STAFF SUMMARY SHEET

	TO	ACTION	SIGNATURE (Surname), GRADE AND DATE		TO	ACTION	SIGNATURE (Surname), GRADE AND DATE
1	DFPY	sig	(sig) Cook, O-6, 12Feb12 (print)	6			
2	DFER	approve	for CIV 22 Feb 12	7			
3	DFPY	action		8			
4				9			
5				10			

SURNAME OF ACTION OFFICER AND GRADE	SYMBOL	PHONE	TYPIST'S INITIALS	SUSPENSE DATE
Dr. Leonard Kahn, AD-23	DFPY	333-8639	LK	Not Applicable

SUBJECT: Clearance for Material for Public Release — USAFA-DF-PA-52 — DATE: 20120213

SUMMARY

1. PURPOSE. To provide security and policy review on the document at Tab 1 prior to release to the public.

2. BACKGROUND.
Authors: Authors: Leonard Kahn (sole author), Associate Professor of Philosophy, United States Air Force Academy

Title: "Rule Consequentialism and Scope"

Description: I sketch the distinction between absolute and relative rule consequentialism and then argue that the latter is preferable to the former.

Release Information: This paper to be submitted to the journal ETHICAL THEORY AND MORAL PRACTICE

Recommended Distribution Statement:
(Distribution A, Approved for public release, distribution unlimited.)

3. DISCUSSION.

4. VIEWS OF OTHERS.

5. RECOMMENDATION.

Title Page

Author: Leonard Kahn

Author's Affiliation: United States Air Force Academy, 2354 Fairchild, Drive, USAF Academy, Colorado 80840-6256, USA, Voice: (719) 338-6791, FAX: (719) 333-2050, Email: leonardkahn@gmail.com

Paper's Title: "Rule Consequentialism and Scope"

Paper's Abstract: Rule consequentialism (RC) holds that the rightness and wrongness of actions is determined by an ideal moral code, i.e., the set of rules whose internalization would have the best consequences. But just how many moral codes are there supposed to be? Absolute RC holds that there is a single morally ideal code for everyone, while Relative RC holds that there are different codes for different groups or individuals. I argue that Relative RC better meets the test of reflective equilibrium than Absolute RC. In particular, I contend that Relative RC is superior because it accommodates our convictions about costless benefits. Some have charged that Relative RC threatens our convictions about the generality of moral codes and that it leads inevitably to what Brad Hooker calls "runaway relativism." I argue that Relative RC has principled reasons for stopping this imagined slide down the slippery slope.

Keywords: Ethics, Consequentialism, Rule, Absolutism, Relativism, Reflective Equilibrium, Parfit, Hooker, Brandt, Singer

Rule Consequentialism and Scope

1. Rule Consequentialism – Absolute and Relative

To a first approximation, rule consequentialism (RC) is the view that it is morally wrong for an agent to do an action if and only if that action violates the ideal moral code, where the ideal moral code (IMC) is the set of rules whose internalization would have the best consequences.[1] But is there only one IMC? Or are there many? And if there are many, how many are there, and to whom does each IMC apply? In short, how is the scope of the IMC best understood within RC? The purpose of this paper is to take a step toward answering these questions.

There are two general lines of response to such questions. In the rest of this section I articulate and clarify both. Let me begin with a brief statement of each line of response:

> Absolute RC: There is a single ideal moral code, and all agents are required to act in accordance with it.
>
> Relative RC: There is more than one ideal moral code, and some agents are required to act in accordance with one, some with another.

It is important to see that Relative RC is not simply the denial of Absolute RC, for the denial of Absolute RC is consistent with there being no IMC at all. But I cannot, and do not, entertain that skeptical possibility here. My concern in this paper is only with the comparative

1 Hooker (2000a 32). Here and throughout, I make no distinction here between consequentialism and utilitarianism. Of course, such a distinction can be made – and in a wide variety of ways. However, these distinctions do not have enough relevance for the purposes of this paper to warrant more than passing mention.

plausibility of these theories, i.e., with the question of whether Absolute or Relative RC is more plausible.

Absolute RC is fairly easy to characterize. As Brad Hooker puts it, there is a single "collective, shared code," which holds for "everyone everywhere" (2000a, p. 32).[2] Despite occasional remarks about making "room for some form of relativity" (2000a, p. 189), his official position appears to be that a "moral code, or at least the most basic moral code should be internalized and followed by everyone, not just by you or by me or by any mere sub-group of the whole" (2000a, p.1).

In contrast to Absolute RC, Relative RC is ambiguous; as such, it is subject to three distinct interpretations in terms of its scope. A good place to begin getting a handle on the first of these is the work of Richard Brandt. According to Brandt, an IMC is a code "which all fully rational agents would support, in preference to all others or to none at all, *for the society of the agent,* if they expected to spend a lifetime in that society" (1979, p. 194, italics added). I certainly concede that there seems to be something natural about thinking of IMCs in terms of societies as Brandt does. However, at least at the outset it seems somewhat arbitrary to limit Relative RC to societies rather than other types of collectivities. The following questions compete for attention: Why should this be the only way to formulate Relative RC? Could there not be versions of the theory that center on moral codes of cultures, of nation-states, of clans, or, for that matter, of senior common rooms? These questions are reasonable, but they should not drive us to formulate countless types of Relative RC: Society Relative RC, Nation-State Relative RC, etc. That way madness lies since we will quickly have more theories than we can possibly evaluate. I propose instead that we use the term "group" to refer to any collective entity of this sort. The term "group" is broad enough, at least for my

2 For another version of Absolute RC, see Parfit (2011). Brandt (1996, pp. 140-141) and Mulgan (2001, pp. 54, 62, and 83; 2006, pp. 255-260) formulate, but do not accept, Absolute RC.

current aims, to encompass sets of agents which are much larger than single societies as well as groups which are much smaller. Hence, we get the first of three interpretations of Relative RC:

> Weak RC: For each relevant group there is an IMC, where G is a group with regard to some set S of agents if and only if G is composed of two or more agents who form a *proper* subset of S. (a.k.a, the Weak Thesis)

On this understanding, Brandt's version of the Weak Thesis which focuses on IMCs for each society is just one of many sub-varieties of the theory. At this stage of the game, I do not offer an argument for preferring any one of these sub-varieties, though I do discuss some in Section 5. For now, I need to continue to engage in taxonomy in order to lay the groundwork for doing so. At any rate, what is of foremost importance for my present purposes is this: According to the Weak Thesis, every sub-variety of RC determines the IMC for each group by comparing the value of a single code which is to be internalized by agents within that group. So, e.g., Archie and Beatrice are in the same group if and only if they will have the same IMC. In other words, the Weak Thesis about RC relativizes IMCs to certain groups. In Sections 3 and 4, I limit myself to showing that as long as "group" is understood as denoting any proper subset of the set of all agents, i.e., something with a smaller domain than that composed of all agents, Relative RC is preferable to Absolute RC. Since doing so shows that Hooker's version of RC needs to be modified, this is not a trivial or an unimportant task; quite the contrary.

