Non-ideal accessibility

I. Introduction

Moral philosophers generally agree that 'ought' implies 'can' for at least an important subset of 'oughts'.[1] But they tend to define 'can' fairly liberally, for example as 'ability plus opportunity' (Vranas 2007 169-70; see also Feldman, 1986 Ch. 2). So long as a person can do what they ought to do, on some reasonably circumscribed understanding of 'can', how likely they are to do it is irrelevant. The main motivation for this line seems to be that we don't want to let people off the moral hook too easily. The recent trend in political theory toward non-ideal theory has brought this motivation into question. Non-ideal theorists want to take seriously some of the non-ideal conditions of the real world:[2] imperfect compliance, corruption, poverty, greed, self-interest, apathy, bureaucracy, and also uncertainty, e.g. about others beliefs, intentions, and willingness to act when it comes to fulfilling collective obligations.[3] We can argue about the priority of ideal over non-ideal theory, but no matter the strength of our concern with what ideally ought to be the case, we should surely still have something to say about what ought to be the case given that certain of our more ideal prescriptions will not be the case. Surely moral philosophy does not 'run out' when it comes to the various kinds of moral failure. One implication of the actual fact of widespread moral failure is that moral and political philosophy should be comprehensive, i.e., should say something about betterness and not just best. (Moral philosophy should tell a person what she ought all-things-considered to do, but it should also make clear what the next-best actions are, 'all the way down' to the worst actions; political philosophy should tell a given collective (the government; an NGO or INGO; perhaps even "society" in general in some cases) what it all-things-considered ought to do, but also what the next-best actions are, all the way down, in the same

---

1 It seems very plausible to me that one role of ideal 'oughts' is to reveal standards and values, rather than guide action, in which case it is not necessary that they imply 'can'. See Lawford-Smith (2010). Peter Graham (2011) has argued recently that positive obligations imply 'can' while negative obligations do not, e.g., we ought not kill, whether or not we can help it, but we ought give to charity, only if we have the resources to. For a good survey of the "ought implies can" debate, see Vranas (2007). See also Fred Feldman's (1986) discussion of accessibility.

2 See the papers collected in (Robeyns & Swift 2008) and the references in (Simmons 2010).

3 On the duties of individuals to form collectives see (Collins forthcoming); on the role of members of collectives in discharging collectives' duties see (Lawford-Smith forthcoming[a]).

way.[4] But there is also a question which is bound up in discussions about the demandingness of morality (Goodin 2009), namely what a realistic, non-ideal morality actually requires of real, non-ideal people. Maybe it doesn't require nearly as much as proponents of 'ought implies can' (and only 'can') have suggested that it does.

Obviously, an outcome being unlikely isn't enough to make it not required. But surely there's a better way to constrain 'ought', between the two extremes of 'can' and 'likely to'. In this paper I will attempt to introduce such a constraint, in order to generate a less ideal set of moral obligations than we get from the standard line which accepts nothing more than 'ought implies can'. I'm going to suggest that ought implies something like 'probably succeeding if you try', building upon a proposal of Geoffrey Brennan and Nicholas Southwood's (Brennan & Southwood 2007). We can call it 'non-ideal accessibility', as opposed to the rather more ideal accessibility discussed elsewhere in the literature (Feldman 1986; Buchanan 2004 61; Dennett 1995 Ch. 5; Hawthorn 1991; Cohen 2001, 2009 56-7). In the course of the paper I shall refine this suggestion and discuss its implications for familiar cases of moral failure.

This project has something in common with discussions had by Holly Goldman (1976), Howard Sobel (1976), and Frank Jackson with Robert Pargetter (1986), about moral imperfection. Goldman and Sobel (independently) propose that if an agent knows that he will fail to do *X*, where *X* is strictly available to him but nonetheless will not be done, then we should say he ought to do *Y*, the best of the options that he will actually do. I will agree with this to some extent, but argue that the agent's knowledge of why he won't is relevant to whether or not the fact that he won't gets him off the moral hook (at least with respect to that obligation, which doesn't mean he isn't immediately obliged to do the next-best thing). I return to this kind of case later in the paper via an example of Frank Jackson and Robert Pargetter's.

II. Non-ideal "oughts"

Consider the following claim: burning fossil fuels is one of the main contributors to climate change, so countries ought to leave all remaining oil in the ground. It's important that moral theory doesn't stop there, however, because chances are that countries won't leave all remaining oil in the ground. So now assume that our

---

4 That is not to suggest that these jobs exhaust moral and political philosophy. Moral philosophy should also tell us what a good person is like, and how persons ought to *be*; political philosophy should tell us what a good society is like, and how both domestic and international political culture ought to be. Both tell us about what is good or valuable, as well as what ought to be the case.

theory tells us that, given that countries won't leave the remaining oil in the ground, they ought to use clean energy technologies to convert the oil into permanent or recyclable materials, like plastics (rather than allowing it to be burned as fuel) (Schwarz 2003). Making that claim is just like saying that it would be better if the remaining oil were converted into permanent or recyclable materials, than if it were burned as fuel. That can be true even if it's also true that it would be best for the oil to be left in the ground (or even just better). Consider an alternative claim: Peter ought to give 10% of his salary to charity. But, if the most he will give to charity is 9%, then he ought to give 9%. But, if the most he will give is 8%, then he ought to give 8%... and so on.[5] It is better that Peter give 9% of his salary to charity than that he give 8%, even if it is better still that he give 10%.[6]

There are two ways to interpret such claims. The first is as follows. Given the option set containing A, B, and C, it ought to be that A (A is the best option); given the option set containing only B and C, it ought to be that B (B is the best option). Using restricted option sets to handle apparently conflicting obligations originated – I think – with Frank Jackson (1985), and was further developed in a collaboration between Jackson and Robert Pargetter (1986). The second interpretation is that the 'given that' claims are part of the full description of the better and worse ways the world could be. On that interpretation, the structure of the full description is a long conjunction, starting with the more ideal obligation statements and ending with the less ideal. What is true is 'it ought to be that A. And, if the best that will obtain is B, it ought to be that B. And, if the best that will obtain is C, it ought to be that C...' and so on. I am not interested in the first interpretation, on which we ought to do B rather than A only because we're looking at an option set that excludes A. It sidesteps the question of which option set is actually available to the agent. If the set containing A is available, then simply restricting our attention to the set excluding A leaves open the possibility that the agent ought, all things considered, to do A. I am interested in the second story, which features a long conjunction of successively less ideal obligation

---

5 In general, conditional obligation is not well understood. It cannot involve the material conditional, because the logic of that conditional makes both the "wide-scope" and the "narrow-scope" readings implausible. Perhaps conditional obligation is best understood as primitive; the issue needs further discussion. See discussion in (Lewis, 2000).

