

# YOU DON'T HAVE TO DO WHAT'S BEST\*

(A problem for consequentialists and other teleologists)

*S. Andrew Schroeder (schroed@fas.harvard.edu)*

*draft of 1 November 2010*

Consequentialists typically say not just that giving more to famine relief is very good, but that it's *required*. Teleologists who admit agent-relative value tell us not just that killing one to save five is very bad, but that it's *forbidden*. In this paper, I'll argue that claims like these, at least as they're usually understood, are unjustified. Even if the consequentialist is right that giving more to famine relief is very good, it doesn't follow that it's required; and even if the agent-relative teleologist is right that killing one to save five is very bad, that doesn't mean it's forbidden. From facts about what is good and bad, we can't reach conclusions about what is required and forbidden.

## I. INTRODUCTION

Many philosophers believe that for an action to be morally required is for it to be best, or for it to bring about the most value, morally speaking. More generally, many philosophers believe that deontic facts or norms (obligations, permissions, rights, etc.) hold in virtue of evaluative facts or norms (value, goodness, etc.).<sup>1</sup> Call theories according to which this is true *strongly teleological*. Standard versions of consequentialism are of course the first strongly teleological theories to come to mind, but they're far from the only ones. In *The View From Nowhere*, for example, Nagel argues that you ought not twist a child's arm in order to help your injured friends because of

---

\* I thank Selim Berker, Roger Crisp, Christine Korsgaard, Kristi Olson, Simon Rippon, Thomas Scanlon, Lucas Stanczyk, the members of Harvard's Workshop in Moral and Political Philosophy, Brown's Metaethics Reading group, and audiences at Fordham University and the College of William & Mary for helpful comments on earlier drafts of this paper. Discussion at the Arizona Workshop in Normative Ethics was extremely helpful and improved the paper greatly.

<sup>1</sup> The argument of this paper also applies, I believe, to many theories which hold that an action is required because it is supported by the balance of reasons. (The order of explanation is important: the argument certainly does not apply to theories according to which an action is supported by the balance of reasons because it is required.) The basic idea is that on many theories, reasons correspond to goodness or value: to be reason-giving is to have value. Such views are essentially equivalent to teleological theories. Theories which don't make reasons correspond to value may also be subject to my argument. Say that something has *shmalue* iff it is reason-giving. If *shmalue* shares with value certain key features, then the argument of this paper will apply. I believe that this is in fact the case for many theories which ground requirements in reasons, especially those which think of reasons on a weighing model.

reasons which arise from the personal point of view and which outweigh the impersonal reasons given by your friends' plight (1989 175-185). For Nagel, reasons correspond to values (1989 154), and so his view is that you ought not twist the child's arm because not twisting the arm brings about more agent-relative value. Some egoistic theories also fit the teleological pattern, saying an agent ought to do whatever would be best for her (Moore 1903 98-9).

Many theories, then – consequentialist, deontological, and egoistic – are strongly teleological. In fact, strongly teleological views can seem virtually inescapable. How could morality require anything other than what is morally best? Thoughts like this have led several philosophers to argue that *all* minimally plausible moral theories are strongly teleological. James Dreier (using 'consequentialist' to mean what I mean by 'teleological') says,

The simple answer...is that every moral view is consequentialist [i.e. teleological], that we common sense moralists as much as anyone are out to maximize the good. Of course, our understanding of the good may be an agent centered one, whereas the typical challenger has an agent neutral understanding, but this contrast will have to be engaged by some other argument. We don't have to be embarrassed by the charge that we are ignoring the good, because the charge is just false. (1993 24-5, *cf.* 2011)

Michael Smith agrees:

[W]hen we reduce moral norms to norms of reason and rationality...the consequentialist structure will be preserved. Meta-ethicists should therefore adopt...the working hypothesis...that all moral theories are, at bottom, forms of consequentialism. (2005 26, *cf.* 2003)

A similar argument can be found in Louise (2004), and the impulse also seems to be present in Broome (1991), Sen (1982), Portmore (2003), and perhaps Zimmerman (2008 2-5).

This strongly teleological picture yields an attractive simplicity at the foundational normative level. The basic normative "building blocks" are all evaluative, with deontic norms constructed out of them. For those with skeptical concerns, this is important: it means that we need to bridge the gap between the normative and non-normative only once, or need only one foundational argument for normativity. On the teleological model, once we provide a justification for evaluative properties, we essentially get deontic ones for free. Since deontic facts hold in virtue of evaluative facts, they don't need an independent foundation. Considerations such as these have frequently been cited as among the principal virtues of teleology.<sup>2</sup>

Among those who reject a strongly teleological analysis, many grant that

---

<sup>2</sup> See e.g. M. Smith (2003 576), Sinnott-Armstrong (2009 440-441), and, in a slightly different sense, Pettit (1991 237-8).

there is a *prima facie* obligation to do good, but believe that sometimes this obligation can be overridden by a “side constraint” or agent-relative permission. An agent-relative teleological approach can mimic these constraints and permissions, but according to these objectors it grounds them in considerations of the wrong sort.<sup>3</sup> Call a theory *weakly teleological* if it says that, in the absence of independent deontic norms, the deontic status of an action holds in virtue of evaluative facts. Most weak teleologists hold that an agent is required to do what is best, unless doing so is forbidden by an independently-grounded constraint or some other action is permitted by an independently-grounded option.<sup>4</sup> For the remainder of this paper, I’ll consider only situations in which it is not plausible to think that there exists a constraint and in which I’ll assume an agent has no desire to exercise an option. In such cases, strong and weak teleology will be equivalent.

Teleology, in both its strong and weak varieties, encompasses a large collection of views in moral philosophy. In this paper, though, I’ll argue that teleological theories, despite their intuitive appeal, theoretical simplicity, and apparent inescapability, are mistaken. There isn’t any interesting sense in which the consequentialist can say that an affluent citizen is *required* to give more to famine relief, the teleological deontologist can say that we’re *forbidden* from twisting the child’s arm, or the teleological egoist can say that we’re *obliged* to go to the dentist. In general, I’ll argue that there isn’t any interesting sense in which an agent is required to do what is best or most valuable.<sup>5</sup>

After laying some necessary groundwork (§II), I’ll first argue that the teleologist is unable to specify how the deontic is supposed to depend on the evaluative without relying on deontic presuppositions. This will show that the teleologist’s proposed analysis of the deontic in terms of the evaluative doesn’t work: teleology can’t be the whole story about moral requirements, because there must be at least some non-teleological requirements (§§III-IV). Nevertheless, it might seem as if there is a sense in which morality could still have a substantial teleological element. I’ll argue, however, that this isn’t the case. Even if we grant the teleologist the foundational deontic premises she needs, there still isn’t any interesting sense in which agents are required to do what is best (§V). I’ll conclude by offering two positive suggestions for those attracted to a teleological outlook (§VI).

## II. PRELIMINARIES

---

<sup>3</sup> See e.g. Mack (1998) and M. Schroeder (2007).

<sup>4</sup> Kagan (1989 47) and Portmore (2000), among others, argue that common sense morality takes this form.

<sup>5</sup> This leaves out non-maximizing versions of teleology. I’ll briefly discuss such theories in note 10, where I’ll explain why I think they don’t look likely to fare better than maximizing versions. But I lack the space in this paper to give satisficing approaches a full discussion.

1. *Deontic concepts and practical reasoning*

Throughout this paper, I'll assume a particular view of how deontic concepts operate in practical reasoning. To believe that some action is morally forbidden is to believe that that option is ruled out, according to morality. That is, an agent (at least one whose sole aim is to act morally)<sup>6</sup> can't consistently do some action while at the same time believing it to be morally forbidden. This seems to me to be indisputable, on the usual way of understanding 'morally forbidden'. What could it mean to think that some action is morally forbidden, other than to think that it is morally out-of-bounds? Hare puts the point this way:

One could not without raising logical as well as moral eyebrows say 'I must not let him down,' and at that same moment deliberately do just that. If anybody said this, the 'inverted commas' would be palpable.  
(1981 24)

The same idea is what motivates the thought that rights function as "trumps" (Dworkin 1984).

I think it's also the usage had in mind by teleologists when they make the deontic claims with which I opened this paper. When the consequentialist tells you you're required to give more to famine relief, she means to assert that the question of whether to give more to famine relief is settled. It wouldn't make sense to reply, "I understand that I'm morally required to give more, have no intention of acting immorally, and also have no intention of giving more." To say that would clearly reveal your irrationality, or at least that you hadn't understood what was being said. Similarly, you couldn't count as agreeing with Nagel that twisting the child's arm is forbidden if you also intended to do just that.

So, I'll assume that moral deontic terms function in practical reasoning to rule in or rule out certain options, at least for agents whose aim is to act morally. Notice that this does not require taking a stand on whether a moral theory must serve as a decision procedure. It's compatible with what I've said here that a moral theory might direct its agents to reason in a way that makes no use of deontic concepts. I'm only making the very weak claim that *if* an agent believes that some action is morally forbidden, then she is rationally obliged to rule the action out as a live option (or else to give up the belief that it is forbidden).<sup>7</sup> Denying that claim would mean accepting that an agent can, clear-headedly, both agree that some action is morally forbidden, and then out of a concern for morality do it anyway. If that's the case, I lose my grip on the idea and the importance of moral obligation.