Let me turn now to the second interpretation of Relative RC. Though as we have seen Brandt prefers a version of the Weak Thesis, he entertains the possibility of an alternative as

a legitimate fall back position if this theory turns out to be untenable (1979 p. 188)[3]. Call this:

> Strong Relative RC: For each agent there is exactly one IMC (a.k.a., the Strong Thesis)

The Strong Thesis differs from the Weak Thesis in holding that the ideal moral code is indexed to individuals, rather than groups. According to the Strong Thesis, Relative RC determines which moral code is ideal for *each* moral agent by comparing the values of a variety of codes which is to be internalized by that moral agent. Simplifying somewhat, if a rule against lying is part of Archie's morally ideal code but not part of Beatrice's, then it is wrong for Archie to lie but not for Beatrice to do so. If two agents have similar IMCs, then that fact is an accident as far as the theory is concerned. The Strong Thesis as such does not require such similarity across agents, though it does not strictly prohibit it either.

The third and final interpretation of Relative RC is a compromise between the Weak Thesis and the Strong Thesis:

> Moderate Relative RC: Some morally ideal codes hold for groups of agents while others hold for a single agent. (a.k.a., the Moderate Thesis)

At first glance the Moderate Thesis seems somewhat unprincipled and perhaps downright wishy-washy, but it has some virtues, as will become clear later in this paper.

Before continuing, it is worth pausing for a moment in order to distinguish the object of my concern in this paper from another with which it is easily confused. Recall from above

3 Also see Brandt (1996, pp. 141-142).

Hooker's claim that an IMC is a set of rules "whose internalization by the overwhelming majority of everyone everywhere in each new generation has maximum expected value." In this passage, Hooker commits himself to two claims which are, I think, conceptually distinct. One of these claims concerns the size and nature of the set of individuals to whom an IMC is meant to apply, and *that* is the question in which I am interested here.[4] Absolute RC is one response to this; Relative RC is another. The other claim to which Hooker commits himself concerns whether we are to evaluate the IMC as if it were internalized by all members of the set or whether we are to allow some slack and assume that it is internalized by, as Hooker puts it, "the overwhelming majority," understood as 90% of the population. The latter question has already been considered at length elsewhere, and I have nothing new to say about it in this paper.[5] For the sake of simplicity, I assume throughout that all members of the given set internalize the relevant IMC. However, everything I say here holds, *mutatis mutandis*, even if the overwhelming majority but not some minority, rather than everyone, internalizes the IMC. A final point: I understand "X internalizes code C" to imply both "X accepts C" *and* "X complies with C." Acceptance and compliance come apart in unusual and philosophically interesting cases, but I am not interested in such outliers here.[6]

---

4 The question of determining the scope of a given moral code is hardly unique to RC. It is a question that must be answered, e.g., by any form of social contract theory, whether it is a general account of morality (see Gauthier, 1986 and Scanlon, 1998) or a more specific account of some part of it, say, justice (see Rawls, 1971 and 1993). Work by Nussbaum (2005) and others have done much to clarify how deeply problematic it is for contractualists to deal with scope.

5 See especially Brandt ([1965] 1992), Lyons (1965, pp. 138-142), Hooker (2000a, pp. 80-85), Arneson (2005), Mulgan (2006, pp. 136-137), Ridge (2006; 2009), Hooker and Fletcher (2008), Miller (2009), and Ross (2009) as well as Kahn (Unpublished).

6 Contra Hooker (2000a, p. 76), Mulgan (2001, p. 62), and Rosen (2009), I do not think that we are forced to choose between either acceptance of or compliance with these rules. Both are aspects of internalization in normal moral agents. Note that sense this understanding of *internalization* includes both acceptance and compliance, it does not open RC to the so-called "collapse objection." See Mulgan (2006, pp. 146-150). Moreover, even if one wished to take a hard line and insist that "internalization" is not plastic enough to encompass compliance, an advocate of RC could circumvent the issue simply by formulating her theory in such a way that the morally ideal code is the one that would lead to the best consequences if it were *both* internalized *and* complied with.

2. Evaluating Ethical Theories

Advocates of RC offer a wide variety of justifications for their theory. Some appeal to basic consequentialist goals and argue that accepting RC would be the most effective means of attaining them.[7] Others assert that possible grounds for RC can be found in what we would desire if we were fully informed, what we would agree to in terms of a social contract, what our society would be most rational in accepting, or even what is required because of our status as ends-in-themselves.[8] But in this paper I want to direct attention to another possibility, one to which I am quite familiar – namely, Hooker's claim that RC "does a better job than its rivals at tying together our moral convictions" (1994, p. 29), i.e., that it best meets the test of reflective equilibrium.[9] Since there are disagreements about precisely how to understand this test, let me clarify what I have in mind. On this approach to evaluating ethical theories, we begin by formulating what I here call *reflectively endorsed beliefs*, i.e., "attractive general beliefs about morality" or "moral convictions" which "we have [even] after careful reflection" (Hooker, 2000a, p. 4). Though one's reflectively endorsed beliefs "include judgments at all levels of generality" (Scanlon, 2002, p. 150),[10] I pay especially close attention to fairly broad beliefs in this paper. Next, we pair these reflectively endorsed beliefs with particular ethical theories and determine which theory best explains and justifies them "from an impartial point of view" (Hooker, 2000a, p. 4). In doing so, we also ask ourselves which

---

7 For discussion, see Brandt (1965 [1992], p. 117), Parfit (1984, pp. 30-31), Shaw (1999, pp. 164-167; 2001, p. 1074), Hooker (2000b, pp. 223-225), Kagan (2000, p. 134), Riley (2000, pp. 40-45), and Mulgan (2001, pp. 55-56). Of course, we must distinguish between RC and forms of act consequentialism that advocate the adoption of "rules of thumb." See, e.g., Hare (1981), Mulgan (2001, p. 64), and Crisp (2006),

8 For discussion, see Harsanyi (1977; 1985), Brandt (1979), Copp (1996, pp. 165-166), Cummisky (1996, pp. 62-64), Kagan (1998, pp. 198-199), Mulgan (2001, pp. 57-59; 2006, pp. 131-132), Gibbard (2008, pp. 59-82), Miller (2009), and Parfit (2011).

9 On this point, see Hooker (1996, pp. 543-546; 2000a, pp. 4-31; 2000b, pp. 101 and 222-238), Miller (2000, p. 156), Thomas (2000, p.181), Montague (2000, p. 205), and Driver (2002). I differ from Mulgan (2001, p. 59) in thinking of reflective equilibrium as a "practical justification" for RC, but the details of our disagreement are not important here.

10 See also Hooker (2000a, p. 88).

theories best "help us deal with moral questions about which we are not confident" (Hooker, 2000a, p. 4). Of course, this is a dynamic and on-going project, but, ultimately, the correct ethical theory is the theory that best fits with these reflectively endorsed beliefs. To repeat, there are disagreements about how best to understand reflective equilibrium, but this is not the place to settle them.