6 Some statements about moral obligation imply that 'closer is better', while others are silent on what ought to be the case given their failure to be realised. Percentage of income is linear; it seems clear that if you ought to give 10%, and yet you won't, the next best thing is 9% (assuming for simplicity that the next percentage is not an integer). But what if you ought to 'help the others rescue that drowning man'? This says nothing about what to do if the others refuse to cooperate, or you yourself are unwilling or unable to play your part.

statements. On that story, we can show that 'the best that will obtain is *B*', and from that conclude that 'it ought to be that *B*'. It ought to be that *B* precisely because it won't be that *A*, and at all the best not-*A* worlds, *B* occurs. The question then becomes, when are we justified in saying that it is *B* that ought to be the case, rather than *A*, even though *A* is strictly (ideally) accessible? Now is the time to elaborate upon the idea that instead of "can", ought implies "probably succeeding if you try" (or non-ideal accessibility).

III. "Ought" implies "probably succeeding if you try"
Once we have the full set of successively less ideal obligation statements, we then need a way of settling on the statement with non-ideal, genuinely action-guiding, normative content. The orthodox view is that the thing you all things considered ought to do is the best of the "can"-constrained options. My proposal is that we can be more realistic than that. We need to show that "the best that will obtain is *B*" is true, which licences inference to "it ought to be the case that *B*". This goes on until we reach an antecedent that is false, e.g. "the best that will obtain is *C*". There are two things to note here. One is that the statement "if the best that will obtain is *B*, then it ought to be the case that *B*" can be true regardless of why A won't be the case. That is because it is a statement that is part of the full description of what morality requires, what the better and worse worlds, relative to both the ideal world(s) and the actual world, look like. But, secondly, knowing the full description of what morality requires doesn't tell us what to do. We can't take all 'won'ts' as normatively relevant. We're not dealing in normativity at all if the upshot is that only things that will be the case ought to be the case. That is just a vindication of the status quo, but we know the status quo has been, and is, very much not the best we can do, even on the most tightly circumscribed of understandings of 'best'. So we need to figure out what kinds of 'won'ts' do some work in defining non-ideal accessibility. Allow me to introduce the following definition:

$$\text{NIA}x\,(S) \text{ iff } \max_A P\,(\,S \mid A\,) > z$$

Which is to say, a state of affairs (S) is non-ideally accessible (NIA) for an agent (x) if and only if the probability (P) of the state of affairs given[7] the agent's action[8] (A)

---

7 Standardly P (S | A) is defined as P (S & A) / P (A), but it might be better to read it 'subjunctively' - see discussion in (Joyce 2009 Ch. 5). It could also be read as the probability of the counterfactual conditional, namely, A $\square\!\rightarrow$ S.
8 More precisely: the proposition that the action occurs.

most likely to bring the state of affairs about (max_A) is greater than ($>$) a specified threshold (z).[9] Some of the elements require further explanation. NIAx (S) is a tensed proposition. It should be understood to be shorthand for 'S is NIA for x at time t'. Max_A gives S the best chance of being done. We get very strange results if we conditionalize on just any old action of the relevant agent. For example, compare the chance of my breaking a window, conditional upon picking my nose, with the chance of my breaking a window, conditional upon throwing a rock at it. Throwing a rock is causally efficacious in breaking a window in a way that picking my nose (normally – of course we can set up a special case in which that is the secret signal for someone else to throw a rock) is not. We conditionalize on the action most likely to produce S in order to obtain the most charitable answer to the question of what the chance of S is.

I use 'states of affairs' in the definition where it is more standard to talk about 'outcomes', and I use 'actions' to cover both mental and physical acts. Instead of the usual two-place relation between actions and outcomes, there is a three-place relation, between mental acts, physical acts, and outcomes. I take mental acts to include trying, choosing, willing, deciding, etc., and physical acts to extend to the standard range of 'doings'. When preparing for a game of cricket, you might ask what the chance of your swinging the bat at the ball, given that you try, is; you might also ask what the chance of your hitting a six is, given that you swing the bat at the ball.

The definition employs a conditional probability, the probability of S given A. I assume that the probabilities are objective–epistemic, determined in relation to the (best) available scientific data. The question of what those probabilities are comes apart from the question of whether, and to what extent, we can access that information. Because in general we, as individuals, face epistemic limitations in accessing that information (it is difficult enough to assess the probabilities of historical counterfactuals, let alone future-oriented counterfactuals), most of what I say in the paper will be about access, about the educated guesses we can make about objective conditional probabilities.

Finally, I introduced a threshold into the definition. This is to follow the ordinary language usage of words like 'can', or 'accessible', or 'feasible', in which a state of affairs is on / off.[10] There are two options when it comes to setting the

---

9 This definition of accessibility differs substantially from the account of feasibility that I have defended elsewhere, which refuses any threshold. See (Lawford-Smith forthcoming_b; Gilabert & Lawford-Smith 2012).

10 The alternative is to make it scalar, in which case instead of blocking a requirement (if it's not the case that S is non-ideally accessible, then it's not the case that x ought to A), it would *count*

threshold. One is to fix it in one place for all purposes, for example to say that S is non-ideally accessible if, conditional upon A, it has a probability greater than 0.5. Modulo the usual worries about line drawing (namely, if the probability of S_1 given A_1 is 0.5 and the probability of S_2 given A_2 is 0.49', why should we think that S_1 is accessible and S_2 not?) we accept that any S with conditional probability lower than 0.5 is non-ideally inaccessible, and therefore, in line with the new assumption that "ought implies non-ideal accessibility", not obligatory. The other option is to set the thresholds relative to the context. This is more plausible, because what the threshold for non-ideal accessibility should be really seems to depend on the case under consideration.

Compare two cases. In the first, a blackmailer calls me and tells me he has kidnapped my four closest friends and will torture and then kill them if I don't run the full 42km in tomorrow's city marathon (set aside the question of why this psychopath takes such a particular interest in my physical fitness...). In the second case, my colleagues agree to sponsor me in the city marathon for charity, together donating $1,000 if I run the full distance. Let's say the probability of my running the full distance, given that I try, is 0.4. If we had taken the option of setting the threshold for non-ideal accessibility at 0.5 across all cases, then my running the marathon would come out as non-ideally inaccessible in both the charity and the torture case. This option is attractive because it emphasises the importance of my physical capacity to what is non-ideally accessible. Surely what matters in establishing whether running a marathon is non-ideally accessible for me is my chance of running it, given a sincere effort to do so. But the alternative is even more attractive, because it gives us the means to say that running is non-ideally accessible in the torture case even if it is non-ideally *in*accessible in the charity case.

