so that she cannot play the next evening, I am involved in a form of coercion even if it is not signalled to her that I have hidden it. Thirdly, the analysis in the literature most compatible with the account of practical reason offered here is Mark Fowler's “normative conception” (Fowler 1982). Fowler writes: “A coerced agent must be faced with the choice of yielding to a threat or acting contrary to practical reason. He is forced to perform his deed because he literally has no reasonable alternative.” (331) I would analyse ‘contrary to practical reason’ in terms of overriding reasons. So part of the proposal would be: I coerce someone when I confront them with a threat that constitutes an overriding reason for them to do the deed I wish them to do.

As regards consent, on the other hand, I would endorse an account of its conditions that *requires* it to be signified in a communicative act to the relevant parties. This condition is requisite if consent is to work its “moral magic” (Hurd 1996) of transforming the morality of a situation. A fairly standard account of its other conditions would also be appropriate: that it be voluntary, intentional, informed, and given by one competent to do so (Kleinig 2010).

<sup>14</sup> I would also advocate, but here cannot justify, the claim that the analysis of practical reason offered here can substantially explain why it is that the phenomena referred to by these terms are immensely important to our ethics of conflicts of interest.

such fanatics are very rare, and concludes that the best strategy one can take against them is one of “persistent attrition” (1963, §9.8, 180), forcing them continually to empathise and to reconsider their position in the light of the interests that they would have were they in the position of their victims.

Hare construes the conflict between the Nazi and his victims here as one between interests and ideals. What entitles me to make a further logical move against the fanatic’s position is the placing of both interests and ideals into the more general category of desire. Hare himself undertakes a parallel move in *Moral Thinking* where he analyses the relevant conflicts as of *preferences* (1981, Chapter 10). But there his defence is confused by his appeal to utilitarian considerations (cf. note 2 above). The further logical move against the fanatic is to point out that, as regards first-order practical reasoning, he cannot judge that it is permissible to coerce his victim since he could not rationally permit himself to be so coerced by somebody else’s fanaticism. If he wishes – ultimately – to remove all Jews from the European continent then necessary to the first-order task is an overriding commitment to negotiation. Of course, such a further move may not in fact *convince* the fanatic.

If Hare had been writing today, he may have used the suicide bomber as his example. Such a person is a real-life example of someone who is willing to give up literally everything earthly for the sake of the cause of his ultimate desire. Could it not be objected that on my account his willingness to die legitimises his killing of others, since he is being rationally consistent? Since his overriding ultimate desire is for martyrdom, he does not seem to be snared in irrationality when he acts to kill others since, given the impartiality principle, his desire to kill is consistent with his desire to die.

13-16, however, enable us to analyse such cases in terms of conflicts of interpersonal ultimate desires. When it comes to reasons overriding, perhaps we really do have a dead heat. The suicide bomber has a fanatical desire to kill, which, given the consequences he is prepared to accept, appears to rival the desire his victim has to stay alive. But whilst he is prepared to accept being killed, he cannot, on pain of irrationality, accept the coercive or forcible thwarting of the fulfillment of his plan. But then he cannot without contradiction coercively or forcibly prevent others from fulfilling their ultimate desires. By failing to attempt to negotiate consent (whatever that would be), he fails to meet the conditions for his desires to count properly as normatively overriding.

### 3.2.3. Beneficence Principle

17. In many circumstances, one desires that others act to help one get what one ultimately desires.

The proposition that one desires that others act to help one get what one ultimately desires is an empirical assertion of social psychology that, plausibly, holds in very many circumstances for any agent.

18. [From 1, 2, 10, 17] The *beneficence principle*: When others are in relevantly similar circumstances to these, one must act to help them get what they ultimately desire.

This grounds norms of kindness or care. Whilst it is uncontroversial that such norms pertain, and, indeed, perhaps account for a great part of traditional morality, recent discussions have focussed on their scope and extent. Do they imply that I have obligations towards those at risk of famine or conflict in regions of the world remote from me? If so, how much am I obliged to help? Do I ever have such obligations to strangers above friends and family members? Whilst I will not be able to do sufficient justice to these questions here, it is worth remarking:

Contemporary utilitarians like Singer and, perhaps to a lesser degree, contractualists such as Scanlon advocate the greatest extent of the reach of this sort of obligation.<sup>15</sup> Indeed, on the traditional utilitarian approach, one should, for example, continue to give donations to relieve the lives of those in poverty until doing so would cause as much suffering to oneself as would be relieved by one's gift. This has been formulated as the 'over-demandingness objection' to utilitarianism: surely, it is argued, it goes wildly against our moral intuitions to suggest that we should give so much. Ensuing debate has focussed on the difficult-looking task of fixing a precise limit, or, at least, a formula for one, that marks off the extent of the obligations of beneficence from what is supererogatory.