It should be noted, if only in passing, that the test of reflective equilibrium is not limited to ethical theories. In one guise or another, it has also been applied to theories of logic (e.g., by Nelson Goodman, [1955] 1983), political theories (e.g., by John Rawls, 1971), and scientific theories (e.g., by Richard Boyd, 1988); and there is no reason it could not be applied to other theories as well. Yet in each of these cases, there are important differences. One of these differences concerns the content of the reflectively endorsed beliefs when using reflective equilibrium to evaluate, for instance, theories of logic rather than ethical theories. The reflectively endorsed beliefs that are relevant to evaluating ethical theories include beliefs about keeping promises and about special obligations to one's family and friends. These reflectively endorsed beliefs – however attractive they might be – do not have any probative force when evaluating theories of logic. Another concerns the goals of the kind of theory being evaluated. All other things being equal, a scientific theory is the worse for failing to predict observable phenomena; an ethical theory is not. But the aim of finding plausible equipoise between one's reflectively endorsed beliefs and a theory which explains and justifies them remains the same in all applications of reflective equilibrium.[11]

We can further clarify the role of reflective equilibrium in evaluating ethical theories by briefly considering recent work on the subject by Katarzyna de Lazari-Radek and Peter Singer. Who, de Lazari-Radek and Singer ask, are we supposed to think of as having the

11 Note also that in some cases it is worthwhile to distinguish between narrow and wide reflective equilibrium, but not here. See Daniels (1996, pp. 21-46), Miller (2000), and Hooker (2000b).

reflectively endorsed beliefs which form the basis reflective equilibrium? They consider two answers to this question, and on the basis of these answers they present a dilemma for those who think that reflective equilibrium favors RC (though their challenge, if valid, would generalize to *any* use of reflective equilibrium to evaluate ethical theories). The first horn is that reflectively endorsed beliefs are held by members of what they call the "general public" (2010, p. 46). Unsurprisingly, de Lazari-Radek and Singer are not sanguine about the credibility of this scenario; they write,

> Encouraging members of the general public to reflect carefully on their moral convictions is all very well, but it is unlikely to diminish to any significant extent the influence of a variety of factors that are irrelevant to the soundness of a moral theory, including false religious beliefs, cultural and ethnic prejudices, and the innate predispositions that are a legacy of the process of millions of years of evolutionary selection. (2010, p. 46)[12]

In short, if the reflectively endorsed beliefs are those of the general public, then they have no special place in the evaluation of ethical theories, as the method of reflective equilibrium presupposes.

Consider now the second horn of de Lazari-Radek and Singer's dilemma. According to it, it is "those of us who are evaluating various moral theories in order to decide which one to accept" who are to be thought of as having these beliefs. If this is so, de Lazari-Radek and Singer allow that these beliefs *do* have a special place in the evaluation of ethical theories. "*Of course*," they tell us, "we should evaluate rival moral theories in terms of their ability to

12 Oddly enough even some *rule* consequentialists think that roughly something like this is true. See Brandt (1979, pp. 16-23; 1996, pp.120-121 and 134).

cohere with the convictions in which we have the most confidence after due reflection" (2010, p. 46, italics in the original). However, these facts do not show what Hooker thinks they do, at least according to de Lazari-Radek and Singer; instead of supporting RC, these facts support a version of act consequentialism. Hence, one way or another reflective equilibrium fails to support RC, de Lazari-Radek and Singer conclude.

Only a partial reply is required here. Without question, not everyone agrees with Hooker that considerations of reflective equilibrium better support RC than its rivals. De Lazari-Radek and Singer are not alone in disagreeing with Hooker about what theory reflective equilibrium best supports. Some, like David Brink (1989, pp. 122-142), agree with them that it better supports another form of consequentialism. Others, like Norman Daniels (1996, p. 127), claim that reflective equilibrium does not support any variety of consequentialism as well as it supports some forms of deontology. While I tend to side with Hooker in this matter, I do not hazard an argument for this thesis here. Instead, let me remind the reader that the question I consider in this paper is narrower than the one that Hooker, Brink, and Daniels have in mind. I contend only that reflective equilibrium better supports one version of RC over its rival. One should be able to agree with my main line of argument in this paper, even if one thinks that reflective equilibrium ultimately better justifies some form of, say, act consequentialism, contractualism, or virtue ethics than any form of RC; this is true even of de Lazari-Radek and Singer. So the second horn does not represent a danger for my present purposes, and their dilemma need not be a much of a concern for me.

That said, further attention to the first horn of de Lazari-Radek and Singer's dilemma also pays dividends since their characterization of reflective equilibrium is deeply flawed in at least two ways. First, it deals in false alternatives. Defenders of reflective equilibrium are not forced to identify the holders of reflectively endorsed beliefs either with the general public as a

whole or with professional philosophers only, as de Lazari-Radek and Singer seem to assume. Certainly, the former option is closer to the truth than the latter; one need not have a Ph.D. in philosophy in order to have one's reflectively endorsed beliefs taken seriously in the evaluation of an ethical theory! Indeed, it is wholly unclear whether having this sort of background is really much of an advantage when it comes to the accuracy of one's moral beliefs or behavior.[13] But there are more alternatives than de Lazari-Radek and Singer admit. To be sure, not every member of what these authors call "the general public" is capable of formulating relevant moral beliefs. Those who are insane, grossly undereducated, incapable of doing the requisite reflection in a reliable manner, etc., are excluded – and properly so. Moreover, de Lazari-Radek and Singer are correct to point out that, e.g., cultural and ethnic prejudices can be distorting influences on one's moral beliefs. Who would disagree? But that does not show that philosophers are the only ones we should consult about their reflectively endorsed beliefs. It only shows that we should not ask those with such biases (including philosophers, if they have such biases). Let me put my cards on the table: I doubt that there is anything like a tidy classification for denoting those whose beliefs are relevant to reflective equilibrium. Categories such as "the general public" or "professional philosophers" are far too simple minded, and more plausibly complex categories are not given to concise expression. Nevertheless, I do not think de Lazari-Radek and Singer present any reason to suppose that we need such a categorization in order to make progress with the evaluation of ethical theories.

Here is a second flaw in de Lazari-Radek and Singer's characterization of reflective equilibrium. The authors appear to presuppose that according to advocates of reflective equilibrium there will be *unanimity* about reflectively endorsed beliefs. And even Hooker

13 Schwitzgebel (2009), for one, raises a number of worthwhile questions about this matter.

himself emphasizes the "shared" nature of these beliefs (2010, p. 113). But it is a fantasy to expect unanimity – and an unnecessary fantasy at that. Unanimity would make the task of evaluating ethical theories easier, of course, but it is wildly out of character for sublunary types like ourselves to attain this level of agreement. It is enough that the vast majority of reasonable people agree about a given moral conviction. We need not wait until everyone is convinced. Majoritarian democracy is not quite the right model for the formation of reflectively endorsed beliefs, but neither is an American jury in a case of criminal law. In fact, many reflectively endorsed beliefs are little more than moral commonsense, and I make fairly liberal use of them below.[14] So the first horn of de Lazari-Radek and Singer's dilemma fails, not only as an obstacle to my limited purposes, but as an obstacle to the use of reflective equilibrium to evaluate ethical theories in general.[15]

There is a third issue here as well, though I only flag it here and leave a serious consideration of it until later in this paper. Though de Lazari-Radek and Singer do not use the term "Archimedean points," they often speak of our reflectively endorsed beliefs as if they were invariably decisive either for or against a given ethical theory.[16] De Lazari-Radek and Singer are not alone in doing so, but as I argue in Section 4 the relationship between reflectively endorsed beliefs and ethical theories is far more complex, though not unmanageably so.