---

<sup>6</sup> I actually think this proviso isn't necessary, but including it won't affect my argument.

<sup>7</sup> In fact, I'll only consider cases in which an agent has *knowledge* of deontic truths, so we need not worry about cases of false and/or unjustified beliefs.

## 2. *Arguments for teleology*

Given how widespread teleological theories are, it is surprising how little has been said on their behalf. It's typically taken for granted that to show an action is required it is sufficient to show that it is morally better than the alternatives. And it's nearly universally accepted that something's being best is at least a *prima facie* or defeasible consideration in favor of its being required. Presumably for this reason, defenses of the basic teleological outlook are frequently nothing more than a rejection of alternatives. This is especially evident in discussions of consequentialism, where what count as arguments for the theory are frequently just rebuttals of objections, or arguments against non-consequentialism.<sup>8</sup> The positive considerations that are offered usually don't defend the distinctively teleological aspects of consequentialism. Instead, they argue for a particular version of teleology – for example, for the view that all moral value is agent-neutral, rather than agent-relative.

Deontic claims, however, seem to me to be in special need of a positive defense. To say that an action is forbidden is to say that it's off the table, or out of play, morally-speaking. Since to an agent contemplating such an act the act does *seem* to be an option, a moral theory needs to explain in what sense the action is forbidden. The story need not, of course, be such as to convince any skeptic. But a moral theorist owes us some account of what it means to say that an action is forbidden. Why, or in what sense, is it true that I *must* not do this? What makes it the case that this action is off the table, morally speaking?

What positive considerations, then, can be offered on behalf of teleology? Just about all positive arguments in the literature focus on the idea of rationality. It's taken to be a requirement of *reason* that an agent do what is best.<sup>9</sup> Intuitively, the thought goes like this: suppose you're painting your bedroom and are about to buy Name Brand paint. I come along, however, and show you that Store Brand paint is better in all the relevant respects. It's cheaper, more durable, and easier to apply. It was produced in a more environmentally responsible way, and the color will better suit your room. At that point, it would seem crazy – irrational – for you to buy Name Brand. If Store Brand is better in every relevant respect, reason requires that you buy Store Brand. Thus, from the fact that Store Brand is *better*, we arrive at a *requirement* of reason.

Here is a somewhat similar but more abstract argument, loosely adapted from Haines (2008 §4a-b; cf. Lawlor 2009 113): for something to be good (or

---

<sup>8</sup> Hare (1981) is a notable exception.

<sup>9</sup> See M. Smith (2003 598), Smart (1973 42), Sidgwick (1907 382), Pettit (1991 238), Freeman (1994 315-16, 334), Scheffler (1985 414), Shaw (2006 16), Hare (1981 133), and (though he eventually argues for a satisficing theory) Slote (1989, e.g. at 115), among many, many others. Some philosophers distinguish between reason and rationality, but I'll use the terms interchangeably here. So, "requirement of reason" should be taken to mean "requirement of reason or of rationality."

valuable) is for there to be a reason in favor of it. Thus, for something to be best is for there to be most reason in favor of it. Suppose that X is best and Y isn't. It follows that there is more reason in favor of X than Y. It would therefore be contrary to reason, or irrational, to do Y instead of X. An agent is therefore rationally required to do what is best.

There are a number of differences between these two arguments, and neither is close to air-tight. But I think both of them are faithful to the basic idea motivating the teleologist: the thought that it's irrational to choose something worse, when you know that a better option is available. Let's call this thought the *core teleological intuition*, or CTI.<sup>10</sup> We needn't here worry about the precise form of the CTI, but for the bulk of this paper I'll grant the teleologist that something in its neighborhood is justifiable.<sup>11</sup> Important for our purposes is that the CTI purports to make the requirement to do what is best a requirement of reason or rationality. If I'm right that the CTI is what teleologists take to lie behind teleological requirements, then a successful defense of teleology will use the CTI to justify a principle showing exactly how the deontic is supposed to depend on the evaluative. The majority of this paper will be focused on looking for such an argument.

### III. TELEOLOGICAL PRINCIPLES

A teleological theory must begin by specifying a *teleological principle* – a principle showing exactly how the deontic is supposed to depend on the evaluative. In this section, our aim will be to formulate a plausible teleological principle, which we'll later seek to justify using the CTI. Let's begin, then, with the most obvious teleological principle, the one most closely tied to the CTI, and the one most frequently mentioned by teleologists:

*Simple teleological principle (STP).* An agent is required to do

---

<sup>10</sup> Notice that if the CTI is the basis for teleology, that would count against satisficing versions. In addition, then, to the obvious problem that faces any satisficer – that of specifying the cutoff point between the permissible and the forbidden – there is a second, perhaps deeper problem. The satisficer can't appeal to the same intuition which I believe has made teleology so attractive, so she'll need to provide a different foundation for her approach. Since in this paper I'll focus on arguments involving the CTI, I'll largely ignore satisficing proposals. But I do think that this points to a serious problem for satisficers. (Similar comments might also apply to rule-oriented versions of teleology. See Foot (1985 198) and Portmore (2003 304n5) for thoughts along these lines.)

<sup>11</sup> See Foot (1985 198), Scheffler (1985 414-415), and Portmore (2003 314-320) for defenses of the CTI.

what is best.<sup>12</sup>

The STP has a number of well-known problems. Perhaps the most clearly decisive objection, though, comes from scenarios like this one:

*Hurricane.* You're on a rescue boat in the middle of 26 islands, conveniently labeled A through Z. You know that 99 people are on A. Of B through Z, one of them (you don't know which) has 100 people, and the other 24 have zero. A hurricane is approaching and will surely kill anyone remaining on an island. You have time to get to one, but only one, island before the hurricane arrives. Where should you go?<sup>13</sup>

I take it that in *Hurricane*, any sane moral agent will head straight for island A. It would truly be a reprehensible (or at least seriously misguided) person who went anywhere else. According to the STP, however, going to A is forbidden and you know this. One of the islands has 100 people, and it's better to save 100 lives than to save 99. Therefore, the STP says that going to the island with the 100 is required, and that all other actions – including going to A – are forbidden. Since you know this, you are rationally bound to remove A from consideration. (This follows from our earlier discussion about the role of deontic concepts in practical reasoning.) In other words, if you accept the STP you should think to yourself, *Well, I know that A is out. The real choice is among islands B-Z.* This is absurd. On any reasonable moral theory, you are required to go to A, not cross it off from the start.

Cases like *Hurricane* aren't at all unusual. In fact, if you're an agent who would donate your winnings to famine relief, you face a *Hurricane*-like situation every time you have the opportunity to buy a lottery ticket. Jackson says,

[I]t is easy to slide into thinking that consequentialism holds that people should aim at the best consequences...whereas in fact most of the time we should select an option which we know for sure does not have the best consequences. Most of the time we are in the position of the person who declines to bet. The right option is a "play safe" one chosen in the

---

<sup>12</sup> Throughout, I'll somewhat awkwardly use terms like 'requirement', 'obligation', and 'duty', when 'ought' and 'should' would be more natural. I do this to emphasize that my objection is specifically to the deontic claims of teleologists. 'Ought', notoriously, has many senses, some of which are deontic and some of which aren't. It may be true that there is a non-deontic sense in which agents indeed ought to do what is best. If so, I suspect that teleologists may sometimes unintentionally trade on the deontic/non-deontic ambiguity in 'ought'. I thank Jamie Dreier, Mark van Roojen, and Mark Schroeder for discussion on this point.

<sup>13</sup> This case is structurally similar to Parfit's (*forthcoming*) much-discussed Mine Shaft case and Jacob Ross's Three Envelope problem. The basic form is much older, though. See, for example, Jackson (1991 462-3, 466-8).

knowledge that it does not have the best consequences and in ignorance of which option does have the best consequences. (1991 468)

Most teleologists agree with this, and for that reason say that there is at least some important sense in which an agent is required to maximize not actual, but instead expected value. Frequently, they appeal to a distinction between objective and subjective requirements. *Objectively* you are required to do what in fact maximizes value (and go to whichever island has the 100), but *subjectively* you are required to maximize expected value (and go to island A).

If this is the teleologist's proposal, let me stipulate that for the remainder of this paper, I'll be concerned only with subjective requirements. Earlier, I argued that deontic concepts have a specific function in practical reasoning: to rule in or rule out certain options. Objective requirements, if they're given by something like the STP, can't have this function, since cases like *Hurricane* show that it must be consistent to say, "I know that objectively this action is morally forbidden, yet out of a concern for morality I'll do it." But a consequentialist presumably doesn't mean to allow for this response when, for example, she says that I'm forbidden from treating myself to an expensive dinner. Since, then, objective requirements don't have the kind of practical function we ascribe to deontic concepts in this sort of discourse, for the remainder of this paper I'll be interested only in subjective requirements.<sup>14</sup>

Here, then, is the principle that teleologists usually propose for subjective requirements:

TP#2. An agent is required to do what maximizes expected value.

TP#2 easily solves *Hurricane*. The expected benefit of going to island A is saving 99 lives. The expected benefit of going to any of B through Z is saving 4 lives, since there's a 1 in 25 chance of saving 100 and a 24 in 25 chance of saving zero. Obviously, saving 99 lives is much better than saving 4 lives, and so TP#2 sensibly says that your duty is to go to A.