My account here sidesteps the need to formulate an objective such limit, i.e., one that holds for all in all cases. Rather, on my account, the reach of obligations of beneficence will be set *relative* to each individual and each particular case according to the different values assigned to the respective interests that are relevant. For example, suppose I am criticised for failing to remember the birthdays of my close family and friends. Whether or not I am rationally obligated to try harder to remember these birthdays will depend, *ceteris paribus*, upon (a) the importance my family and friends place on their birthdays being remembered, and (b) the importance I place on not having to remember their birthdays. If I find it immensely irksome to remember birthdays and my family and friends don't mind too much that I forget, then I am not rationally conflicted if I don't try harder to remember their birthdays. Of course, if the remembering of birthdays is construed as a generic act of kindness, then it is going to be less likely that I can escape the obligation to try harder to remember birthdays since it is going to be very unlikely that I am a person who does not value kindness. If I merely enjoy the actions of those being kind to me, I ought, *ceteris paribus*, to be similarly kind to others.

Whilst such a response will only partly answer the challenge, since setting such *subjective* limits in particular cases may be just as difficult as setting one objective limit, I am at least able to provide a formula for very simplified cases: one has a rational obligation to give up until the point at which the welfare or

---

<sup>15</sup> Singer's classic statement is his 1972; for Scanlon's account, see his 1998, Chapter 5, §8.

benefit facilitated by the gift is equally important to the fulfilment of the beneficiary's ultimate desires as not giving is to the fulfilment of one's own. If I am not wealthy, or if I am wholeheartedly committed to my family, profession, or other projects, etc., this point may arrive very early on, and I may thus be obligated to give comparatively little. On the other hand, most of us in the HICs of the world would, by the lights of this formula, be obligated to give much more than we do.<sup>16</sup>

#### 4. Concluding Remarks

The most significant outcomes of the foregoing blueprint I take to be twofold. Firstly, I hope to have indicated how we might be able to have our cake and eat at least most of it – that is, one way in which we might be able to be internalists and instrumentalists about practical reason and at the same time endorse, with a rigorous grounding, many of the intuitive commitments of the 'morality system.' This amounts thus to showing the way towards an answer to the challenge posed by the so-called 'central problem' of internalism. Presented in this way, the mysterious *aura* that philosophers like to impute to moral obligation is deflated to plain instrumental practical rationality; despite not recognising special or 'queer' moral properties, fairly standard obligations concerning our treatment of others remain largely intact.

Secondly, I hope to have provided sufficient justification for the claims that, on internalist and instrumentalist assumptions, we have enkratic obligations towards ourselves, and that our obligations to others can be understood as indirect obligations to ourselves. The enkratic principle expresses our fundamental obligation towards ourselves, but it also founds our obligations to others since it is an essential part of the reason why one is irrational if one refuses to be held to account by the interests of others. Given the impartiality principle, it is because it is irrational not to pursue what one has overriding reason to pursue that it is irrational not to respect what others have overriding reasons to pursue.

This view is counterpoint to the liberal tradition that asserts that, if there are any duties to ourselves, they are to be construed as indirect duties to others. It is surprising that a tradition in which Kant has been so influential has arranged things this way around:

---

<sup>16</sup> It might be thought that approaching the problem of beneficence through an analysis of the normativity of practical rationality in this way leads to the conclusion that there are no supererogatory beneficent acts, since giving more than one has reason to is contrary to practical reason in a similar way to not giving as much as one has reason to. This may be a counter-intuitive result, since, in general, the existence of supererogatory beneficence is rarely doubted. The account here given, however, does not suggest that supererogatory beneficence cannot exist. Since a variety of actions may, in a given circumstance, be equally rational, there may well be permissible actions of beneficence that are not obligatory. It should also be noted that, by making the integrity of the agent central to the account, the proposal here is better suited than utilitarianism to weighing against one another possibly conflicting obligations to family, friends, etc., and to strangers.

[T]here is no question in moral philosophy which has received more defective treatment than that of the individual's duty towards himself... All moral philosophies err in this respect... Far from ranking lowest in the scale of precedence, our duties towards ourselves are of primary importance and should have pride of place... It follows that the prior condition of our duty to others is our duty to ourselves; we can fulfil the former only insofar as we first fulfil the latter... Our duties towards ourselves constitute the supreme condition and the principle of all morality. (Kant 1963, 117-121)

But, as Kant recognised, the view that one has obligations to oneself that are grounded independently of one's obligations to others is suggested by much common sense moral thinking: conscience bites not only when we conceive ourselves as having done wrong to others, but also and perhaps even more painfully when we fail, due to weakness of will, to fulfil the projects most dear to ourselves.