The torture case involves high stakes, while the charity case involves relatively low stakes. In the torture case, whether or not I run the full marathon will determine whether my four closest friends are tortured and killed, while in the charity case whether or not I run the full marathon will determine only whether

against a requirement in some to-be-specified way (e.g. the more non-ideally inaccessible S is, the less that x ought to A). But then accessibility ceases to play a negative role by restricting the set of actions / states of affairs that are candidates for what ought to be the case, and starts playing a positive role in what ought to be done / brought about. But that would require a much fuller discussion of how much accessibility should matter in deciding what to do. It can't be straightforwardly that the more accessible an option is the more it ought to be done, because that would show a strong bias toward the status quo, or states of affairs and actions that are easy to bring about / do. For a nearby discussion on scalar feasibility see (Gilabert & Lawford-Smith 2012, Lawford-Smith forthcoming).

$1,000 goes to charity (and let's assume that I could just donate $1,000 to charity myself, in lieu of running). Epistemologists, wanting to analyse knowledge as credence over a given threshold, set the threshold for knowledge higher in high stakes cases (e.g. when sceptical possibilities are salient, or when a lot depends on the outcome), and lower in low stakes cases. We can do the converse, setting the threshold for non-ideal accessibility low in high stakes cases, and high in low stakes cases. For example, we might set the threshold at 0.2 in the torture case, because the lives of my friends are at stake, and at 0.8 in the charity case, because it doesn't matter much if my colleagues fail to donate $1,000 (especially assuming that I could donate it in their stead). As above, I'm assuming that the chance of my running the marathon given a sincere effort is 0.4. If we set the threshold for non-ideal accessibility relative to context in this way, it comes out that running the marathon is non-ideally accessible in the torture case, and non-ideally inaccessible in the charity case. That entails that in the charity case, I am not under an obligation to run, while in the torture case, I am (see also discussion in Section VI).

One complication with this option is that in high stakes cases I might be inclined to try harder, to push myself to limits I would not have gone to in low stakes cases. Perhaps in a run where $1,000 for charity is at stake I would run until moderate exhaustion, while in a run where the lives of my four closes friends are at stake I would run until I literally collapsed and could not get myself up again. If more is possible in high stakes cases because high stakes force an expansion of the set of available actions, we could revert to the first option of fixing the threshold at the same point across all cases. The charity case would come out as non-ideally inaccessible, because the best action (i.e. trying with moderate effort) won't be very likely to result in my completing the marathon, while the torture case would come out as non-ideally accessible, because the best action (i.e. trying with extraordinary effort) will be likely to result in my completing the marathon. It does seem plausible that extreme circumstances expand the range of actions available to an agent. Nonetheless, there are many non-extreme cases where the A giving S the best chance of success is fixed across cases of differing normative significance, so that the conditional probability of S is also fixed. In those cases we need to define the threshold relative to the context.

Fred Feldman has suggested that we ought to do the best we can, which means, we are each morally obliged to bring about the intrinsically best of the possible futures accessible to us (Feldman 1986 Ch. 2). This sounds promising in terms of non-ideal morality, so long as we take the 'we' part seriously (incorporating the constraints that face each actual agent), because if a world is

not accessible to us, then we are not obliged to make it the case. But as it turns out, the view actually departs very little from orthodoxy. The possible futures accessible to an agent are those that she can "bring about" in some intuitive sense. If there is at least one accessible world in which she does *X*, then it can be true that she ought to do *X*. What happens in all the 'nearest' worlds is irrelevant. But this is a very near neighbour of the old view that 'ought implies can'. Surely what happens at all the nearest worlds *is* relevant, or at least, what happens at the nearest of a relevant subset of worlds. The proposal I put forward above was that a state of affairs is non-ideally accessible for an agent if the chance of that state of affairs given an action (mental, usually 'trying', or physical) is above some contextually-specified threshold. This suggests that the relevant worlds are those where the action is undertaken. Suppose we want to know whether a future in which I hit a six in the cricket game this afternoon is non-ideally accessible to me. We're asking about the likelihood of my hitting a six, given trying to (where 'trying' might involve concentrating, aiming the bat in the right way, swinging the bat at the ball with the appropriate amount of effort, and so on). Then we need to look at the unique closest world[11] in which I try to hit a six, and see whether I in fact hit a six at that world. For Feldman, what matters is only that there is at least one world among the accessible worlds in which I hit a six; on my view, what matters is whether the nearest world in which I try to hit a six has me in fact hitting a six. The two views come apart, because it might be that in most of the nearby worlds[12] where I try to hit a six, I fail, including at the nearest, even though there is one world, further out, in which I succeed. For Feldman it could thereby be obligatory for me to bring about that world, while on the view I proposed above, it could not be obligatory for me to bring about that world (the chance of S given A would fall below the threshold).[13]

IV. Agents' actions

There are two ways to be more realistic about what agents ought to do. One is to

---

11 Whether there is one unique closest world is controversial. See discussion in Feldman (1986 17).

12 I assume that we can make sense of an intuitive idea of 'most of the nearby worlds', even though strictly speaking possible worlds are infinite.

13 Here I am assuming that there is a unique closest world where A is the case, and that the objective probability of S given A at the actual world is equivalent to the objective probability that the whole counterfactual conditional 'if it were the case that A, it would be the case that S' is true at that unique closest A world. But this is controversial; in general it is an open question whether we can translate talk about probabilities into talk about closeness of worlds. David Lewis has proved that conditional probability is not equivalent to the probability of a conditional (Lewis 1976). Robbie Williams (ms.) has an interesting discussion about this result for counterfactual conditionals.

look at whether we are too optimistic about the kinds of actions we take to be available to agents (the other is to think about what causes S given A to have low probability, which I turn to in Sec. V). We cannot figure out the chance of S given A_1 if A_1 is not in the agent's option set. We conditionalize upon available actions, so we only ask about the chance of S given some A_2 actually in the agent's option set. And that chance might be much lower – low enough, for example, to make the action not obligatory. Circumscribing sets of available actions more narrowly is one way of being less idealistic about what agents can be obliged to do. I suspect that part of the reason we are so permissible about available actions is that we think about agents in abstract rather than concrete terms. When x is 'an agent' we think that very much more is possible for her than when x is a particular person (your father, your grandmother, your boss, for example). In this section I want to explore some of the features of persons that might make us more inclined to judge that certain actions are not in their option sets. I consider inductive evidence based on past actions, (more or less severe) pathology, personality, and beliefs.

It's not controversial that if you won't A_1, and I can only A_2 if you A_1, then it's not true that I ought to A_2. I ought to do as I do in the intrinsically best worlds accessible to me – what is accessible to me is different from, and can be constrained by, what is accessible to you (Feldman 1986 Ch. 2). But the received wisdom has it that the fact that you won't A_1 doesn't make it false that *you* ought to A_1. "Won't" is no barrier to "ought". But that is part of what is under investigation in this paper. In this section I want to agree with the common sense claim that when my action relies upon yours, I can't if you won't (and I ought only if I can), but disagree with the claim that you ought even if you won't. I want to suggest that if you know you won't A_1, it might not be true that you ought to A_1; knowing that you won't might in some cases be sufficient to trigger your obligation to A_2 instead.[14] The proposal is that exactly the information that gets me off the moral hook when my action depends on yours should get you off the moral hook too.