3. The Costless Benefit Argument for Relative RC

Let me turn now to the task of arguing in favor of Relative RC. It is worth recalling that

14 Compare the derivation of "intuitively plausible principles" by Mulgan (2006, pp. 133-137).
15 See also Singer (1972; 1974; and 2005).
16 See, e.g., Daniels (1996, pp. 47-64), Verwijj (1998, p. 31), and Brom (1998, p. 193). The expression "Archimedian point" is sometimes used in the evaluation of ethical theories outside the tradition of reflective equilibrium. See. e.g., Williams (1985, pp. 22-29). However, that is not at all what I intend to pick out here.

this paper is a species of what is sometimes called an "immanent critique" - i.e., an attempt to show that even on Hooker's own way of looking at things, Relative RC is superior to Absolute RC. As a result, the case for Relative RC is also, in part, a case *against* Absolute RC.

I'll begin with what I take to be one among many of our reflectively endorsed beliefs about morality – namely,

> Costless Benefit: All other things being equal, a moral code is superior if it leaves some better off, when doing so leaves no one worse off.

Before I offer the argument for Relative RC, a little more setup is necessary. So consider the set of all possible moral codes as well as the set of all possible worlds in which there are moral agents. First, let C be the set of all possible moral codes {Code 1, Code 2, Code 3,..., Code n}, where a given set of rules is a moral code if and only if it meets the following two conditions:

(i) the rules are consistent; that is, there is no action such that the conjunction of rules requires an agent both to do and not to do that action;

(ii) for any morally relevant action that might be undertaken by an agent, the conjunction of the elements of either requires an agent to do an action F, permits but does require an agent to do F, or neither requires nor permits an agent to do F.

This is not meant to be a reductive definition or anything remotely close to it. Here I rely on a pre-theoretic understanding of the distinction between morally relevant and morally irrelevant

actions. Second, let W be the set of all possible groups of moral agents[17] {World 1, World 2, World 3,..., World m}. Obviously, it is not practical to represent anything even close to all of the elements of both sets here. Rather, for the sake of ease of presentation, I assume that there are only three possible moral codes and four possible worlds. However, the precise numbers of codes and agents make no difference to the matter at hand, provided that both sets are countable, if not necessarily finite.[18] Finally, turn to what I call the *value of each world under each code*.[19] By this phrase, I mean the value for everyone that results from having the agents in a certain possible world internalize a given moral code and act in accordance with it. Let Table 1 summarize all of this information, with the cardinal number in each cell representing for the value of each world under each code.

	Code 1	Code 2	Code 3
World 1	50	70	30
World 2	20	10	80
World 3	45	25	10
World 4	90	30	40
Sum of all Worlds	205	135	160

Here, then, is the argument for Relative RC: Absolute RC is preferable to Relative RC if and only if for all agents there is a single IMC. And there is a single IMC *simpliciter* only if there is a single moral code that is ideal for each of Worlds 1 through 4. But there appear to be different IMCs for at least some of these worlds. For Costless Benefit tells us that all others things being equal, a moral code is superior if it leaves some better off, when doing so leaves

17 I assume a fairly undemanding conception of transworld identity, such that, e.g., Obama in World 1 and Obama in World 2 are counterparts, not numerically identical agent in different possible worlds. See Lewis (1986).

18 Whether knowledge of this sort is available to the likes of us is not at stake here. But for useful discussion, see Breakey (2009).

19 Hooker (2000a, pp. 72-75) thinks that we should be concerned with expected value, instead of actual value. Contrast Mulgan (2001, p. 66; 2006, pp. 142-146). The issue between Hooker and Mulgan is a live one, worth careful attention. However, it is an unnecessary complications for the purposes of this paper, I remain neutral on this point and simply speak of "value" *sans* the phrase. But see Kahn (Unpublished).

no one worse off. And internalizing Code 1 in all four worlds leaves those in Worlds 1 and 2 worse off than internalizing Code 1 only in Worlds 3 and 4 while imposing Code 2 in World 1 and Code 3 in World 2. Moreover, these are distinct possible worlds, so they have an effect upon one another, ensuring that the all-others-things-being-equal condition is met. Therefore, it is false that Absolute RC is preferable to Relative RC and true that, all other things being equal, Relative RC is superior to Absolute RC.

In fact, this line of argument can be extended. Though the version of the argument considered a moment ago makes use of a comparison of possible worlds, the argument does not rely essentially on this fact. In order to see that this is the case, limit the data to a single possible world in which there are 4 populations. In this particular world, these populations do not interact and are not even aware of each other's existence. Table 2 represents all of the relevant information.

	Code 1	Code 2	Code 3
Population 1	50	70	30
Population 2	20	10	80
Population 3	45	25	10
Population 4	90	30	40
Sum of all Populations	205	135	160

The quantitative reasoning here is identical. If agents in all 4 populations internalize Code 1. In the second, agents in Population 3 and Population 4 internalize Code 1, but agents in Population 1 internalize Code 2, and agents in Population 2 internalize Code 1. Since the numbers in Table 1 and Table 2 are identical, the conclusion is identical as well. On the basis of Costless Benefit, the second scenario, in which some populations internalize different codes from others is superior. Once again, Relative RC is superior to Absolute RC in at least

some cases.[20]

4. The Abandonment Argument against Relative RC

Nevertheless, the Costless Benefit Argument for Relative RC is only part of the story. As I mentioned at the beginning of the last section, the case for Relative RC is, in part, a case against Absolute RC. And so it is – but only in part. In addition to showing that Relative RC is superior to Absolute RC in at least some cases, we must also show that there are no countervailing reasons of equal or greater force to prefer Absolute RC to Relative RC. In short, we must show that there are not cases in which Absolute RC is superior to Relative RC. Obliviously, I cannot consider every conceivable reason to prefer Absolute RC to Relative RC, I can challenge the two arguments for this conclusion which appear most often in the literature. I consider one such line of argument in this section and a second in the next two.

Call the first of these the *Abandonment Argument.* This argument begins with the claim moving away from Absolute RC to Relative RC "comes at a steep price," for it deprives the advocates of RC of "one of the chief attractions and motivations for adopting rule-consequentialism" in general, as Doug Portmore (2009) puts it. We might think of this as another reflectively endorsed belief – namely,

Generality: A moral code ought to be internalized by everyone.[21]

The worry, then, is that the move away from Absolute RC requires that "one abandon this intuition," i.e., Generality (Portmore, 2009) and that abandoning it is too costly.[22]

---

20 There are some similarities between this second scenario and a few thought experiments offered in Portmore (2009). I consider Portmore's arguments in Sections 4 and 5.