Unfortunately, though, TP#2 is under-specified. The expected value of an action varies depending on what body of information we take as a starting point. For example, although it maximized expected value for you to go to A (since you didn't know where the 100 were), it would maximize expected value for me to go to the 100, if I know where they are. In order to calculate expected value, we therefore need to know what body of information to

---

<sup>14</sup> This is not to say that objective requirements are of no importance. It could be, for example, that subjective requirements are somehow defined in terms of objective ones, or that objective and subjective requirements are both species of the same basic modal operator, with different parameters specified. (See e.g. Dowell (*forthcoming*) and Björnsson and Finlay (*forthcoming*)). In either case, the teleological principles I discuss below aim to capture whatever type or form of requirement has the practical function I've identified.



work from. So we might try:

*TP#3.* An agent is required to do what maximizes expected value, in light of her current beliefs.<sup>15</sup>

TP#3 has the virtue of being better-defined than TP#2, but there are three problems with it. First, it may still be undefined. Second, it's too strong (in one sense). And third, it's too weak (in another). I'll explain these objections in turn.

First, agents frequently have contradictory or incomplete beliefs. In such cases, it's often not clear how to calculate expected value. Suppose I believe that island A has 99 people, B has 25 families of four, and A has more people than B. What maximizes expected value, given those beliefs? Or, what if I believe that A has 99, but I have no beliefs at all about how many people B has? Once again, there doesn't seem to be any way to even begin the calculation. I suspect this is a very serious problem, but I'll set it aside for now.<sup>16</sup>

Next, let's see why TP#3 is too strong. Suppose in *Hurricane* you get a phone call from a trustworthy but eccentric mathematician. He informs you that, by using the tools of advanced topology, set theory, and chaos theory, he has confirmed that the location of the 100 is in fact logically entailed by your current beliefs. You, of course, don't have the slightest idea how to figure out where the 100 are, but you do nevertheless now know that their location is entailed by your current beliefs. Therefore, going to the island with the 100 is what maximizes expected value, in light of your current beliefs. Or: going to the island with the 100 is what would maximize value, were the world as you currently believe it to be. TP#3 thus requires you to go that island and (accordingly) forbids going to A. You know this, and so you're rationally obliged to remove A from consideration – exactly the absurd conclusion that doomed the STP.

Of course, the solution to this problem is obvious. Instead of taking into account *all* the logical consequences of your beliefs, we need to weaken TP#3, somehow restricting the kinds of inferences allowed. That the answer could be determined with advanced topology is irrelevant to what *you* are required to do. That it could be determined by multiplying 4 and 25, or by using a single, obvious instance of *modus ponens* is not. The difference seems to be that it's reasonable to think that a typical agent is morally responsible for completing – *ought* to complete – simple inferences, but it's unreasonable

---

<sup>15</sup> For versions of this principle see Smart (1973 47) and Jackson (1991).

<sup>16</sup> While it may be implausible to suppose that I have no beliefs at all about how many people are on island B, I very plausibly have no beliefs about the far downstream consequences of my actions. TP#3 will therefore frequently yield an indeterminate verdict for any theory (such as most versions of consequentialism) which takes into account the distant effects of an action.

to think that a typical agent is morally responsible for making complex mathematical calculations that far outstrip her training. In order to answer this objection, then, we'll need to appeal to a normative standard. I'll return to this shortly.

The third objection to TP#3 is that it is too weak. In addition to the beliefs that an agent in fact has, there are other beliefs that it seems an agent *ought* to have, or is responsible for having. You ought to know your best friend's birthday. You ought to remember that you promised to feed your neighbor's cat. You ought to know whether there is someone behind your car before you back up. And, if teleology is going to work, you ought to know that saving 99 lives is better than saving 4 lives.<sup>17</sup> A failure to have those beliefs will, according to TP#3, automatically excuse forgetting a birthday, breaking a promise, running over a pedestrian, or failing to go to island A. I don't think this is what most teleologists have in mind. That you have forgotten a promise doesn't mean that you need no longer fulfill it; ignorance may sometimes preclude obligation, but it doesn't always do so.<sup>18</sup>

These latter two objections, then, are both solved in the same way: by introducing normativity into our analysis.<sup>19</sup> This idea isn't completely alien to teleologists. Parfit, for example, says that (subjective) rightness is based on "what [a] person believes, or *ought to believe*, about [an action's] effects" (1986 25, emphasis added). And Shaw says that an agent ought to do "what is likely to have the best results as judged by what a *reasonable and conscientious* person in the agent's circumstances could be expected to know" (2006 8,

---

<sup>17</sup> This may seem like a surprising addition, since many similar views in moral philosophy stipulate that we work only from an agent's non-normative beliefs. For example, an ideal observer theorist might imagine a being with knowledge of all non-normative facts, and then ask what (normative) conclusions she would arrive at. In the present case, however, working from only an agent's non-normative beliefs would be unmotivated. Recall that we're seeking to make the teleological requirement one of reason. If basic normative beliefs (e.g. that saving 99 lives is better than saving 4) are such that it's not the case that an agent ought to have them, then there would be nothing irrational about an agent failing to have them, and accordingly nothing irrational about an agent failing to act on them. So, if we're to say that an agent is required to save the 99 because it would be better to do so, and if that requirement is supposed to be one of rationality, then the teleologist must say that such normative beliefs are ones an agent ought to have. (See Zimmerman (2008 38-39) for some related points.)

<sup>18</sup> Some teleologists may be willing to bite this bullet. Unless, however, they're also willing to bite the bullet in the mathematician case (or have some other analysis of that case), I think they'll still need to accept something like the fourth teleological principle I propose below. Once they've done that, there seems to me no remaining reason to bite the bullet here. Also, biting this bullet comes with other steep consequences. Suppose I believe (falsely, according to the consequentialist) that personal and impersonal values are incomparable. It seems to follow that I *never* have an obligation to give to Oxfam. (Giving to Oxfam wouldn't be better than blowing my cash at a casino, given my actual beliefs; they'd be incomparable.)

<sup>19</sup> Normativity also helps with, though it doesn't completely solve, the first objection. In at least some situations, we can say of an agent with contradictory beliefs that she ought not have them, and of an agent with incomplete beliefs that she ought to have certain beliefs.

emphasis added).<sup>20</sup> The argument here has shown that normativity enters into the analysis in two places. We need to identify certain basic beliefs as ones an agent ought to have, and also certain inferences as ones she ought to make. If we understand 'ought to believe' as covering both of these cases, and if we're careful not to read 'in light of' as bringing in the logical consequences of a set of beliefs, we could say:

*TP#4.* An agent is required to do what maximizes expected value, in light of what she ought to believe.

There are a number of teleological principles similar to TP#4. I won't worry about adjudicating among them, since what's important for our purposes is that they'll all include a normative component. In specifying what an agent is required to do, they'll all make reference to what an agent ought to believe.<sup>21</sup>

#### IV. UNDERSTANDING TP#4

In its analysis of moral obligation, TP#4 includes a reference to what an agent ought to believe. The thought that certain roles or activities come with epistemic responsibilities is a familiar one. Drivers ought to know about pedestrians in their path. Hikers ought to know the boundaries of public lands. Pharmacists ought to know about drug interactions. Promisors ought to remember their promises. Friends ought to remember each others' birthdays. And so forth. 'Ought' claims like this raise a number of interesting issues, but most important for our purposes is that *prima facie* they

---

<sup>20</sup> For similar normative analyses of expected value, see Smart (1973 47), Pettit (1991 233), and Kagan (1998 65). Normativity enters other analyses of expected value covertly. Oddie and Menzies, for example, suggest that an agent should act on her "best estimates, in light of her information" of the value of certain options (1992 532). If 'best' means anything there, it would seem to establish a normative standard: a quick-and-sloppy estimate isn't adequate. Zimmerman (2008 36) also offers an avowedly normative proposal, saying that agents should act based on available evidence, which is "evidence of which someone *can*, in some sense, and *ought*, in some sense, to avail himself." Unfortunately, he continues, "I confess that the exact senses of this 'can' and 'ought' elude me." He stipulates that the 'ought' must be epistemic (and therefore not moral), but in the next section, I'll argue that we shouldn't accept that assertion.

<sup>21</sup> In her paper in this volume, Holly Smith argues against principles like TP#4. I can't hope to do justice to her arguments in this short note, but let me briefly say that I believe some of her examples provide good reason for not allowing the 'ought to believe' clause in TP#4 to range over everything that an agent ought to believe. Rather, only certain kinds of beliefs are eligible for inclusion – perhaps (roughly) those that could be arrived at through a kind of introspection or calculation, but not those which required past information gathering. Examples like that are sufficient to drive my argument, I think. Also, I should mention that while I think that this is the *best* teleological principle, I also believe it ultimately fails. So I don't find it troubling that compelling arguments can be made against it.

are deontic.<sup>22</sup> When we say that you ought to know whether there is someone behind your car before you back up, we mean that morality *requires* it of you – that it is unacceptable, morally speaking, to back up without knowing. Promisors have a *duty* to remember their promises.