It is generally accepted that agents are in a privileged epistemic position with respect to their own actions. They know more about their own capacities, and they have more control over what they can do, can get themselves in a position to be able to do, and will do. But this creates an unusual situation with respect to prediction. It generally means that an agent cannot use useful information about himself that others have access to. For example, consider Frank

---

14 For the decision-theoretic discussion about taking beliefs about one's own future actions into account, see e.g. (Rabinowicz 2002; Spohn 1977; and Levi 1989).

Jackson's case of the procrastinating professor[15] (let's call him "David"). David is constantly being sent requests to referee papers, review books, examine students' theses, give papers at conferences, and contribute to edited collections. He generally fails to fulfil some proportion of those things he commits to. Now let's imagine that David has an administrative assistant, call him Adam, who notices that one thing David consistently fails to do is review books. When Adam receives an email asking David to review a new book, he has access to the information that David has never succeeded in completing a review in the past. He might reason that reviewing the book (or any book) is just not something that is accessible to the professor. The idea is that from observing David's behaviour over a period of time, and reasoning inductively from past to future cases, Adam has a better idea of what the professor can and can't (realistically) do, than the professor has himself. Adam can decide on the basis of this information to reject the request to review the book on the professor's behalf; but if the professor himself were to reject the review, citing past failures to complete, he'd have moral philosophers clamouring about his letting himself off the hook. Surely he *could* review it if he actually chose to!

I am inclined to think that the first- and third- person perspectives should be consistent with each other. The professor should have access to the same information his assistant has access to. People take information about their tendencies into account all the time. For example, I know that if there is candy within grabbing distance I will eat it. Predicting this about myself, I make sure not to have candy within grabbing distance whenever I can help it. But if I can take my predictions about my own behaviour into account in reasoning about what to do in some cases, why not all? Why can't David say "I never review books when I commit to doing so, so I shouldn't accept any more book reviews"? Adam can say "David never reviews books when he commits to doing so, so I shouldn't accept any more book reviews on his behalf", and it is odd that they shouldn't both be able to make use of that information. Thus a proposal: we go for consistency between first- and third- person predictions, and we use such predictions as a way of ruling certain actions out as live options for an agent, for the purposes of assessing non-ideal accessibility.

All sorts of complications arise here, including what counts as a sufficient evidential base for the prediction, and whether such predictions should be limited to a subset of agents' actions. With respect to the former, what if David had only

---

15 (Jackson 1985; Jackson & Pargetter 1986). Those papers are presumably inspired by the earlier discussion in (Goldman 1976).

ever been asked to review one book, and failed? With respect to the latter, isn't there a difference between a person who always litters because they can't be bothered finding a rubbish bin, on the one hand, and a person who often offends people because she is forthright in personality, or a person who often fails to keep her commitments because she suffers from anxiety, on the other hand? Let's say:

> S is non-ideally inaccessible to x if, every time x has tried in the past to bring about S by $_{max}A$-ing (for some intuitively sufficient number of attempted $_{max}A$-ings) she has failed.

Presumably a person who litters simply because they can't be bothered finding a rubbish bin hasn't tried not to litter. So this formulation should rule out cases of predicting mere bad habits. If the first- and the third- person predictions are to be consistent with one another, then third- persons must also think about whether the actions the relevant agent undertakes (fails to undertake) are done (not done) *in spite of* some effort not to do them (to do them). An agent can know herself well, sometimes better and sometimes worse than someone else who knows her well, but when either of them is interested in what is non-ideally accessible to her, their predictions should be not about what she will in fact do, but what actions are available to her in this more circumscribed way, and what the probabilities of various states of affairs conditional upon those actions are.

But predictions based on inductive evidence are not the only way that we can rule out more actions as being available to agents in the first place. We can also think of agents' option sets as constrained more closely by pathology, by personality, and by belief. If David's failure to complete book reviews is due to some phobia, compulsion, addiction, obsession or the like, then we shouldn't think that he has actions available to him (sitting down to review, beginning to read, making notes, for example) that would give the state of affairs of his having reviewed a book any positive probability. In Jackson's original case there is no pathology – the issue is procrastination, which is a tendency or habit rather than a hard psychological constraint (although it too can be more or less severe). It's generally uncontroversial among moral philosophers that pathologies block action, but pathologies are not 'on' or 'off', they come along a spectrum from extremely severe to very mild. One way of being less ideal about obligation would be to let more than just the extreme end of the spectrum rule out actions as being available to agents. For example, we could take anything from moderate to severe pathology as ruling out actions:

S is non-ideally inaccessible to x if, some moderate to severe pathology (phobia, compulsion, addiction, obsession) prevents x from A-ing, and S depends on A-ing.

What about personality? Imagine Blake, who is introverted, and awkward about social interactions in general. Someone tells her 'you ought to stand on street corners recruiting people to join and donate money to Amnesty International'. On the standard account of accessibility, Blake strictly speaking can stand on street corners recruiting people to Amnesty – we can't use 'cannot' to dissolve the obligation. But but because it difficult, given her personality, for her to do so, an account of non-ideal accessibility might prefer to deny that recruiting on street corners is something she can do. Blake herself, and others who know her well, would presumably deny that she could undertake that action.

This claim that personality can act as a constraint upon available actions is contingent upon further empirical research into the psychology of personality. How malleable is human personality? Are some traits fixed, or is everything open to change? How hard is change? Some research suggests that a large part of where a person sits in relation to the six major dimensions of temperament (Introversion-Extroversion, Neuroticism-Stability, Psychoticism-Socialization) is genetically determined (Eysenck & Eysenck 1975). But a weaker claim will suffice: if personality is not entirely malleable, then we may use it as a basis upon which to constrain agents' option sets for thinking about non-ideal accessibility. It is also worth noting that personality can be more or less fixed without being impossible to change. An action can be unavailable now because an agent's personality is a certain way, even though she could change her personality over time and cause the action to be available. Then we have:

S is non-ideally inaccessible to x if, given her personality, she would have to overcome a very significant amount of psychological resistance in order to A, and S depends on A-ing.

In thinking about both pathology and personality as grounds for narrowing option sets, we can characterize them both in terms of costs. The more serious the pathology, and the more fixed the personality (whether certain aspects are genetically predetermined, or whether it has become entrenched over time even if it were malleable at some earlier time), the higher the psychological costs of resisting it are, in which case an action requiring resistance would end up costing more than a non-ideal morality would ask.[16]

It might be impossible to bite this hard on a bullet without risking a dislocated jaw. To borrow an example, lets imagine that a psychopath, call her M., captures a victim and intends to torture him. And now we want to be able to say to her 'you ought not torture that man'. But I have just said that past failures, pathology, and personality can all be used as a basis for restricting option sets. It might then be that all the actions non-ideally accessible to M. at this time feature torturing as a component. In that case it will be false that she ought not torture that man, on the grounds that it violates the strong principle of non-ideal ought being restricted by non-ideal accessibility that we are working with. It's false that she ought not, because she *cannot* not. Do I have to accept that?