21 See also Mulgan (2006, p. 132).

22 Of course, Portmore rejects RC in all its forms, but we can ignore this fact for the moment. His argument

Now, there is an important point in the Abandonment Argument, but it needs to be addressed cautiously, and it does not show that Relative RC is inferior to Absolute RC or simply untenable as a whole – the opposite, in fact, is true. Showing why this is the case takes a bit of doing, but in the process we will also get a better idea of how reflective equilibrium is meant to evaluate ethical theories.

Recall, as before, that on our chosen way of evaluating ethical theories, we begin with reflectively endorsed beliefs about morality. While I cannot enumerate and defend all such beliefs here, several excellent candidates are obvious:

> Harm-to-the-Innocent: A moral code ought to protect the innocent from being harmed.

and

> Lying: A moral code ought to prohibit lying.

So we are bound by the terms of reflective equilibrium to use these two beliefs to evaluate ethical theories. Again, on this approach, ethical theories must show how these beliefs cohere with one another. However, it is impossible not to notice that in at least some circumstances Harm-to-the-Innocent and Lying pull in different directions regarding the evaluation of these theories.

Let me explain. On the one hand, if one ethical theory makes preventing harm a high priority, even when the only way to do so is to lie, then Harm-to-the-Innocent will count in

---

here concerns the superiority of Absolute RC to Relative RC.

favor of it, though Lying will count against it. On the other hand, if another ethical theory makes not lying imperative, even when that means that the innocent might be worse off as a result, Harm-to-the-Innocent will count against it, while Lying will count in favor of it. So it would be a mistake to think that any interesting set of reflectively endorsed beliefs about morality will speak univocally in favor of one theory and against others. But it would also be a mistake to think that such univocality is necessary. Using the test of reflective equilibrium to evaluate ethical theories requires us to find the *best* possible balance among our reflectively endorsed beliefs. It would be unrealistic to imagine that any combination of the two will be completely without tension. Because of this, I suggest that reflectively endorsed beliefs such as Harm-to-the-Innocent and Lying be understood as providing *prima facie* oughts which contribute to different degrees to the justification of one theory or another, rather than all-things-considered which determine once and for all whether or not a theory should be accepted. As a result, we do not need to reject either Harm-to-the-Innocent or Lying; rather, the best ethical theory must strike the right balance between them.

While all of this seems quite reasonable, one might wonder whether any reflectively endorsed beliefs provide all-things-considered oughts, or whether all provide only *prima facie* oughts. I count myself as fortunate in as much as nothing on this paper depends on the answer to this question. As I argue below, I need only assume that both Costless Benefit and Generality provide *prima facie*, instead of all-things-considered, oughts. It might be possible to narrow the scope of some reflectively endorsed beliefs to such an extent that we should reject any theory that is inconsistent with it. E.g., Tim Mulgan claims that there are at least some beliefs "that any acceptable moral theory must accommodate," though such beliefs are to be distinguished from others "that [simply] mark distinctive features of different theories" (2009, p. 116). One of Mulgan's examples of a belief that any acceptable moral theory must

accommodate is that "It is wrong to gratuitously create a child whose life contains nothing but suffering" (2009, p. 117). We should note that Mulgan stipulates that the action is *gratuitous*; i.e., there is no reason which justifies it. Since morality requires such reasons, it is hard to imagine an acceptable ethical theory which does not rule out such an action.[23]

However, caution is warranted even in a case such as this. We cannot know *a priori* whether or not every ethical theory faces at least some counterintuitive results. I would not be inclined to say that if this were so, we should reject every ethical theory. Ultimately, it is the comparative virtues and vices of ethical theories that must be used to judge them, not their absolute virtues and vices (if there really are such things). So even saying that Mulgan's "decisive intuition" provides an all-things-considered ought might be a source of regret (2009, p. 116). At any rate, the mere fact that there is some tension between a theory and a considered judgment is not sufficient to justify the rejection of one or the other. Far more than that would have to be shown, a point to which I return in just a moment.

Return now to the Abandonment Argument. According to it, the Costless Benefit Argument against Absolute RC in Section 3 requires us to "abandon" Generality. But it should now be clear that it does no such thing. Of course, Generality tells us that, *prima facie*, an ethical theory ought to be general, it ought to apply to everyone. But it does tell us that, all-things-considered, it ought to. There are other matters that must be taken into consideration as well. Moreover, it should also be clear that in Case-1 and Case-2 that Costless Benefit has priority over Generality. In order to avoid needless clutter, I condense the argument for this point by concentrating only on Population 2 and 3 in Case 2. Begin with Costless Benefit.

---

23 One might also understand Mulgan as claiming, not that the gratuitous action is not justified, but that the action is not undertaken for this or any other reason. However, if that it is the case, then I doubt that an action of this kind can never be ethically permissible. E.g., if 1,000 children will be tortured if I do not gratuitously torture one, then I suspect that I am ethically permitted, even required to do so if I can, though the usual problems (from the Toxin Puzzle, etc.) with getting myself to act for no particular reason threaten to make this very difficult.

Recall that Population 2 is significantly better off as a result of internalizing Code 3 rather than Code 1, and Population 3 is made no worse off since it still internalizes Code 1, Costless Benefit clearly applies to this case. Now turn to Generality. In spite of sharing a possible world, these populations are causally and even epistemically isolated from one another. Population 2 and Population 3 do not interact and do not even know of each other's existence. In light of this fact, it is difficult to see why "morality should be thought of as a collective, shared code" between these two groups. In fact, there is nothing that these two groups can share in any meaningful sense of the term. Moreover, considerations of equality does not at all appear to justify invoking the Generality in either case which Morey considers. The benefits that go to Population 2 do not leave Population 3 at a comparative disadvantage in their interactions since they do not interact at all. So I cannot imagine why Costless Benefit would have priority over Generality.

5. The Runaway Argument

Let me turn now to a second line of argument for preferring Absolute RC to Relative RC, in at least some situations. Hooker claims that accepting Relative RC leads to "runaway relativization" which "flies in the face of the idea that the same rules should apply to everyone" (2005, p. 267). Indeed, he goes so far as to say that

> The idea of relativizing codes to groups is on the road to relativizing them to sub-groups, and at the end of that road is relativizing them to individuals. To go down that road is to turn our backs on...the idea that morality should be thought of as a collective, shared code. (1998, pp. 28-29)

Portmore seems to buy into this view as well, claiming that the only way to avoid the criticism raised in Section 4 "is to modify [RC] so that it relativizes codes to individuals" (2009). Even Brandt comes close to accepting this view in the passage quoted in Section 1 (though he doubts that this fact counts against RC). Call the argument for this claim *the Runaway Argument*.