## References

- Aquinas, St. Thomas. 1948. *Summa Theologica*. Notre Dame: Ave Maria Press.
- Aristotle. 1984. "Nicomachean Ethics." In *The Complete Works of Aristotle*, edited by Jonathan Barnes. Princeton: Princeton University Press.
- Badiou, Alain. 2001. *Ethics: An Essay on the Understanding of Evil*. London: Verso.
- Bayles, Michael. 1972. "A Concept of Coercion." In *Nomos XIV: Coercion*, edited by James Pennock and John Chapman, 16-29. Chicago: Aldine-Atherton.
- Broome, John. 2013. "Enkrasia." *Organon F. International Journal of Analytic Philosophy* 20(4): 425-436.
- Chang, Ruth. 1998. *Incommensurability, Incomparability, and Practical Reason*. Cambridge, Massachusetts: Harvard University Press.
- Fehige, Christoph. 2001. "Instrumentalism." In *Varieties of Practical Reasoning*, edited by Elijah Millgram, 49-76. Cambridge, MA: MIT Press.
- Finlay, Stephen, and Mark Schroeder. Winter 2015 edition. "Reasons for Action: Internal vs. External." *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. URL = <<http://plato.stanford.edu/archives/win2015/entries/reasons-internal-external/>>.
- Foot, Philippa. 2003. *Virtues and Vices*. Oxford: Clarendon Press.
- Foucault, Michel. 1998. *Ethics: Subjectivity and Truth*, edited by Paul Rabinow. New York: The New Press.
- Hare, Richard. 1952. *The Language of Morals*. Oxford: Clarendon Press.
- . 1963. *Freedom and Reason*. Oxford: Clarendon Press.
- . 1971. *Practical Inferences*. London: Macmillan.
- . 1981. *Moral Thinking: Its Levels, Method and Point*. Oxford: Clarendon Press.
- Heidegger, Martin. 2001. *Sein and Zeit*. Tübingen: Max Niemeyer Verlag.
- Holton, Richard. 1999. "Intention and Weakness of Will." *Journal of Philosophy* 96(5): 241-262.

- Holton, Richard, and Joshua May. 2012. "What in the World is Weakness of Will?" *Philosophical Studies* 157(3): 341-360.
- Hurd, Heidi. 1996. "The Moral Magic of Consent." *Legal Theory* 2(2): 121-146.
- Kant, Immanuel. 2012. *Groundwork of the Metaphysics of Morals*, edited by Mary Gregor and Jens Timmerman. Cambridge: Cambridge University Press.
- . 1996. *The Metaphysics of Morals*, trans. Mary Gregor. Cambridge: Cambridge University Press.
- . 1963. *Lectures on Ethics*, trans. Louis Infield. Indianapolis: Hackett Publishing Company.
- Kleinig, John. 2010. "The Nature of Consent." In *The Ethics of Consent: Theory and Practice*, edited by Franklin Miller and Alan Wertheimer, 3-24. Oxford: Oxford University Press.
- Kolnai, Aurel. 1961. "Deliberation Is of Ends." *Proceedings of the Aristotelian Society* 62: 195-218.
- Korsgaard, Christine. 1986. "Skepticism about Practical Reason." *The Journal of Philosophy* 83(1): 5-25.
- Lacan, Jacques. 2007. *The Ethics of Psychoanalysis*. London: Routledge.
- McIntyre, Alison. 2006. "What Is Wrong with Weakness of Will?" *Journal of Philosophy* 103(6): 284-311.
- Millgram, Elijah. 2001. "Practical Reasoning: The Current State of Play." In *Varieties of Practical Reasoning*, edited by Elijah Millgram, 1-26. Cambridge, MA: MIT Press.
- Nietzsche, Friedrich. 1999. *Die Fröhliche Wissenschaft*. Kritische Studienausgabe, edited by Giorgio Colli and Mazzino Montinari. München: Deutscher Taschenbuch Verlag de Gruyter.
- O'Neill, Onora. 1985. "Consistency in Action." In *Morality and Universality: Essays on Ethical Universalizability*, edited by Nelson Potter and Mark Timmons, 159-186. Dordrecht: Reidel.
- Reisner, Andrew. 2013. "Is the Enkratic Principle a Requirement of Rationality?" *Organon F. International Journal of Analytic Philosophy* 20(4): 437-463.
- Richardson, Henry. 1990. "Specifying Norms as a Way to Resolve Concrete Ethical Problems." *Philosophy and Public Affairs* 19(4): 279-310.
- Scanlon, Thomas. 1998. *What We Owe to Each Other*. Cambridge, Massachusetts: Harvard University Press.
- Schmidtz, David. 1994. "Choosing Ends." *Ethics* 104(2): 226-251.
- Schroeder, Mark. 2007. *Slaves of the Passions*. Oxford: Oxford University Press.
- Singer, Marcus. 1959. "On Duties to Oneself." *Ethics* 69(3): 202-205.
- Singer, Peter. 1972. "Famine, Affluence and Morality." *Philosophy & Public Affairs* 1(3): 229-243.
- Williams, Bernard. 1981. *Moral Luck*. Cambridge: Cambridge University Press.
- . 1985. *Ethics and the Limits of Philosophy*. London: Fontana.