Firstly, it's not clear that torturing is a part of personality. It's also not clear for any given psychopath how much they have tried not to torture in the past and ended up torturing in spite of themselves. This leaves pathology. The cases I discussed above were negative, in that there was something an agent ought to have done but failed to do because a pathology prevented her. The torture case is positive, in that an agent does something she ought not do because a pathology compels her to. But that can't make a difference, because in more innocuous cases (such as with Obsessive Compulsive Disorder) it seems right that positive pathologies rule out certain actions. There are two things to say. The first is that it's probably not true that *all* the actions available to M. feature torturing as a component. It is presumably open to her to turn herself over to psychiatric care, where she will be forcibly prevented from torturing. In that case, she ought to do that. This is similar to my earlier claim that because I will eat any candy within grabbing distance I should make sure there are no treats around to eat – just more macabre. The second is that in the rare cases in which it really is true that the only accessible options feature torturing, it is *correct* that a non-ideal morality should recommend the best torturing action. Imagine we told M. 'you ought not torture that man, but if you do, you ought to use an anaesthetic'.[17] If it's a matter of fact that all the worlds non-ideally accessible to M. are worlds where the man is tortured, then the best world accessible from where she is will be those where an anaesthetic is used. This doesn't mean that torturing a person with anaesthetic ought ideally to be the case – of course it ought not. But it might mean that it ought non-ideally to be the case. Perhaps that choice takes us to the best world we

17 I borrow this example from (Louise 2009). Louise argues that both the oughts are relevant, because people genuinely find themselves caught between doing all that they ought to do, and failing in familiar ways and doing only what they ought to do in light of those failures (see esp. p. 346).

can get to from from an antecedently bad starting place. Morality cannot be about avoiding 'dirty hands' – sometimes bringing about the best of the accessible worlds requires us to do things that are not themselves morally good, or to implicate ourselves causally in harms.

Finally, agents' beliefs. We might want to say that an action is ideally accessible to an agent *whether or not she knows it*. But it seems a key feature of a non-ideal morality that it be sensitive to the agent's epistemic position. If she has been heavily indoctrinated into a particular religion or, more controversially, a particular society's positive morality, we might want to exclude certain actions from her option set – not actions that her religion or morality tell her are *wrong*, because people can choose to do what is wrong, but actions she does not believe are *possible*. Note that this restriction does not map onto the distinction between the subjective and the objective ought (Goldman 2010). Imagine that a small child has climbed over the safety fence of Matt's pool and is drowning. Perhaps we want to say that objectively, he ought to go out and rescue the child, but that subjectively, it's not true that he ought, because he doesn't know there is a child drowning. But ignorance about reasons for action is not grounds for saying that actions are not available to agents. Nothing makes it non-ideally inaccessible that Matt go out to the pool and rescue the child – no pathology (e.g. agora- or hydro-phobia), no aspect of his personality, and no beliefs that rescuing is impossible. Matt doesn't go outside because he doesn't have a reason to, but going outside is an action in his option set, whether ideal or non-ideal. Thus:

S is non-ideally inaccessible to x if, according to her beliefs A is impossible and A is the necessary means to S.

In this section I have explored the possibility of making obligation more practical by limiting agents' option sets. If we can say of any agent that some action is ruled out because she has failed at a sufficient number of past attempts, she has some moderate to severe pathology that prevents it, because it is in conflict with fixed aspects of her personality, or because she does not believe it possible, then we can say that that action is not available to her – not in her set of options. Then we cannot even *ask* what the probability of some state of affairs conditional upon that action is. (We can ask that only when there is an action). This is one way to strengthen what counts as non-ideally inaccessible, and what is therefore not non-ideally obligatory. In the next section I turn my attention to those actions that are accessible, even in this more circumscribed non-ideal sense, but which give an

agent's action low chance of producing the state of affairs in question.

V. From actions to states of affairs

In this section I want to concentrate on the chance of S given A where A is an action available in the more narrowly circumscribed way discussed in the last section, and in fact on a particular kind of A, namely the act of 'willing' or 'trying'. Insufficient attention has gone to the role of mental actions in determining physical actions for the purposes of the "can" in "ought implies can".

David Estlund (2011) talks about a case involving Bill, a man who consistently fails to deal with his rubbish in the required way. Bill, Estlund says, *wants to* deal with his rubbish correctly, but just can't will himself to do so. Despite his best efforts, Bill just drags bags of his trash out to the street and dumps them wherever he fancies. Not on the required day, not to the appropriate location, not in the council-appointed bags or the council-appointed bins. Estlund explicitly denies that there is any pathology involved in Bill's failure. His story is that Bill cannot will himself to deal with his rubbish correctly.

Let's try to run this case with the definition of non-ideal accessibility introduced in §III. The state of affairs we're interested in is Bill's willing himself to deal with his rubbish correctly (assume that if he succeeds in this, his willing will be efficacious in his acting). The action then must be his trying to will himself to deal with his rubbish correctly, the (antecedent) mental act of producing a (subsequent) mental act. What is the chance of Bill's willing, given that he tries to will? If it is lower than the contextually-defined accessibility threshold, then his willing is non-ideally inaccessible (and then it's not true that he ought to will himself to deal with his rubbish correctly). By stipulation, Bill cannot will himself to deal with his rubbish correctly. When he tries to will it, he will with a very high probability fail. Therefore, unless the threshold is set inordinately low (as it might be in a high-stakes case), Bill's dealing with his rubbish correctly will be inaccessible.

It's not clear to me that failure to will actually constitutes a special case distinct from weakness of will. Perhaps the former can be subsumed under the latter. For Richard Holton (2009), weakness of will is the over-ready revision of intentions formed explicitly in order to block future intentions. For example, take two people, Mickey and Tiki, both overweight. Imagine that their doctors tell them to lay off the junk food, for health reasons. Tiki has a fairly strong will; if he forms the intention now to eat no more junk food, then he will eat no more junk food. He forms an intention designed to block a future contrary intention, viz., to

eat junk food. His prior intention will be efficacious in blocking the later intention: that is what it means to have a strong will. Mickey, on the other hand, has a fairly weak will; he forms an intention now to eat no more junk food, but tomorrow, or later, when junk food is salient, he simply revises that intention, and forms the new intention to enjoy the junk food. The original intention was supposed to block the future contrary intention, but it did not do its job, because that's what it is to be weak-willed. ("Strong-willed" is not necessarily an accolade, incidentally: one can refuse to revise intentions that one really ought to revise, just as one can too readily revise intentions that one ought not to revise) (Holton 2009). Now imagine that we say to both Mickey and Tiki "you ought not eat any more junk food". Set the threshold for non-ideal accessibility arbitrarily at 0.5. Probably, it will be accessible for Tiki to do what he ought, and inaccessible for Mickey to do what he ought. It remains true that Tiki ought not eat any more junk food, but it becomes false that Mickey ought not. Particular states of affairs are less accessible to those with weaker wills. (Incidentally, this is one way to diagnose Jackson's procrastinating professor case. We can say that the professor has a weak will, so that no intention he forms today can make it the case that he writes the book review later. In light of that, his writing the book review is not really accessible to him. No action of his gives his writing the review a sufficiently high chance of occurring, so writing it will generally be inaccessible for him.[18] I say 'generally' because of course it is possible that his writing it involves extremely high stakes, e.g. if he doesn't write it then thousands of people will die, in which case we set the threshold so low that even the tiny chance of his current intention actually blocking the later intention not to write is higher than the threshold).