It is important at the outset to distinguish the Runaway Argument from the Abandonment Argument. The Abandonment Argument commits one to claiming that any ethical theory which does not treat Generality as providing an all-things-considered ought must be rejected for this very reason. However, the Runaway Argument does not commit one to any such claim. Rather, it proceeds roughly as follows: [1] If Relative RC is correct, then there are no non-arbitrary grounds for holding that moral codes should be determined at any level other than that of the individual; so [2] if Relative RC is correct, then the Strong Thesis is correct as well. Though neither Hooker nor Portmore makes it as explicit as this, I believe that they must think that the Runaway Argument ought to be expanded along the following lines: [3] If the Strong Thesis is correct, then Relative RC does not treat Generality even as a *prima facie* reason; so [4] Relative RC must be rejected. Though I argued in Section 4 that using the method of reflective equilibrium to evaluate ethical theories does not require that one give priority to moral convictions such as Generality in *every* case, it certainly requires that we do so in some cases. As a result, the Runaway Argument does pose a threat to Relative RC even though the Abandonment Argument does not. However, it is a threat which can be overcome because [1] is false, as I argue in the rest of this Section.[24]

---

24 One might ask why the Runaway Argument stops with individual agents. Why doesn't its logic force us to consider what codes would be appropriate to time-slices of agents? Though this question falls outside of the scope of the paper, the essence of my response to it can be reconstructed easily by the end of this section. At any rate, the possibility that the Runaway Argument takes us beyond individual agents is not a problem for me because, as I contend here, the argument is unsound.

Before proceeding to an explanation of why this is so, it is worth our while to clarify two points about the Strong Thesis. First, it has been suggested to me in conversation that if the Strong Thesis is correct, then Relative RC collapses into ethical egoism. Roughly speaking, ethical egoism is the view that an action is morally wrong if and only if it fails to have the best consequences for the agent herself.[25] Now, it is true that if the Strong Thesis is correct, then morally ideal codes are determined at the level of individuals, just as ethical egoism does. However, the Strong Thesis is concerned with more than just the self-interest of the agent. As mentioned in Section 1 of this paper, all forms of RC – or at least all forms of RC which I consider here – seek to maximize value for everyone. The agent's self-interest is part of this aggregation, but it does not play any privileged role in its determination. The only circumstances in which both ethical egoism and Relative RC would judge precisely the same actions to be wrong are ones in which the agent's actions can affect no one except herself. These are, suffice it to say, unusual circumstances. Clearly, then, ethical egoism and the Strong Thesis are distinct.

Second, it has also been suggested that if the Strong Thesis is correct, then Relative RC collapses into act consequentialism. As it is usually understood, act consequentialism makes rightness and wrongness of actions a function of each individual's actions. Hence, if Albert's doing F rather than any other alternative on this occasion will produce a greater amount of good, then it would be wrong for him not to do F. But, contrary to initial appearances, the Strong Thesis does not commit one to this claim. The question for this version of RC focuses on which moral code would, if internalized by the agent, bring about the greatest amount of value. On any given occasion, acting in accordance with this moral code might lead to sub-optimal results, at least from the perspective of act consequentialism.

25 See Brandt ([1972b] 1992, pp. 99-103), Parfit (1984, pp. 3-5), and Brink (1997, pp. 96-129).

As a result the two theories produce distinct accounts of which actions are right and which are wrong.[26]

But let me return to the main question of this section: Does relative RC lead to "runaway relativization"? To put the same question in slightly different terms, does Relative RC entail the Strong Thesis? Unsurprisingly, I argue that the answer is "no." One way to test whether Relative RC is true only if the Strong Thesis is true is to see how internalization would best work in typical early 21st century humans. In particular, assume the truth of Relative RC in the case of typical early 21st century humans and then ask whether the Strong Thesis follows. In doing so, we need to pay especially close attention to the question: "Would the results be better if codes were internalized at the level of groups or at the level of individuals?"

Recall that when determining the value of any set of circumstances under any code, we must take into consideration the cost of getting the code internalized.[27] So it is not just a matter of looking at the value of the consequences of having a certain group or individual act in accordance with a particular code; it is also a matter of the costs involved with getting the group or individual to accept this code as well as with maintaining their internalization. Hence the question that we have to address first is this: What are the implications of internalization costs for creatures like us (i.e., early 21st Century humans)? Note that for internalization to be possible for the likes of us, the code must meet several condition:

- Codes must be fairly simple.

26 Of course, some, e.g., Lyons (1965) and Smart (1973) have argued that *any* form of RC ultimately collapses into act consequentialism. But this argument is not a special problem for the Strong Thesis, and, at any rate, it has been answered at length elsewhere. See Brandt ([1963] 1992) and Hooker (2000, esp. pp. 93-99).

27 E.g., Brandt (1979, p. 287) and Hooker (2000b, pp. 78-80).

- Codes must be fairly general.

- Codes must be highly resistant to change.

These points require further discussion.

Why should a code be simple and general? The answer is, as Brandt puts it, that the code must "be suited to the level of intelligence and education of" those who are to internalize it (1979, p. 180). While human beings can be fairly clever and well-educated, we suffer from significant cognitive limitations. In particular, complex codes are too difficult for us to employ in many circumstances, and specific codes require us to internalize too many rules. The codes which most of us internalize when young – e.g., do not lie, do not steal, do not treat others in ways that you would not want to be treated –are models of simplicity and generality. While we learn to apply these rules with increasing sophistication and subtlety as we mature, we do not reject or replace the rules. Though it might be possible for many of us to increase the complexity or specificity of the codes which we internalize, doing so would come at a high cost.[28] For we face diminishing marginal returns from goods such as these. Why should codes be resistant to change? Humans face considerable temptations to violate their own moral codes in order to address their self-interest or simply to avoid what might be socially uncomfortable situations. Unless internalization is permanent (or close to it), we will need to re-internalize codes from time to time, repeating the initial costs of doing so.

Given the facts that for 21st Century humans, codes must be fairly simple, fairly general, and highly resistant to change, it is highly unlikely that anything like the Strong Thesis is correct. To begin with, there is a strictly limited number of simple general rules that

28 See, e.g., Shaw (1999, p. 164) and Hooker, Mason, and Miller (2000, pp. 1-2).

we can learn. E.g., with regard to lying, there are only three basic possibilities: [A] internalize a code that prohibits lying, [B] internalize a code that permits lying, [C] do not internalize any code with regard to this behavior. Both [B] and [C] are likely to have disastrous affects on community, personal relationships, business dealings, individual trustworthiness, and the like, so they are almost certainly not live options for any form of RC. It seems far more plausible that the best consequences would result from everyone in circumstances like our own internalizing a rule which prohibits lying. Though we can imagine rules that prohibit lying except in certain situations, these will very quickly become unmanageably complicated and specific and, therefore, unsuitable for internalization. At best, we might internalize a rule that prohibits lying except where not doing so will lead to grave harm to the innocent. But the point is clear enough. At least for creatures much like ourselves, matters will go best if all of us, rather than one of us, internalize this rule.