This creates a problem for the teleologist. She was trying to analyze deontic norms in evaluative terms, but has apparently ended up with a deontic term in her analysis. That doesn't mean that TP#4 is false or unhelpful. TP#4 says that a very general class of deontic truths can be understood using a combination of evaluative facts and a very narrow class of deontic truths. Such an analysis would be, if correct, quite significant. That said, it would come at a cost for the teleologist. As we saw, one of the main attractions of the teleological project was its promise of a simplified normative picture at the foundational level. If all deontic norms were grounded in evaluative ones, the basic normative “building blocks” would all be evaluative. But if TP#4 includes a deontic component in its analysis, this project will have failed: we won't be able to eliminate the deontic from the foundational normative level. There will be at least some non-teleological deontic norms.

In order to avoid this conclusion and defend a fully teleological approach, the teleologist has two options. She can argue either that the ‘ought’ in TP#4 can be analyzed away, or else that it isn't really a deontic, practical norm. The former route is clear enough. Teleologists generally try to analyze deontic norms in evaluative terms, and so it seems natural for the teleologist, if she agrees that this ‘ought’ is deontic, to try to analyze it in evaluative terms. This strategy, though, seems unlikely to work. In the last section, we found ourselves unable to give an analysis of deontic norms in evaluative terms; we ended up with an (apparently) deontic remainder. The same thing seems likely to happen here. If we try to analyze this deontic ‘ought’ teleologically, we'll for the same reason end up with a deontic remainder.<sup>23</sup>

The second option, to argue that this ‘ought’ isn't a deontic, practical one, seems more promising. Since the ‘ought’ concerns belief, it's natural to wonder whether it might be epistemic. A driver ought to look before backing up because she ought to know that there is a significant chance of hitting a pedestrian if she doesn't. An agent who didn't know that

---

<sup>22</sup> Among the concerns one might have is that it is inappropriate to require agents to have certain beliefs, since agents don't have voluntary control over what they believe. This, though, isn't a serious problem in the examples I'll consider. First, in each case the belief is true and evidence for it is or was available to the agent. Second, in at least some cases we can understand these obligations as necessary conditions for engaging in actions, rather than as simple obligations to have beliefs: a pharmacist ought not dispense drugs, *unless she knows the relevant drug interactions*.

<sup>23</sup> In my (2009), I argue for this claim in more detail. In that manuscript, I also argue against the possibilities that the ‘ought’ might be evaluative, or might be some third kind of practical norm, neither deontic nor evaluative.

pedestrians are frequently behind parked cars would, it seems, be *epistemically* defective.

This is, I think, the most likely way for the teleologist to deal with the leftover 'ought' in TP#4. But in the remainder of this section, I'll suggest that this proposal ultimately doesn't work. The 'ought' in TP#4 isn't, or at least isn't entirely, epistemic. (The considerations here aren't entirely decisive, but, fortunately, most of what I say in the remainder of the paper won't depend on this argument.)<sup>24</sup>

My argument will proceed by showing that in order to reach the kinds of conclusions she needs (e.g. that agents ought to know about pedestrians), the teleologist will need to distinguish between propositions for which the evidence is comparable. That is, she'll need to say that agents are required to believe proposition P but aren't required to believe proposition Q, where the evidence for P and Q is equally compelling. But purely epistemic requirements should be based on evidence and will therefore be content-neutral in the following sense: whether or not an agent epistemically ought, or e-ought, to believe some proposition won't be affected by the proposition's moral content or importance. The fact that a proposition is *morally* significant can't by itself affect the *epistemic* reasons concerning it. If an agent is obliged to believe P because P has moral content or because P plays an important role in a piece of moral reasoning, then that requirement seems more naturally viewed as moral or practical, rather than epistemic. And if the obligations we're worried about are at least partly moral, then the teleologist won't be able to eliminate deontic norms from the foundational moral level.

Let's begin with two simple scenarios. In the first, someone promises a vacationing neighbor that she'll feed his cats; and in the second, a driver is preparing to back out of a driveway across a stretch of sidewalk in an unfamiliar part of the city. Of the promisor, the teleologist might say that the fact that the agent consciously made the promise, that the event was phenomenologically vivid, that she hasn't recently suffered a head injury, and so forth all collectively make it the case that she e-ought to believe that she promised to feed her neighbor's cats. TP#4 says that she is required to do what would maximize expected value, given what she e-ought to believe – in this case, that she's promised to feed the cats. Given that she's promised to feed the cats, it's likely that no one else will. If no one else will feed the cats, they'll die if she doesn't feed them. This is a very bad thing, and hence it would maximize expected value for her to feed the cats. Thus, TP#4 delivers the intuitively correct conclusion, that she is required to feed the cats.

---

<sup>24</sup> The argument beginning in the next section requires only that TP#4 includes a normative standard which agents sometimes fail to meet, and agents do violate epistemic requirements. So, even if my argument here fails, I think that with only minor modifications the rest of the paper could remain largely intact. That said, a number of other philosophers who have considered the question of 'ought's like this have agreed that they must be practical. See e.g. H. Smith (2011), Rosen (2003), and Guerrero (2007).

Similarly, in the driving case the teleologist could say that given the agent's frequent experience walking around the city, she e-ought to know that at any given time there is a non-trivial chance that a pedestrian will be on any given stretch of sidewalk. TP#4 says that she is required to do what maximizes expected value, given that there is a non-trivial chance that a pedestrian will be on the stretch of sidewalk behind her car. Given that assumption and the trivial cost of looking before backing up, it would maximize expected value to look before backing up. Thus, TP#4 directs her to look before backing up.

But things aren't quite so simple. Suppose that last Wednesday you ordered pasta in the faculty dining room. At the time, the event may have been just as phenomenologically vivid as a typical episode of promising, and you (I hope) haven't since suffered a head injury. It would seem to follow, by reasoning parallel to that above, that you e-ought to believe you ordered the pasta for lunch last Wednesday. Similarly, if the driver's experience in the city means that she e-ought to have beliefs about the general prevalence of pedestrians, your experience walking around the halls of the philosophy building would seem to suggest that you e-ought to have beliefs about how many of your colleagues have office hours on Thursdays, have detached earlobes, or are left-handed. But those conclusions don't seem correct. It doesn't seem like you e-ought to remember what you ordered for lunch last week or to know what fraction of your colleagues have office hours on Thursdays. Given your experiences, you certainly may have epistemic license, or *permission*, to have those beliefs. But it doesn't seem right to say that you have an epistemic *obligation* to do so.

Indeed, if we did think that you e-ought to know what you ordered for lunch last week or when your colleagues have office hours, we'd run into problems. Suppose we discover that a food-borne pathogen has been spreading, and it is therefore crucially important for you to remember what you ate last week for lunch. We wouldn't say that you'd violated any requirement, epistemic or moral, I think, if you couldn't remember. TP#4, however, implies that you've violated requirements of both types. Or, suppose a fire is rushing through the university on Thursday. You're outside the philosophy building and firefighters ask you whether any of your colleagues were holding office hours this morning. If we said that you e-ought to know your colleagues' office hours, then it would follow from TP#4 that you'd be morally required to send the firefighters in if there are colleagues holding office hours, and required not to send them in if there aren't. But it doesn't seem to me that we'd necessarily think there was anything morally deficient with you if you told the firefighters you didn't know. And we *certainly* wouldn't put any such failing in the same category as backing up a car carelessly. There is an asymmetry, then, between remembering promises and knowing about pedestrians, on the one hand, and remembering lunch orders and knowing about office hours, on the other, which the current proposal doesn't capture.

The difference, of course, is that promises are likely to be of critical moral importance, whereas lunch orders aren't. And the prevalence of pedestrians is predictably of huge moral significance for drivers, whereas Thursday office hour scheduling isn't likely to be of moral significance for professors. *That* is why agents ought to remember promises but not lunch orders, and why drivers ought to know about pedestrians but professors aren't required to know about colleagues' office hours. The content-neutrality of epistemic norms, however, precludes taking this difference into account in any direct way. If the teleologist said that agents ought to remember promises but not lunch orders because the former are likely to be morally significant, that would amount to endorsing a non-epistemic norm.

If the teleologist is to pursue the epistemic strategy, then, she'll need to do justice to the thought that the reason agents ought to remember promises is that promises are likely to be of moral importance. Given content-neutrality, the only way to do that is to build the importance into the content of the belief. That is, the teleologist will need to say not that agents e-ought to remember their promises, but instead that they e-ought to believe that remembering their promises is likely to maximize expected value. Drivers e-ought to believe not that pedestrians are likely to be around, but instead that having accurate beliefs about the presence of pedestrians is likely to maximize expected value.

If the teleologist makes those claims, she can get the proper conclusion in all four cases. If an agent e-ought to believe that remembering a promise would maximize expected value, then TP#4 says she ought to do what would maximize expected value, given (what she e-ought to believe) that remembering the promise would maximize expected value. It follows trivially that she ought to remember the promise. On the other hand, it's not likely to matter, morally, if you remember your lunch orders, so surely it's not the case that you e-ought to believe that remembering your lunch orders will maximize expected value. We therefore are able to properly distinguish the cases.