Notice that the point requires a dynamic decision sequence. There have to be at least two temporal nodes – in the above, there was the point at which the professor decided whether to accept the review, and the point at which, having accepted it, he should have started writing it. Moreover, the actions at each node must be non-identical, and more than trivially. The difference cannot be merely that between 'writing the review' and 'having written the review', for example. That is because when we conditionalize some action upon that action itself, we get probability 1. What is the chance of David writing the book review if he writes the book review? Obviously 1. If we take some action and plug it in as both S and A, the state of affairs and the action it is conditional upon, the state of affairs will always come out as maximally accessible. The special case shows that we have to be careful about what we take as 'action' and what we take as 'state of affairs' when

18 I am grateful to Michael Smith for discussion on this point.

assessing accessibility. What we plug in as an 'action' should be distinct from, and maximally causally relevant to, the state of affairs whose accessibility we are interested to assess (that is why I used max_A instead of just any A in the definition). What this shows is that weakness of will doesn't work as a constraint upon non-ideal accessibility in static decisions. An agent either has an action in her option set, or not (relevant to the constraints discussed in Sec. IV, as well as the usual constraints),[19] and if she does have, then the only question we can ask is what the chance of some state of affairs is, *given* that action. If it is higher than the contextually-defined threshold, then the state of affairs is accessible, whether the agent will actually decide to undertake that action or not.

Some might think that a non-ideal morality should make room for both weakness of will in acting upon the intentions one has formed and weakness of will in selecting certain actions that are strictly speaking non-ideally accessible. But remember the connection between non-ideal accessibility and blame. If an action is not in an agent's option set then she is not blameworthy for failing to bring about a state of affairs that depended upon such an action, nor is she blameworthy if the probability of her bringing about a state of affairs dependent upon an action that actually is in her option set is below the contextually-specified threshold. If we make weakness of will in *selecting* available actions a constraint, then agents won't be blameworthy even when an action is non-ideally accessible to her. But we generally think that weakness of will is not a good excuse when it comes to static decisions: it's not on a par, psychologically, with personality or pathology or belief. It plays a role in dynamic decision sequences only because the agent at the earlier node does not have full control over what his later self will do. This is false when we're only concerned with a single node. Part of what it means to be an agent is to have the ability to choose among one's accessible options which of them to actually do. Even on this more narrowly circumscribed account of that option set, it would be taking things too far to start permitting decisions not to act, even when these are disguised in the form of decisions to do *something else than that action instead*, as weakness of will in the form of procrastination often is.[20] So on the picture I have defended of restricted obligations, some "won't"-facts dissolve obligation, but not so many that we get all the way to "(non-ideal) ought implies will".

---

19 On the usual constraints, see discussion in Gilabert & Lawford-Smith (2012).
20 See Johanna Cordes-de Waal's (1996) discussion of omissions, and the correlation between omitting agents' blameworthiness when the omission is intended compared with when it is a non-decision (i.e. the time in which a decision could have been made has simply run out).

VI. Some worries

I want to deal with four main worries. The first is that the current proposal lets agents off the moral hook. The second is that the non-ideal accessibility threshold should not be normative in the way that it is. The third worry, related to the second, is that maintaining this view forces us into either cynicism or contradiction. Finally, I note that although controversial, the proposal might not be controversial enough, as it misses out an important class of non-ideal cases. I take these concerns in order.

If an action is wildly counter to my personality, then it may not count as an action of mine at all, and then no state of affairs that would require that action can be one I am obliged to bring about. But if that is true, it must be true that I can get myself off the moral hook by developing the kind of personality that is unable to perform many of its *prima facie* obligations. For example, imagine that Milo, a committed environmentalist, is extroverted and gregarious. By virtue of his character traits, Milo falls under an obligation to persuade others to be environmentalists. And let's say that for whatever reason, Milo dislikes this task. One way he can get himself off the hook (which is to say, rid himself of the obligation, rather than merely fail to satisfy it) is to make himself the kind of person who does not have 'recruiting others to a political cause' as an option. Milo steadfastly becomes less and less sociable, until he is such that others who know him would say that recruiting others to environmentalism is not something Milo could do. Now, setting aside worries about whether persons can really 'remake their personalities' in this way, should we really allow that Milo can make himself unable to fulfil his prima facie obligations? It seems clear that the answer to this question is 'yes'. There is a difference between saying 'Milo ought not to have made himself socially awkward', on the one hand, and saying 'Milo, that socially awkward guy, ought to go out and recruit people to environmentalism'. The former is clearly true, the latter clearly false. What persons ought to have done or are blameworthy for having done are importantly separate questions from what persons ought to do. Facts about how the world is, including facts about how people are, non-ideal as they may be, are what determine the set of practical obligations.

The second problem with the proposal defended here is that it tramples over efforts to maintain a strict separation between what can be done, on the one hand, and what ought to be done, on the other (Cohen 2001, 2009; Gilabert & Lawford-Smith 2012; cf. Räikkä 1998; Gilabert forthcoming). My proposal appears to make accessibility dependent on what we happen to care about. We

happen to care a lot about torture, so we set the stakes high in the torture case; we don't care too much about charity (especially if I can pay the equivalent of whatever would have been donated on my behalf), so we set the stakes low in the charity case. Normative considerations are necessary to determine whether the stakes are low or high, and so where to set the threshold. But accessibility is meant to be a consideration on the empirical side of things, telling us what can with some particular probability be done. Setting the threshold contextually, relative to what we care about, also risks (albeit non-vicious) circularity: to determine what is not obligatory we have to know what is accessible, but to determine what is accessible, we have to know what we care about (which some might say amounts to roughly the same thing as what is obligatory). We can answer this worry in the same way the epistemologists answer an analogous challenge to contextualism about knowledge. We do not gain knowledge merely by ignoring sceptical possibilities (or other high-stakes -makers), just as we do not gain accessibility merely by caring more about what is at stake. Rather, we can be interested in different strengths of knowledge, and accessibility. In ordinary cases we are interested in a less demanding kind of knowledge, in high-stakes cases in a more demanding kind. We rarely have the demanding kind of knowledge necessary to defeat sceptical possibilities, but fortunately, we rarely need it. Likewise in ordinary cases we are interested in a less demanding kind of accessibility, in high-stakes cases a more demanding kind. When a lot is at stake, we're not interested in the less demanding kind of accessibility, rather we want to know whether there is some chance, perhaps any chance, of an agent succeeding in producing a state of affairs. When less is at stake, we will be interested in the less demanding kind of accessibility, as we are in the charity case. We want to know whether there's some reasonable chance of the agent's success. This response does assume a further premise, something like "when a lot is at stake, we should be interested in even actions with a low chance of success". That is an entirely plausible assumption. Normally, we wouldn't think that the world where the agent escapes a burning building alive by jumping from the third floor is accessible to her, but if the building is on fire and jumping is her best chance of survival, we might.[21]