Another reason to think that codes for humans must meet all three conditions is found in the fact that internalization of moral codes is a social process. We use a wide variety of social institutions to internalize codes in the young – families, schools, religious institutions, legal institutions, and the like. We are most likely to internalize a rule that, say, prohibits theft, if all of these institutions are in agreement about the rule. That is to say, an agent is more likely to internalize a rule that prohibits theft if her family members, her teachers, her religious leaders, etc., aim at internalizing the same rule. It might be possible for an entire society to tailor-make a set of rules for each individual, but doing so is likely to be extraordinarily expensive. A far more cost-effective approach would be for the institutions of a social group to work together on internalizing a common code for all.

A third reason is provided by the need to reduce coordination costs within a social group. If everyone believes that everyone else has internalized a common code, then it will

be easier for members of the social group to coordinate actions. Example would be avoiding one-off prisoner's dilemmas. Though this is just an example.

Yet another reason has to do with the stability of codes. Differences in moral codes are likely to be seen as undercutting the codes in each person because of unfairness. It might be possible to internalize within each individual norms in such a way that they do not lead to this result, but that is likely to be quite expensive too.

A final reason: It is expensive to internalize codes, so we don't want to do this more than once.

Recall that if the Strong Thesis is correct, then for each agent there is a morally ideal code. Morey's reflections of the limitations of 21st Century humans clearly shows that the Strong Thesis is false. We humans are agents, and there are many circumstances in which it is simply false that the best results would be obtained by having each of us internalize a unique moral code. And there is nothing morally arbitrary about the fact Relative RC recommends different codes for social groups rather than individuals. RC is, after all, a form of consequentialism, and, at least in the case of humans like us, the best consequences to be obtained by internalizing codes comes at the group, rather than the individual, level.

The falsity of the Strong Thesis leaves open the question of whether the Weak Thesis or the Moderate Thesis is correct, but that is a question for another day.[29]

29 I first formulated many of the ideas explored in this paper in conversation with Roger Crisp, Brad Hooker, and Derek Parfit, and if I have managed to say anything worthwhile here it is largely thanks to them. I am also very grateful for both generous comments and insightful criticisms to audience members at the Central Division Meeting of the American Philosophical Association (March 2012), the Joint Session of the Aristotelean Society and the Mind Association (July 2012), and the annual meeting of the British Society for Ethical Theory (July 2012). Finally, I am much in debt to members of my own Department at the U.S. Air Force Academy and of the Colorado Springs Philosophy Discussion Group, especially James Carey, Marion Hourdequin, and Ivan Meyerhoffer. Finally, my thanks to Kimberly Kahn for comments and questions on the final draft of this paper.

References

Arneson, Richard (2005) "Sophisticated Rule Consequentialism: Some Simple Objections," *Philosophical Issues* 15, 235–251.

Badmington, Neil (2000) *Posthumanism* (London: Palgrave).

Bostrom, Nick (2005) "In Defense of Posthuman Dignity," *Bioethics* 19, 202–214.

Boyd, Richard (1988) "How to be a Moral Realist," in Geoffrey Sayre McCord (ed.) *Essays on Moral Realism* (Ithaca, New York: Cornell University Press).

Brandt, R.B. ([1965] 1992) "Some Merits of One Form of Rule-Utilitarianism," *University of Colorado Studies*. Reprinted in Brandt (1992), *Morality, Utilitarianism, and Rights* (Cambridge, UK: Cambridge University Press), 111-136.

Brandt, R.B. ([1972a] 1992) "Utilitarianism and the Rules of War," *Philosophy and Public Affairs* 1, 145-162. Reprinted in Brandt (1992), *Morality, Utilitarianism, and Rights* (Cambridge, UK: Cambridge University Press), 336-353.

Brandt, R.B. ([1972b] 1992) "Rationality, Egoism, and Morality," *Journal of Philosophy* 69. Reprinted in Brandt (1992), *Morality, Utilitarianism, and Rights* (Cambridge, UK: Cambridge University Press), 93-108.

Brandt, R.B. (1979) *A Theory of the Good and the Right* (Oxford: Oxford University Press).

Brandt, R.B. ([1988] 1992) "Fairness to Indirect Optimific Theories," *Ethics* 98, 341-360. Reprinted in Brandt (1992), *Morality, Utilitarianism, and Rights* (Cambridge, UK: Cambridge University Press), 137-157.

Brandt, R.B. (1996) *Facts, Values, and Morality* (Cambridge, UK: Cambridge University Press).

Breakey, Hugh (2009) "The Epistemic and Informational Requirements of Utilitarianism," *Utilitas* 21, 72-99.

Brink, David (1989) *Moral Realism and the Foundations of Ethics* (Cambridge, UK: Cambridge University Press).

Brink, David (1997) "Self-Love and Altruism," *Social Philosophy and Policy* 14, 122-157.

Brink, David (2006) "Some Forms and Limits of Consequentialism" in D. Copp (ed.) *Oxford Handbook in Ethical Theory* (Oxford: Clarendon Press).

Brom, Frans W.A. (1998) "Developing Public Morality: Between Practical Agreement and Intersubjective Reflective Equilibrium"in Wilbren van der Burg and Theo van Willigenberg (eds.) *Reflective Equilibrium: Essays in Honor of Robert Heeger* (Klewer Academic). 191-202.

Caplan, Arthur (2002) "Review of *Our Posthuman Future* by Francis Fukayama," *American Journal of Bioethics* 2, 57 – 61.

Copp, David (1996) *Morality, Normativity, and Society* (Oxford: Oxford University Press).

Crisp, Roger (2006) *Reasons and the Good* (Oxford: Oxford University Press).

Cummisky, David (1997) *Kantian Consequentialism* (Oxford: Oxford University Press).

de Lazari-Radek, Katarzyna and Singer, Peter (2010) "Secrecy in Consequentialism: A Defense of Esoteric Morality," *Ratio* 23, 34-58.

Daniels, Norman (1996) *Justice and Justification: Reflective Equilibrium in Theory and Practice* (Cambridge, UK: Cambridge University Press).

Driver, Julia (2002) "Review of *Ideal Code, Real World* by Brad Hooker," *Notre Dame Philosophical Reviews*, http://ndpr.nd.edu/review.cfm?id=1136.

Driver, Julia (2003) *Uneasy Virtue* (Cambridge, UK: Cambridge University Press).

Foot, Philippa (1985) "Utilitarianism and the Virtues," *Mind* 94.

Gauither, David (1986) *Morals by Agreement* (Oxford: Oxford University Press).

Gibbard, Allan (2008) *Reconciling Our Aims* (Cambridge, UK: Cambridge University Press).

Goodman, Nelson ([1955] 1983) *Fact, Fiction, and Forecast*, Fourth Edition (Cambridge, Massachusetts: Harvard University Press).

Harrison, Jonathan (1979) "Rule Utilitarianism and Cumulative Utilitarianism," *Canadian Journal of Philosophy* 5.

Harrod, R.M. (1936) "Utilitarianism Revised," *Mind* 45, 137-156.

Harsanyi, J. (1977) "Rule Utilitarianism and Decision Theory," *Erkenntnis* 11, 25-53.

Harsanyi, J (1985) "Does Reason Tell Us What Moral Code to Follow and, Indeed, to Follow Any Moral Code at All?" *Ethics* 96 (1), 42-55.