Is it true, however, that agents e-ought to believe that remembering promises and having accurate beliefs about pedestrians is likely to maximize expected value? As before, these are plausibly beliefs to which normal agents have epistemic *license*, but it's not clear to me that there is an epistemic *requirement*. If those beliefs are e-required, it is presumably because they're so obvious. *Of course* remembering promises and knowing about pedestrians maximizes expected value. Those claims, though, are actually quite complex. They involve empirical judgments about the world (that backing into a pedestrian causes injury, that cats can't live a week without food), substantive abstract judgments (that injuries and dead cats are bad, that the badness of a dead cat is comparable to the badness of personal inconvenience), probabilistic assessments (how likely is it that a pedestrian will be struck? how likely is the victim to receive an insurance windfall?), and mathematical calculations of expected value (sometimes involving very small probabilities

and very high values). In addition to being complex, these are also claims that very few people – moral philosophers of a certain stripe possibly excepted – have ever thought about. Given all of that, I don't think it would be reasonable to expect many people to instantly affirm the truth of such claims.<sup>25</sup>

The sense, then, in which these beliefs are obvious – if indeed they are obvious – is that an agent who reflects on them e-ought to affirm them. That is, an agent who considers whether or not remembering promises maximizes expected value e-ought to conclude that it does.<sup>26</sup> Consider a number of other beliefs which are similar in that respect. How many windows are in your house? If two triangles have two sides of equal length and one angle of equal measure, does that guarantee the triangles will have the same area? Which cards should be turned over in the Wason Selection Test? Could 1000 M&Ms fit into a one liter soda bottle? Most people, I think, haven't (recently) considered most of these questions, and few have tip-of-the-tongue answers to any of them. But they are questions that we might expect an intelligent person could, upon reflection, answer correctly.

Now, it doesn't matter to me whether we want to say that this means agents have beliefs about such matters. If beliefs, for example, are dispositions to affirm or deny statements, then it might turn out that you do believe your house has eighteen windows, since that's what you'd say if you were asked. This could be true even if you've never thought about how many windows your house has. Let's stipulate, however, that an *immediate belief* is one an agent would be disposed to more-or-less immediately affirm or act upon. So, you perhaps e-ought to immediately believe that Shakespeare wrote *Hamlet*, but it's not necessarily the case that you e-ought to immediately believe that your house has eighteen windows or that remembering promises maximizes expected value – even though those are claims that, if you considered them, you e-ought to come to immediately believe.

---

<sup>25</sup> Daniel Star has suggested to me that a teleologist might try to get around this problem by insisting that the “obvious” belief in question is the (much simpler) belief that remembering promises is of great moral importance. This may help, though it seems subject to two problems. First, it is no longer so obvious what TP#4 directs an agent to do. If it is of great moral importance that agents remember promises, would it maximize expected value to remember some promise? Not necessarily. Second, many agents have much more specific beliefs about the moral importance of promises that would make the general belief unhelpful for teleological analysis. An agent, for example, might believe that it is morally important to remember promises, *in the sense that doing so pleases God*. That kind of “moral importance” isn't relevant to the teleological calculation. It's not enough, therefore, that agents believe that remembering promises is morally important; they need to believe it *in the sense that doing so maximizes expected value*, for it to have clear relevance for the teleologist.

<sup>26</sup> Feldman comes to a similar conclusion concerning epistemic obligation in general. He argues that epistemic duties will be duties to have certain doxastic attitudes concerning a proposition *for agents who have any doxastic attitude concerning that proposition* (2000 679). That is, if an agent has some belief about proposition P, *then* she might be e-required to believe or disbelieve P, depending on her evidence for or against P. But if an agent has no doxastic attitude towards P, then she'll be under no e-obligations with respect to P.



Now, if the 'ought to believe' clause in TP#4 covers only immediate beliefs, then the teleologist's epistemic strategy won't work, since it's not the case that agents e-ought to immediately believe that remembering promises maximizes expected value. So, the teleologist pursuing the epistemic strategy needs to read the 'ought to believe' clause more widely. In particular, it looks like it will need to cover not just immediate beliefs, but also beliefs like those we discussed above – beliefs an agent could come to immediately have, were she to reflect on them.

Unfortunately, this can't be the proper interpretation of TP#4. To see why, suppose that an evil demon threatens to kill one hundred people unless you solve a difficult math problem in ten minutes. The problem is difficult enough so that it's not the case that you e-ought to be able to solve it, and in fact staring at it for ten minutes is unlikely to help, and you know this. The demon adds, however, that if you admit to him right away that he's stumped you, he will kill only ten people. What should you do? Surely the teleologist will say that you are obliged to admit defeat. You'll pass up a small chance at saving the hundred, but in return you can guarantee saving ninety. (So far, the case is just like *Hurricane*.)

To make things a bit more interesting, though, suppose the demon gives you something of a hint. He provides you with one hundred much simpler math problems. You can solve, and e-ought to be able to solve, any of these problems in a minute. The demon tells you, truthfully, that one of the hundred easy problems is such that if you solve it, you'll immediately see how to solve the difficult problem. (The solutions, perhaps, mirror one another in some obvious way.) Now, intuitively, you should still admit defeat. Your chances of saving the hundred have gone up a bit, but it's still unlikely that you'll hit upon the single easy problem which would let you solve the difficult one. But if the 'ought to believe' clause in TP#4 ranges over the things an agent could, upon reflection, come to immediately believe, TP#4 says you are required not to admit defeat. One of the things you could, upon reflection, come to immediately believe is the solution to the relevant simple math problem. Given that belief, it would then be the case that you could easily solve the difficult problem, and doing so would maximize expected value. So, TP#4 would say that you're required to solve the difficult problem – which is wrong.

The problem here is, once again, clear. Although there is something that you e-ought to be able to figure out that would let you solve the difficult problem, you have no way of knowing what that thing is. Unless you know which simple problem is relevant to the solution of the difficult problem, your situation isn't really much better than if the demon hadn't given you the "hint". The 'ought to believe' clause in TP#4 therefore can't range over all the things an agent e-ought to be able to come to (immediately) believe, but can at most cover those things that an agent both (1) e-ought to be able to come to believe and (2) e-ought to believe would be relevant. The demon's "hint" doesn't change your normative situation unless he gives you a problem

that you (1) e-ought to be able to solve and (2) e-ought to believe would be applicable to the difficult problem.

Okay – let's return to the case of the forgetful promisor. We said it's not the case that agents generally e-ought to believe that remembering promises maximizes expected value. But it is (perhaps) the case that they e-ought to be able to come to have that belief upon reflection. We've now seen, though, that that isn't enough for the teleologist. Even if it's true both that remembering promises maximizes expected value and that an agent, were she to consider that proposition, would be e-required to believe it, that isn't sufficient to include it under the 'ought to believe' clause of TP#4. It will also need to be the case that an agent e-ought to believe that considering that proposition would be relevant. Since 'relevance' here is relevance to the project of value maximization, that means it will need to be the case that the agent e-ought to believe that considering whether remembering promises maximizes expected value itself maximizes expected value. Needless to say, if most agents haven't considered whether *remembering promises maximizes expected value*, they certainly haven't considered whether *considering whether remembering promises maximizes expected value itself maximizes expected value*. That very complicated proposition therefore can't be one that an agent is e-required to immediately believe (though, once again, believing it might perhaps be e-required upon reflection). We could continue to consider further iterations, but given the growing complexity of the propositions such a search seems likely to be fruitless. It doesn't look like the teleologist will be able to give the 'ought' in TP#4 a purely epistemic reading.

In the end, the problem with the epistemic strategy, I think, is that there just aren't enough epistemic duties out there, or at least not enough duties of the right kind. Agents don't have beliefs about many things, even very obvious ones, and it seems to me that it's not the epistemologist's job to tell an agent what things she ought to have beliefs about. Rather, if there are epistemic duties they'll generally be like the ones we've just been discussing: they'll be duties to have certain beliefs *for agents who have considered the matter*.<sup>27</sup> Or they'll be duties *not* to have certain beliefs which are unsupported by evidence. So, you don't have an epistemic duty to have any belief about how many windows are in your house. But you do have a duty to believe there are eighteen windows, if you consider the matter. Or, you have a duty *not* to believe that you have seventeen or nineteen windows. Epistemic duties like this, though, can't do the work the teleologist needs. The teleologist needs to

---

<sup>27</sup> There is a similarity here to fitting attitude (FA) theories of the good. We've been looking at proposals that attempt to link belief with truth. FA theories try to link desire with goodness. (FA theorists propose the opposite order of explanation, saying that something is good because it ought to be desired, but that difference doesn't matter right now.) But very few FA theorists say that agents are required to desire certain things. Rather, they say that there are certain things such that *if an agent were to think about them*, she would be required to desire them (Chisholm 1986 52), or certain things that it would be *fitting* or *appropriate* for an agent to desire. I suspect that truths about what we e-ought to believe work the same way.

say not just that the driver, *if she were to think about it*, ought to believe it would maximize expected value to look before backing up. The teleologist needs to say that the driver ought to have that belief, or at least ought to think about the matter.

It is plausibly true that drivers have a duty to know about the danger cars pose, that pharmacists have a duty to know about drugs' side effects, that hikers have a duty to know where hiking is allowed, and that promisors have a duty to know about the meaning and consequences of a broken promise. But these duties won't be grounded purely in epistemic considerations. Rather, they'll be grounded at least in part practically. Drivers ought to know about pedestrians and promisors ought to know about promises because it is, from a moral perspective, very important that they do so. The 'ought' in TP#4 incorporates, at minimum, a duty of moral conscientiousness and reflection. It therefore can't be purely epistemic.<sup>28</sup>

## V. CAN TP#4 BE JUSTIFIED?

I began this paper by asking what consequentialists could mean when they say that we are required to give more to famine relief, or what teleological deontologists could mean when they say that we're forbidden from killing one to save five. In §III we looked for a teleological principle that could justify such claims. I argued that any plausible teleological principle would have a normative component, requiring an agent to do what maximizes expected value, in light of what she ought to believe. In the last section I argued that this 'ought' was deontic and couldn't be understood epistemically. If those arguments have been correct, it follows that the teleologist doesn't have any way to get rid of the 'ought' in TP#4. Any teleological theory, then, will need to recognize at least some non-teleological, deontic norms.