For those who are not fully satisfied by the epistemologists' answer, an

---

21 I borrow this example from Pablo Gilabert. The other way to go, I suppose, is to say that escaping the office building is inaccessible in the low-stakes case and *remains* inaccessible even in the high-stakes case, but that *sometimes we should do inaccessible things*. Because I use accessibility as a way to rule-out obligations, I prefer not to go this way. If the actions are things we should do, then they're things that it is accessible for us to do (because of the assumption that "ought implies non-ideal accessibility"). They can't be *both* inaccessible and something we ought to do.

alternative is to distinguish the kind of normativity relevant to setting the threshold.[22] That is a response to the third worry, that the view defended here pushes us into either cynicism or contradiction. The worry is this: "Blake ought not torture that man" and "Blake ought to use anaesthetic when she tortures that man" cannot both be true. If only the latter is true, we are pushed into cynicism – can't we aspire to better than that? Even if Blake will torture that man, surely the point of moral theory is to be able to say that she ought not. Moral theory should reflect our values and standards, not make hopeless concessions to the extremes of human behaviour. But if both are true we have a contradiction, because the obligation claims recommend distinct and exclusive courses of action. We can resolve this worry by distinguishing between different kinds of 'oughts'. This is not a novel proposal; Fred Feldman (1986 Ch. 2) for example distinguishes moral, prudential, social and civic oughts, and John Broome for example distinguishes a rational, moral, prudential, conventional, and all things considered ought (although in the end he chooses to refer to the first four as 'sources of requirement', and the latter as the 'central normative concept') (Broome forthcoming). But here the distinction should be more fine-grained, between different kinds of moral oughts rather than between different sources of normativity in general.[23]

There is no circularity involved if the moral information relevant to setting the threshold is of a different kind to the moral information relevant to determining what we ought to do (the latter which generates prima facie obligations, which are then constrained by accessibility). One candidate distinction is between 'the right' and 'the good'. Another is that between 'ought-ideally' and 'ought'. Or we can give up 'ought' for the ideal claims, using something like 'would be good if' in their place and reserving 'ought' for the non-ideal and immediately action-guiding. How does this disambiguating of 'oughts' relate non-ideal morality to blameworthiness? Normally, if a person fails to do as she ought, she is blameworthy. The answer depends on whether we think blame should be indexed to ideal or to non-ideal morality. As far as I can see, there would be little point in

---

22 A further alternative is to take the normativity out of the definition of non-ideal accessibility entirely, and put it into the relevant principle. So we could shift to a non-binary definition, e.g. $NIAx(S) = max_A P(S \mid A)$ (the non-ideal accessibility of a state of affairs for an agent is equal to the probability of the state of affairs given that action of the agent's which gives the state of affairs the best chance of success), and modify the relevant principle from "non-ideal ought implies non-ideal accessibility" to "non-ideal ought implies accessibility greater than contextually-defined threshold z".

23 Lennart Åqvist (1991 219) discusses moral oughts indexed to different times, which is one way of their being ideal to differing degrees (as time passes, the space of possibilities narrows).

advancing the case of a non-ideal morality only to have real non-ideal agents lumped with blame for failing to do what they *ideally* ought to have done. What a person ought to do and what she is blameworthy for not doing must go hand in hand. Thus I suggest that we attach blame to non-ideal morality. That is a reason for going with the last suggested disambiguation. No blame attaches to the 'would be good if' locution. If we interpret ideal obligation statements as evaluative, as expressions of standards and values, and non-ideal obligation statements as normative, as bearing upon agents' behaviour directly, then agents are only blameworthy when they fail to do what they non-ideally ought to do. I am sure there will be hard line traditionalists who will want to resist this claim, saying that how weak a person's will is, what kind of personality they have, what phobias and addictions and compulsions they struggle with, and what they have been conditioned to believe, are all irrelevant to what a person ought to do. There is little more that I can say to such people; I can only hope that they will see the merit of at least *thinking about* what a less ideal morality might be like.

Quite aside from the kinds of normativity involved in making the ideal claim and setting the threshold for the non-ideal claim, there is potential contradiction between each and every one of the successively less ideal claims set out in the comprehensive picture of morality outlined early in the paper. We want to say 'Peter ought to give 10% of his salary to charity. But, if he won't give 10%, then he ought to give 9%. But, if he won't give 9%, then he ought to give 8%...' and so on. But if the antecedents are true, then the consequents are true, in which case the claims 'Peter ought to give 10%', '...he ought to give 9%', and '...he ought to give 8%', contradict one another. If he ought to give 10% then it's not true that he ought to give (only) 9%. A solution to the threat of contradiction is to give multiple disambiguations of non-ideal oughts, relative to particular sets of constraints. Formally, each 'ought' will come with a subscript that describes the relevant constraints, e.g. 'Peter ought$_{\{C:\ a,\ b,\ c\}}$ to give 10%', '...he ought$_{\{C:\ a,\ b,\ c,\ d,\ e\}}$ to give 9%'... and so on. Then the interesting question becomes which is the right set of constraints relative to our practical and theoretical interests (but notice how close this is to Jackson & Pargetter's claim that oughts are relative to different option sets). I tried to formulate the comprehensive picture earlier in the paper using antecedents that will turn out to be false in non-ideal conditions rather than true, to avoid the consequents being true. I said 'Peter ought to give 10% of his salary to charity. But, if the most he will give to charity is 9%, then he ought to give 9%. But, if the most he will give is 8%, then he ought to give 8%...' and so on. In that case, if Peter will give 9%, then it is false that the most he will give is

8%, and contradiction is avoided. It remains to be seen whether we can formulate the antecedents in that way in all cases.