Hare, R.M. (1981) *Moral Thinking: Its Level, Methods, and Point* (Oxford: Oxford University Press).

Hooker, Brad (1994) "Rule Consequentialism, Incoherence, and Fairness," *Proceedings of the Aristotelian Society* 95, 21-35.

Hooker, Brad (1998) "Rule-Consequentialism and Obligations to the Needy," Pacific Philosophical Quarterly 79, 19–31.

Hooker, Brad (1996) "Ross-Style Pluralism versus Rule Consequentialism," *Mind* 105, 531-552.

Hooker, Brad (2000a) *Ideal Code, Real World: A Rule-Consequentialist Theory of Morality* (Oxford: Oxford University Press).

Hooker, Brad (2000b) "Reflective Equilibrium and Rule Consequentialism," in Brad Hooker, Elinor Mason, and Dale Miller *Morality, Rules, and Consequences: A Critical Reader* (Edinburgh: Edinburgh University Press), 222-238.

Hooker, Brad (2005) "Reply to Arneson and McIntyre," *Philosophical Issues* 15, 264-281.

Hooker, Brad (2008) "Act-Consequentialism versus Rule-Consequentialism," *Politeia* 24, 75–85.

Hooker, Brad (2009) "The Demandingness Objection," in Tim Chappell (ed.), *The Moral Problem of Demandingness* (London: Palgrave) 148–162.

Hooker, Brad (2010) "Publicity in Morality: Reply to de Lazari-Radek and Singer," *Ratio* 23, 111-117.

Hooker, Brad and Fletcher, Guy (2008) "Variable Versus Fixed-Rate Rule-Utilitarianism," *Philosophical Quarterly* 58, 344–352.

Kagan, Shelly (1989) *The Limits of Morality* (Oxford: Oxford University Press).

Kagan, Shelly (1998) *Normative Ethics* (Boulder, CO: Westview).

Kagan, Shelly (2000) "Evaluative Focal Points," in Brad Hooker, Elinor Mason, and Dale Miller *Morality, Rules, and Consequences: A Critical Reader* (Edinburgh: Edinburgh University Press), 134-155.

Kahn, Leonard (Unpublished) "Rule Consequentialism and Compliance."

Lewis, David (1986) *On the Plurality of Worlds* (Oxford: Blackwell).

Lyons, David (1965) *Forms and Limits of Utilitarianism* (Oxford: Oxford University Press).

Miller, Dale (2000) "Hooker's Use and Abuse of Reflective Equilibrium," Brad Hooker, Elinor Mason, and Dale Miller *Morality, Rules, and Consequences: A Critical Reader* (Edinburgh: Edinburgh University Press), 156-178.

Miller, Richard (2009) "Actual Rule Utilitarianism," *Journal of Philosophy* 106.

Montague, Philip (2000) "Why Rule Consequentialism is Not Superior to Ross-Style Pluralism," Brad Hooker, Elinor Mason, and Dale Miller *Morality, Rules, and Consequences: A Critical Reader* (Edinburgh: Edinburgh University Press), 203-221.

Mulgan, Tim (2001) *The Demands of Consequentialism* (Oxford: Oxford University Press).

Mulgan, Tim (2006) *Future People* (Oxford: Oxford University Press).

Mulgan, Tim (2009) "Rule Consequentialism and Non-Identity," in M.A. Roberts, D.T. Wasserman (eds.), *Harming Future Persons*, International Library of Ethics, Law, and the New Medicine 35.

Nussbaum, Martha (2005) *Frontiers of Justice: Disability, Nationality, and Species Membership* (Cambridge, Massachusetts: Harvard Belkanp).

Parfit, Derek (1984) *Reasons and Persons* (Oxford: Oxford University Press).

Parfit, Derek (2011) *On What Matters*, Volume 1 (Oxford: Oxford University Press).

Portmore, Douglas (2009) "Rule-Consequentialism and Irrelevant Others." *Utilitas* 21, 368-376.

Railton, Peter (1988) "How Thinking about Character and Utilitarianism Might Lead to Rethinking the Character of Utilitarianism," *Midwest Studies in Philosophy* 13. Reprinted in Peter Railton (2003) *Facts, Values, and Norms: Essays toward a Morality of Consequence* (Cambridge, UK: Cambridge University Press).

Rawls, John (1971) *A Theory of Justice* (Cambridge, Massachusetts: Harvard Belkanp).

Rawls, John (1993) *Political Liberalism* (New York: Columbia University Press).

Ridge, Michael (2006) "Introducing Variable-Rate Rule-Utilitariansim," *Philosophical Quarterly* 56, 242-253.

Riley, Jonathan (1999) *Mill on Liberty* (London: Routledge).

Riley, Jonathan (2000) "Defending Rule Utilitarianism," in Brad Hooker, Elinor Mason, and Dale Miller *Morality, Rules, and Consequences: A Critical Reader* (Edinburgh: Edinburgh University Press), 40-70.

Rosen, Gideon (2009) "Might Kantian Contractualism Be the Supreme Principle of Morality?" *Ratio* 22, 78-97.

Ross, W.D. (1930) *The Right and the Good* (Oxford: Oxford University Press).

Scanlon, T.M. (1998) *What We Owe to Each Other* (Cambridge, Massachusetts: Harvard Belkanp).

Scanlon, T.M. (2002) "Rawls on Justification," in Samuel Freedman (ed.) *The Cambridge Companion to Rawls* (Cambridge, UK: Cambridge University Press).

Schwitzgebel, Eric (2009) "Do Ethicists Steal More Books?" *Philosophical Psychology* 22, 711-725.

Shaw, William (2001) "Ideal Code, Real World: A Rule-Consequentialist Theory of Morality, by Brad Hooker," *Mind*.

Shaw, William (1999) *Contemporary Ethics: Taking Account of Utilitarianism* (Oxford: Blackwell).

Singer, Peter (2005) "Ethics and Intuitions" *Journal of Ethics* 9.

Singer, Peter (1974) "Sidgwick and Reflective Equilibrium," *Monist* 58.

Singer, Peter (1972) "Moral Experts," *Analysis* 32.

Skyrms, Brian (2004) *The Stag Hunt* (Cambridge, UK: Cambridge University Press).

Sinnott-Armstrong, Walter (2006) "Consequentialism," in *The Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/entries/consequentialism/.

Smart, J.J.C. (1956) "Extreme and Restricted Utilitarianism," *Philosophical Quarterly* 6, 344-354.

Thomas, Alan (2000) "Consequentialism and the Subversion of Pluralism," Brad Hooker, Elinor Mason, and Dale Miller *Morality, Rules, and Consequences: A Critical Reader* (Edinburgh: Edinburgh University Press), 179-202.

Williams, Bernard (1985) *Ethics and the Limits of Philosophy* (Cambridge, UK: Cambridge University Press).

Verwijj, Marcel (1998) "Moral Principles: Authoritative Norms or Flexible Guidelines?" in Wilbren van der Burg and Theo van Willigenberg (eds.) *Reflective Equilibrium: Essays in Honor of Robert Heeger* (Klewer Academic). 29-20.