This strikes me as a significant blow to the teleological project. The theoretical simplicity which came with recognizing only one form of normativity at the foundational level is gone. Nevertheless, we might – especially if we're weak teleologists – be willing to accept this. If something like TP#4 were correct, morality would still be teleological in spirit, or

---

<sup>28</sup> Feldman once again agrees: "It's surely true that there are times when one would be [epistemically] best off finding new evidence. But this always turns on what options one has, what one cares about, and other non-epistemic factors. As I see it, these are prudential or moral matters, not strictly epistemic matters" (2000 689). And: "There are cases in which one can spend one's time gathering evidence about propositions concerning inconsequential and trivial propositions or about more weighty matters. Evidentialism [the view that one always ought follow one's evidence] provides no guidance about what to do... What topics you ought to investigate depend upon what topics are of interest to you, what investigations can help you to make your own life or the lives of others better, and other such matters. Evidentialism is silent on those moral and prudential issues, and I don't see why it should address them" (2000 690, *cf.* 1988 98).

substantially teleological. That is, although there would be some non-teleological norms indicating what an agent ought to believe, or what kinds of reasoning she ought to complete, the rest of morality – and in particular the norms which guide most of our actions – would still be teleological. When figuring out what an agent's duty is, we would first need to determine what she ought to believe, but from that point forward we could simply calculate expected value, as TP#4 directs.

In this section, however, I'll argue that this picture isn't correct. Even if we grant the teleologist whatever non-teleological deontic norms she needs, she still won't be able to use the Core Teleological Intuition to justify anything like TP#4. When agents violate moral requirements, it isn't typically because they've failed to do what would maximize expected value, given what they ought to believe. Instead, their violations are frequently of non-teleological norms. Morality therefore isn't even substantially teleological.

Unlike the earlier teleological principles, TP#4 does seem to issue reasonable verdicts, and I don't intend to offer an argument to the contrary. So what's the problem? In §II.2, I introduced the CTI, which asserted that there is a sense in which it's irrational to do something when a better option is known to be available. I argued that teleology requires a positive defense and that the CTI is the only thing that has been put forward as the foundation of that defense. The STP was so similar to the CTI that, had it been plausible, we might have been compelled to accept that it was directly justified by the CTI. TP#4, however, doesn't have such a clear connection to the CTI. (The CTI, for example, doesn't say anything about what an agent ought to believe.) Therefore, in order to provide a positive argument for teleology, we now have some work to do, connecting TP#4 to the CTI.

Consider someone who is not doing as TP#4 directs. Perhaps in the original *Hurricane* case you calculate the expected value of going to island B by multiplying 100 and 25, instead of dividing them. You therefore conclude that you have an overwhelming obligation to go to island B. The natural thing for the teleologist to say, I think, is: *You're wrong about the expected value – it's 4, and not 2500. If you'd reasoned the way you should have, you'd have seen that the expected value of going to A is much higher than of going to B. You'd have realized that it would be irrational not to go to A. Therefore, your duty really is to go to A!* Similar reasoning might be addressed to the agent who carelessly backs up her car and hits a pedestrian: *You didn't know there was a pedestrian behind your car, but you should have known that there was at least a chance of it. And had you known that, you'd obviously have seen that backing up so quickly was a dumb thing to do. Therefore, you had a duty not to back up so quickly!*

This sounds like a plausible bit of reasoning, I think. It first points out that the agent ought to know something – how to combine 25 and 100, or the general prevalence of pedestrians. That seems to correspond to the non-teleological deontic facts we're now granting the teleologist. Then it points out that if the agent knew that thing, she'd determine a different option was better, and accordingly that it would be irrational to act as she had planned.

This sounds reminiscent of the CTI. Together, they seem to justify the conclusion that the agent is obliged to act as TP#4 directs. Therefore, if this form of argument is valid, we'll have justified TP#4, given non-teleological deontic facts plus something like the CTI.

Despite its intuitive plausibility, however, the argument isn't a good one. We could schematize it like this:

1. You ought to Q. [non-teleological deontic fact]
2. If you Q-ed, you'd see that A is best. [assumption]
3. If you knew that A is best, you'd be (rationally) obliged to A. [CTI]
4. Therefore, you are (rationally) obliged to A. [conclusion of TP#4]

Below I'll present an example which will show that this form of argument isn't valid, but it may be helpful to quickly write out a simplified version using the notation of deontic logic:

1.  $\Box Q$
2.  $Q \rightarrow S$
3.  $S \rightarrow \Box A$

---

4.  $\Box A$ <sup>29</sup>

An argument like that is valid in most systems of modal logic: necessarily-Q entails Q, which will quickly lead to the conclusion. The characteristic feature of deontic logics, however, is their rejection of the axiom  $\Box Q \supset Q$ . From the fact that something ought to be the case, we can't conclude that it is the case, and so this argument isn't valid when the box operator is interpreted deontically.<sup>30</sup>

---

<sup>29</sup> I use ' $\rightarrow$ ' because there are well-known problems incorporating material conditionals into deontic logic. (See e.g. McNamara 2009 §4.5.) I use  $\Box$  to represent what (depending on the argument of §IV) may be different types of normativity. The possible ambiguity shouldn't matter for our purposes.

<sup>30</sup> On some versions of deontic logic, the above argument could be valid. If  $\Box A$  and  $A \supset B$  entail  $\Box B$ , and if  $\Box \Box B \supset \Box B$  is accepted as an axiom, then the argument would work. However, such versions of deontic logic only have plausibility, I think, when the obligation operator takes propositions as arguments. ("It is obligatory *that Jane goes to the store.*") However, for reasons I don't have the space to go into here, it seems to me that the obligation operator should take action-agent pairs as arguments. ("*Jane is obliged to go to the store.*") In any case, the example I describe below should show the problem with arguments like the one here.

Here is an example which should make the point clear.<sup>31</sup> Suppose you've promised to help me in my lab with my scientific research this evening. My experiment requires that the lab be flooded with a highly toxic gas. Fortunately, there is a counteragent available. If you enter the lab without taking the counteragent, you'll die. But if you enter the lab after taking the counteragent, you'll suffer no ill effects. If, however, you take the counteragent and don't enter the lab, you'll unleash a horrible, contagious illness on the world. (The counteragent needs to be neutralized by the gas in the lab.) Clearly, then, if you enter the lab, you should take the counteragent. And if you don't enter the lab you have a duty not to take the counteragent. Since you've promised to help me, you have a duty to enter the lab. Can we conclude that you should take the counteragent? No! You should take the counteragent only if you enter the lab, and the fact that you've promised you'll enter the lab doesn't guarantee that you actually will enter the lab. If you don't enter the lab, you are obliged *not* to take the counteragent. So, if you don't enter the lab, the following will all be true:

1. You are obliged to enter the lab.
2. If you were to enter the lab, you'd be obliged to take the counteragent.
3. You're obliged not to take the counteragent.

The first two claims are essentially the same as the premises in the argument for TP#4, but the conclusion is exactly the opposite of what the teleologist is looking for.

Intuitively, think of it this way: if you fail to help me, I'll chastise you for failing to show up. You had a duty to help, and you didn't fulfill it. But I surely won't chastise you a second time for failing to take the counteragent. In fact, I'd be relieved that you didn't take it. Morally speaking, things would have been terrible had you done so. In failing to show up, then, you're guilty of only one error, failing to keep your promise. You're not *also* guilty of a second error, failing to take the counteragent. From the fact that you are obliged to help, it doesn't follow that you are obliged to take the counteragent. That only follows if you do in fact help.

What this means is that the argument the teleologist needs in order to justify TP#4 is invalid. TP#4 says that an agent is required to do what would maximize value, given what she ought to believe. But the claims (1) that an agent ought to believe that there is liable to be a pedestrian behind her car and (2) that given such a belief, it would maximize expected value for her to look before backing up, don't entail (3) that she is required to look before

---

<sup>31</sup> This example has the same structure as Chisholm's (1963) famous example of the man who ought to go the assistance of his neighbors, but ought to call only if he in fact goes. I use the example, however, for a different purpose. Whereas Chisholm's aim was to show that existing formalizations of deontic logic are inadequate, my aim is the much more modest one of showing that an informal pattern of inference which might *prima facie* seem plausible isn't a good one.

backing up, as TP#4 says. Reaching *that* conclusion would require the argument we saw was invalid. Similarly, we can grant (1) that you ought to remember that you promised to feed your neighbor's cats, and (2) that if you remembered the promise it would maximize expected value for you to feed the cats. But it doesn't follow (3) that you, having forgotten, have a duty to feed the cats. Of course I think that those conclusions are true – you are obliged to look before backing up and to feed the cats – but we can't reach those conclusions from our non-teleological deontic starting points plus the CTI. Which is to say: we can't justify TP#4 by an appeal to the CTI.