The final worry I have is that the proposal misses out an important class of non-ideal cases. It allows us to rule actions out of agents' option sets when the agents have a pathology towards the severe end of the spectrum preventing that action, when the actions are against fixed aspects of the agent's personality (pending further empirical investigation into personality), when the agent has a history of failing to complete the action when she tries. And it allows us to assign a state of affairs a low chance of success conditional upon actions when the agent has a weak will. What it doesn't allow us to rule out is those actions an agent decides not to do, as mentioned already. In many cases, this is an obvious advantage – after all, we don't want to end up at "ought implies will", which would be normatively uninteresting and philosophically useless. But it seems that there are many non-ideal cases where the non-ideal conditions do rest on decisions. For example, China and the United States should cooperate with other states in cutting carbon emissions. At the Copenhagen meeting, they decided not to. That means there's nothing we can say about what they ought to do instead – we're stuck at saying they ought to cooperate, even though they won't. (We might, for example, want to say "given that they won't cooperate, they ought at the very least not use other countries' emissions cuts as credits".) Maybe it's right to say that they ought to cooperate, because they ideally ought to cooperate and the world where they do cooperate is non-ideally accessible. But perhaps we want to be able to assign non-ideal obligations in decision-based cases too, and if so, the approach outlined here will need to be augmented.

VII. Conclusion

In this paper I have tried to give the constraints for determining a set of practical obligations, differing from the orthodox view which accepts "ought implies can" but nothing stronger. I introduced the stronger constraint that "non-ideal ought implies non-ideal accessibility", and made non-ideal accessibility dependent upon a contextually-defined threshold. In high-stakes cases more is accessible because we're interested in a weaker kind of accessibility; in low-stakes cases less is accessible because we're interested in a stronger kind of accessibility. I suggested that past behaviour, severe pathology, and character are relevant to circumscribing a narrower set of agents' actions for the purposes of assessing accessibility, and I suggested that weakness of will is relevant to making the conditional probability of

states of affairs lower. I think that the general structure of obligation statements can be captured in a long conjunction beginning with the more ideal and ending at the less ideal. The interesting project for moral philosophers interested in non-ideal obligation is to say which antecedents are false (because inaccessible), and so to arrive at the obligation which actually bears in the non-ideal conditions we find ourselves in.

Holly Lawford-Smith
University of Sheffield

References
Åqvist L (1991) Review Essay: Doing The Best We Can. *Philosophy and Phenomenological Research* 51: 215-225.

Brennan G, Southwood N (2007) Feasibility in Action and Attitude. In: Rønnow-Rasmussen T, Petersson, B, Josefsson J (eds) *Hommage à Wlodek. Philosophical Papers Dedicated to Wlodek Rabinowicz.* Online at www.fil.lu.se/hommageawlodek.

Broome J (manuscript) *How to be Rational*. Oxford.

Buchanan A (2004) *Justice, Legitimacy, and Self-Determination.* Oxford, Oxford University Press.

Cohen G (2009) *Why Not Socialism?* Princeton, Princeton University Press.

Cohen G (2001) Why Not Socialism? In: Broadbent E (ed) *Democratic Equality. What Went Wrong?* Toronto, University of Toronto Press, pp. 58-78.

Collins S (forthcoming) Collectives' Duties and Collectivisation Duties. *Australasian Journal of Philosophy*.

Cordes-de Waal J (1996) Intention and the Omission Bias: Omissions Perceived as Nondecisions. *Acta Psychologica* 93:161–72.

Dennett D (1995) *Darwin's Dangerous Idea*. New York, Touchstone.

Estlund D (2011) Human Nature and the Limits (If Any) of Political Philosophy. *Philosophy and Public Affairs* 39:207-237.

Eysenck HJ, Eysenck S (1975) *Manual of the Eysenck Personality Questionnaire*. London, Hodder & Stoughton.

Feldman F (1986) *Doing The Best We Can*. D. Reidel Pub. Co., Dordrecht.

Gilabert P (forthcoming) Feasibility and Socialism. *The Journal of Political Philosophy*.

Gilabert P, Lawford-Smith H (2012) Political Feasibility: A Conceptual Exploration. *Political Studies* [early view].

Goldman H (1976) Dated Rightness and Moral Imperfection. *The Philosophical Review* 85:449-487.

Goldman H (2010) Subjective Rightness. *Social Philosophy and Policy* 27:64-110.

Goodin R (2009) Demandingness as a Virtue. *Journal of Ethics* 13:1-13.

Graham P (2011) 'Ought' and Ability. *Philosophical Review* 120:337-382.

Hawthorn G (1991) *Plausible Worlds*. Cambridge, Cambridge University Press.
Holton R (2009) *Willing, Wanting, Waiting* Oxford, Oxford University Press.

Jackson F (1985) On the Semantics and Logic of Obligation. *Mind* 94:177-195.

Jackson F, Pargetter R (1986) Oughts, Options, and Actualism. *The Philosophical Review* XCV:233-255.

Joyce J (2009) *The Foundations of Causal Decision Theory*. Cambridge, Cambridge University Press.

Lawford-Smith H (2010) Ideal Theory: A Reply to Valentini. *Journal of Political Philosophy* 18:357-368.

Lawford-Smith H (forthcoming_a) The Feasibility of Collectives' Action. *Australasian Journal of Philosophy*.

Lawford-Smith H (forthcoming_b) Understanding Political Feasibility. *Journal of Political Philosophy.*

Levi I (1989) Rationality, Prediction, and Autonomous Choice. *Canadian Journal of Philosophy* 19:339–363.

Lewis D (2000) Semantic Analyses for Dyadic Deontic Logic. *Papers in Ethics and Social Philosophy.* Cambridge, Cambridge University Press, pp. 5-19.

Lewis, David (1976) Probabilities of Conditionals and Conditional Probabilities. *Philosophical Review* 85:297-315.

Louise J (2009) I won't do it! Self-prediction, moral obligation and moral deliberation. *Philosophical Studies* 146:327-348.

Rabinowicz W (2002) Does Practical Deliberation Crowd Out Self-Prediction? *Erkenntnis* 57:91-122.

24

Räikkä J (1998) The Feasibility Condition in Political Theory. *The Journal of Political Philosophy* 6: 27-40.

Robeyns I, Swift A (eds) (2008) Social Justice: Ideal Theory, Nonideal Circumstances: Special Issue. *Social Theory and Practice* 34.

Schwarz W (2003) The Plastic Plan. *Wo's Weblog.* http://www.umsu.de/wo/2003/104#c1758 Accessed 16th January 2012.

Simmons J (2010) Ideal and Non-Ideal Theory. *Philosophy & Public Affairs* 38:5-36.

Sobel H (1976) Utilitarianism and Past and Future Mistakes. *Noûs* 10:195-219.

Spohn W (1977) Where Luce and Krantz Do Really Generalize Savage's Decision Model. *Erkenntnis* 11:113–134.

Vranas P (2007) I Ought Therefore I Can. *Philosophical Studies* 136:167-216.

Williams J, Robert G (manuscript) Counterfactual Triviality: A Lewis-impossibility argument for counterfactuals.