## VI. WHAT REMAINS FOR THE TELEOLOGIST?

The argument, then, which is apparently needed if we are to use the CTI to justify TP#4, is invalid. In order to hold on to TP#4 and accordingly to maintain that morality is substantially teleological, the teleologist has three options: (1) to reject the need for a positive defense of teleology, (2) to find an alternate justification for TP#4 using the CTI, or (3) to find a justification for TP#4 based on something other than the CTI. I've suggested (1) is implausible, I don't know how (2) would go, and I don't even know where (3) would start. In this last section, I'll first respond to one objection, which will give me a chance to comment on what kind of teleological principle the CTI might support, and then I'll conclude by offering two positive suggestions for those attracted to a teleological outlook.

### 1. *A minimal teleological principle?*

I can imagine a teleologist responding: even if the argument of the last section is right, its application isn't very wide. It may be true that the person who divides 100 by 25 and gets 2500 is guilty only of violating a non-teleological norm, but you, Andrew, frequently aren't such a person. You know that your thousand dollar bonus could be doing much more good if it were sent to Oxfam, so the argument you've given doesn't apply in your case. *You*, therefore, are obligated to send your bonus to Oxfam, instead of putting it towards your mortgage, and the explanation for that obligation is substantially teleological.

This objection isn't right, though. When I don't send every extra thousand dollars I earn to Oxfam, it's (at least partly) because I believe that certain values are incomparable. Or it's because I believe that certain values are agent-relative, whereas the consequentialist believes they're agent-neutral. Now, if consequentialism provides the right story about value, then I'm mistaken. These beliefs are false. And, perhaps, we might even grant to the consequentialist that I ought to have the correct beliefs – I ought to believe that values are comparable and agent-neutral. But if that's right, then the argument I've given *does* apply. It's true that I ought to have certain beliefs, and it's true that if I had those beliefs, then I'd see that sending the money to

Oxfam would be best and so would be rationally required to donate my bonus. But the argument of the last section showed that  $\Box P$  and  $P \rightarrow \Box Q$  don't entail  $\Box Q$ . So long, then, as I in fact believe what I ought not – that values are incomparable or that they're agent-relative – we can't conclude that I am required to donate the bonus to Oxfam. I am indeed making a moral mistake, but my error isn't in failing to send the money; instead, it's in believing that values are incomparable. The norm I violate in having that belief isn't a teleological one. Its foundation has nothing to do with the CTI or goodness. The norm I violate is one of the non-teleological norms that I argued the teleologist needs to take for granted. The argument of the last section therefore is widely applicable. Merely believing, for example, that self-regarding and other-regarding values are incomparable – even if that belief is false and even if it's a belief one ought not have – is enough to preclude the application of a teleological principle.

If this is right, it means that in at least a large number of everyday cases, agents who violate moral requirements apparently don't do anything that violates a teleological requirement. What does that mean for the CTI? Can it justify any teleological principle? Are there *any* cases in which agents violate an obligation grounded in the CTI? I think that these reflections can point us towards a better understanding of the CTI and, accordingly, to what kind of teleological principle it might support.

We saw earlier that the CTI purports to make teleology a requirement of rationality. But there is clearly nothing irrational about failing to do what is in fact best. If you have no idea where the 100 are, there's nothing irrational about failing to rescue them. Similarly, there's nothing irrational about buying Name Brand paint unless you have good reason to think that Store Brand is better. All of this suggests that if the CTI can justify any teleological principle, it will be one like this:

*Minimal Teleological Principle (MTP).* An agent is required to do what she believes would maximize expected value.<sup>32</sup>

The MTP is, as its name says, quite minimal. It wouldn't support anything like the obligations most teleologists discuss. Is the MTP justifiable, though? Perhaps, but I have my doubts. Many philosophers (and more social scientists) believe that an agent *will* do what she thinks best. The fact that an agent intentionally does something is taken to be (nearly) conclusive evidence that she thought that action best. If I choose to do A, then I must either think that A is best, all-things-considered, or at least best on some scale which I take to be incomparable with other applicable scales. If this is right, an agent can't violate the MTP. A normative principle an agent can't violate,

---

<sup>32</sup> Perhaps more accurately: *if* an agent believes of some action that it would maximize expected value, then she is rationally required to do it. (In most situations, I suspect we act without any beliefs about what would maximize expected value.)



though, is no normative principle at all.

So, in order to say that something like the MTP is a normative principle, we'll need to allow for the possibility of clear-eyed *akrasia*, saying that an agent can intentionally do something other than what she thinks best. But once we say this, it becomes less clear why an agent is obligated to – why she *must* – do what she thinks best. If it seems to me that I can choose some option that I know to be sub-optimal, and if I can in fact choose it, then in what sense must I choose the option I believe to be best? If clear-eyed *akrasia* is a possibility open to me, then in what sense *must* I not avail myself of that option?<sup>33</sup> This is the question the proponent of the MTP needs to answer. If it can be answered, then there will be one kind of deontic norm, given by the MTP, which holds in virtue of *beliefs* about evaluative norms. That's the closest, I think, the evaluative could come to grounding the deontic.

## 2. *Two ways forward*

If the arguments I've given are correct, they count against most consequentialist theories, as well as a number of deontological and egoistic ones. Do I think, then, that in a handful of pages I've refuted an entire branch of moral philosophy? In particular, do I think I've put a (another?) final nail in the coffin of consequentialism? Of course not. Unlike most objections to consequentialism (or to teleology more broadly), I haven't tried to argue that the substantive conclusions the consequentialist reaches are implausible. I've merely argued that they can't be justified by a teleological principle. That leaves, I think, two important options open to a consequentialist. (In what follows, I'll speak in terms of consequentialism, since the examples I'll draw on will be from the consequentialist literature. But everything I say could also be applied to a non-consequentialist version of teleology.) Let me conclude by briefly laying out these two options. I think they've both received far less attention than they deserve, and I hope that one effect of this paper will be to encourage their further discussion and development.

The first possibility I'll call *non-teleological consequentialism*. I've argued that teleology can't be justified by anything like the CTI. So, even if the consequentialist is right that (say) desire satisfaction is the only thing that's intrinsically good, we can't conclude that because an action would satisfy the

---

<sup>33</sup> Notice that this is not the typical objection that “there is a real danger that in attempting to make causal and conceptual space for full-fledged akratic action one might commit oneself to the rejection of genuine ties between evaluative judgment and action” (Mele 1991 34). In order to defend the MTP, we would need an argument not merely establishing a link between evaluative judgment and action; we would need for that link to be deontic. It would not be enough to show, for example, that agents are rationally criticizable or defective for failing to act in accordance with their evaluative judgments. Rather, the defender of the MTP would need to show that agents *must* act in accordance with those judgments. This is a stronger claim. (See my discussion of scalar consequentialism, below.)

most desire and hence bring about the most good, it is required. But it may still be true that an agent is required to bring about maximal satisfaction of desire *for some other reason*. What's maximally good may still be required, for all I've said, so long as it is required for some reason other than the fact that it's maximally good.

This may sound like a strange proposal, but it actually has several precedents in the consequentialist literature. Harsanyi, for example, believes that justice demands society be organized in whatever way would be agreed to by the parties to a Rawlsian original position (1953, 1975). That is, he agrees with Rawls that deontic norms are grounded in the idea of a hypothetical contract. This, then, is clearly *not* a teleological foundation for deontic norms. Unlike Rawls, however, Harsanyi thinks that the parties to the original position would agree to arrange society so as to maximize average utility. Harsanyi, then, believes that we're in fact required to maximize average utility, but he grounds that requirement in a non-teleological, contractualist way.<sup>34</sup>

Paley's (1785) theological utilitarianism is also non-teleological. For Paley, happiness and unhappiness are good and bad, and this is a fact about the world. Requirements, on the other hand, are facts about what God has commanded. Paley thinks that God in fact commands that we maximize utility, and so we're required to do so. But the requirement holds in virtue of God's command – not in virtue of evaluative facts. Paley's deontic facts are therefore, like Harsanyi's, non-teleological. Other non-teleological versions of consequentialism can be found in Kymlicka, who argues that utilitarianism might be justified “because it [treats] people as equal, with equal concern and respect” (1988 177) and Cummiskey, who tries to derive consequentialism from Kant's Categorical Imperative (1996).

So much for the first way forward for teleologists. The second option, which following Slote I'll call *scalar consequentialism*, is perhaps a bit more radical, but I think is also more interesting. The arguments I've given here have left many characteristically consequentialist claims untouched. For example, nothing I've said counts against the possibility that all value is agent-neutral, that all that matter morally-speaking are pleasure and pain, that a moral saint would maximize utility, that there is no meaningful distinction between doing and allowing, that any amount of pain for me could be morally offset by some amount of pleasure for you, that giving two people

---

<sup>34</sup> Rawls, in fact, argues that any version of utilitarianism which maximizes average, as opposed to total, utility, should be non-teleological: “By choosing [the principle of average utility], the parties [to the original position] maximize their expected well-being as seen from this point of view. Some form of contract theory provides, then, a way of arguing for the average over the classical [total] view. In fact, how else is the average principle to be accounted for? After all, it is not a teleological doctrine, strictly speaking, as the classical view is, and therefore it lacks some of the intuitive appeal of the idea of maximizing the good. Presumably one who held the average principle would want to invoke the contract theory at least to this extent” (1999 143). I thank Tim Scanlon for bringing this passage to my attention.

some amount of pain is morally on a par with giving one person twice as much pain, and so forth. In fact, the only characteristically consequentialist claim I've cast doubt on is the claim that agents are *required* to maximize value.

Scalar consequentialism, which was first explicitly discussed by Slote (1985, 1989) and has since been advocated by Norcross (2006a, 2006b) and Howard-Snyder (1993, 1994), gives up this last claim, while holding onto the earlier ones.<sup>35</sup> It proposes consequentialism as a purely evaluative theory, with no deontic component. As Norcross (2006b 47) points out, in addition to all the claims I noted in the last paragraph, the scalar consequentialist can also say that an agent has reason to do an action to the extent that it promotes the good, that agents are virtuous to the extent that they promote the good, and so forth. To say that an agent is *required to* or *must* do what's best, though, is on this view just a superstition, brought to consequentialism by the pernicious influence of deontologists and an overly-legalistic model of what morality must look like (Norcross 2006b 43). We can tell an indecisive agent that she has most reason to do what maximizes value and that a good person would do so. But if she asks whether she truly *must* do the optimal action, as opposed to some slightly less good alternative, we should concede that of course there's no interesting sense in which she *must* do it. The difference between the optimal action and a slightly sub-optimal one is of no more moral significance than the difference between a sub-optimal action and a slightly more sub-optimal action.

## WORKS CITED

- Björnsson, Gunnar and Stephen Finlay (*forthcoming*). "Metaethical Contextualism Defended," forthcoming in *Ethics*.
- Broome, John (1991). *Weighing Goods*. Oxford: Blackwell.
- Chisholm, Roderick (1963). "Contrary-to-Duty Imperatives and Deontic Logic." *Analysis* 24: 33-36.
- Chisholm, Roderick (1986). *Brentano and Intrinsic Value*. Cambridge: Cambridge University Press.
- Cummiskey, David (1996). *Kantian Consequentialism*. Oxford: Oxford University Press.
- Dowell, Janice L. (*forthcoming*). "Flexible Contextualism about 'Ought': Deliberation and Belief-ascriptions," forthcoming in *Ethics*.
- Dreier, James (1993). "The Structure of Normative Theories," *Monist* 76: 22-40.
- Dreier, James (2011). "In Defense of Consequentializing," OSNE VOL. 1
- Dworkin, Ronald (1984). "Rights as Trumps," in *Theories of Rights* (ed. Waldron). Oxford: Oxford University Press.
- Feldman, Richard (1988). "Having Evidence," in *Philosophical Analysis* (ed. Austin). Dordrecht: Kluwer: 83-104.
- Feldman, Richard (2000). "The Ethics of Belief," *Philosophy and Phenomenological Research* 60 (3): 667-695.
- Freeman, Samuel (1994). "Utilitarianism, Deontology, and the Priority of Right," *Philosophy and Public Affairs* 23 (4): 313-349.
- Foot, Philippa (1985). "Utilitarianism and the Virtues," *Mind* 94: 196-209.
- Guerrero, Alexander A. (2007). "Don't Know, Don't Kill: Moral Ignorance, Culpability, and Caution," *Philosophical Studies* 136: 59-97.
- Haines, William (2008). "Consequentialism," *Internet Encyclopedia of Philosophy*. Available at <http://>

---

<sup>35</sup> Slote (who has since rejected scalar consequentialism in favor of a virtue-theoretic approach) was the first *modern* scalar consequentialist, but I think that Bentham's and Mill's utilitarianisms can also be interpreted without undue strain in a scalar way.

www.iep.utm.edu/c/conseque.htm.

- Hare, R.M. (1981). *Moral Thinking: Its Levels, Methods, and Point*. Oxford: Oxford University Press.
- Harsanyi, John (1953). "Cardinal Utility in Welfare Economics in the Theory of Risk-taking," *The Journal of Political Economy* 61 (5): 434-435.
- Harsanyi, John (1975). "Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls's Theory," *The American Political Science Review* 69 (2): 594-606.
- Howard-Snyder, Frances and Alastair Norcross (1993). "A Consequentialist Case for Rejecting the Right," *The Journal of Philosophical Research* 18: 109-125.
- Howard-Snyder, Frances (1994). "The Heart of Consequentialism," *Philosophical Studies* 76 (1): 107-29.
- Jackson, Frank (1991). "Decision-Theoretic Consequentialism and the Nearest and Dearest Objection," *Ethics* 101 (3): 461-482.
- Kagan, Shelly (1989). *The Limits of Morality*. Oxford: Clarendon Press.
- Kagan, Shelly (1998). *Normative Ethics*. Boulder, CO: Westview Press.
- Kymlicka, William (1988). "Rawls on Teleology and Deontology," *Philosophy and Public Affairs* 17 (3): 173-90.
- Lawlor, Rob (2009). "The Rejection of Scalar Consequentialism," *Utilitas* 21(1): 100-116.
- Louise, Jennie (2004). "Relativity of Value and the Consequentialist Umbrella," *The Philosophical Quarterly* 54 (217): 518-536.
- Mack, Eric (1998). "Deontic Restrictions are Not Agent-Relative Restrictions," *Social Philosophy & Policy* 15(2): 61-83.
- McNamara, Paul (2009). "Deontic Logic," *Stanford Encyclopedia of Philosophy* (ed. Zalta). Available at <<http://plato.stanford.edu/archives/spr2009/entries/logic-deontic/>>.
- Mele, Alfred (1991). "Akratic Action and the Practical Role of Better Judgment," *Pacific Philosophical Quarterly* 72: 33-47.
- Moore, G.E. (1903/1971). *Principia Ethica*. Cambridge: Cambridge University Press.
- Nagel, Thomas (1989). *The View From Nowhere*. Oxford: Oxford University Press.
- Norcross, Alastair (2006a). "The Scalar Approach to Utilitarianism," in *Blackwell Guide to Mill's Utilitarianism* (ed. West). Oxford: Blackwell: 217-232.
- Norcross, Alastair (2006b). "Reasons Without Demands: Rethinking Rightness," in *Contemporary Debates in Moral Theory* (ed. Dreier). Oxford: Blackwell: 38-54.
- Oddie, Graham and Peter Menzies (1992). "An Objectivist's Guide to Subjective Value," *Ethics* 102 (3): 512-533.
- Paley, William (1785/2002). *The Principles of Moral and Political Philosophy*. Indianapolis: Liberty Fund.
- Parfit, Derek (forthcoming). *On What Matters*. Unpublished manuscript.
- Parfit, Derek (1986). *Reasons and Persons*. Oxford: Clarendon Press.
- Pettit, Philip (1991). "Consequentialism," in *A Companion to Ethics* (ed. Singer). London: Blackwell: 230-237.
- Portmore, Douglas W. (2000). "Commonsense Morality and Not Being Required to Maximize the Overall Good," *Philosophical Studies* 100: 193-213.
- Portmore, Douglas W. (2003). "Position-Relative Consequentialism, Agent-Centered Options, and Supererogation," *Ethics* 113: 303-332.
- Rawls, John (1999). *A Theory of Justice (revised ed.)*. Cambridge, MA: Harvard University Press.
- Rosen, Gideon (2003). "Culpability and Ignorance," *Proceedings of the Aristotelian Society* 103: 61-84.
- Scheffler, Samuel (1985). "Agent-Centered Restrictions, Rationality, and the Virtues," *Mind* 94: 409-419.
- Schroeder, Mark (2007). "Teleology, Agent-Relative Value, and 'Good,'" *Ethics* 117: 265-295.
- Schroeder, S. Andrew (2009). *Divorcing the Good and the Right*. Ph.D. dissertation, Harvard University.
- Sen, Amartya (1982). "Rights and Agency," *Philosophy and Public Affairs* 11: 3-39.
- Shaw, William (2006). "The Consequentialist Perspective," in *Contemporary Debates in Moral Theory* (ed. Dreier). Oxford: Blackwell: 5-20.
- Sidgwick, Henry (1907/1981). *The Methods of Ethics, 7th ed.* Indianapolis: Hackett.
- Sinnott-Armstrong, Walter (2009). "How strong is this obligation? An argument for consequentialism from concomitant variation," *Analysis* 69 (3): 438-442.
- Slote, Michael (1985). *Common-sense Morality and Consequentialism*. London: Routledge & Kegan Paul.
- Slote, Michael (1989). *Beyond Optimizing: A Study of Rational Choice*. Cambridge, MA: Harvard University Press.
- Smart, J.J.C. (1973). "An Outline of a System of Utilitarian Ethics," in *Utilitarianism For and Against*. Cambridge: Cambridge University Press: 3-74.
- Smith, Holly (2011). "The Moral Clout of Reasonable Beliefs," OSNE VOL. 1
- Smith, Michael (2003). "Neutral and Relative Value after Moore," *Ethics* 113: 576-598.
- Smith, Michael (2005). "Meta-ethics," in *The Oxford Handbook of Contemporary Philosophy* (eds. Jackson and Smith). Oxford: Oxford University Press, 3-30.
- Zimmerman, Michael (2008). *Living with Uncertainty*. Cambridge: Cambridge University Press.