

THE RATCHETING-UP EFFECT

Vanessa Carbonell
University of Cincinnati

This is the peer reviewed version of the following article: "The Ratcheting-Up Effect", Pacific Philosophical Quarterly 93(2), 2012, 228-254, which has been published in final form at <https://doi.org/10.1111/j.1468-0114.2012.01425.x>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions. This article may not be enhanced, enriched or otherwise transformed into a derivative work, without express permission from Wiley or by statutory rights under applicable legislation. Copyright notices must not be removed, obscured or modified. The article must be linked to Wiley's version of record on Wiley Online Library and any embedding, framing or otherwise making available the article or pages thereof by third parties from platforms, services and websites other than Wiley Online Library must be prohibited.

ABSTRACT. I argue for the existence of a 'ratcheting-up effect': the behavior of moral saints serves to increase the level of moral obligation the rest of us face. What we are morally obligated to do is constrained by what it would be reasonable for us to believe we are morally obligated to do. Moral saints provide us with a special kind of evidence that bears on what we can reasonably believe about our obligations. They do this by modeling the level of sacrifice a person can realistically bear. Exposure to moral saints thus 'ratchets-up' our obligations by combating a type of ignorance that would otherwise defeat those obligations.

Suppose there are people who deserve to be called 'moral saints,' and these people do not have any magical powers or disturbing pathologies. That is, suppose there are otherwise normal human beings whose lives are marked by extraordinary moral accomplishments: they do far more than we think morality requires of them, they exhibit resolve and tenacity when others would give up, and they bear heavy burdens of personal sacrifice. What does the existence of such people mean for the rest of us? Do moral saints serve a purpose other than to make us feel bad about ourselves?

In this paper I argue that the behavior of moral saints serves to increase the level of moral obligation the rest of us face. They do this by providing us with crucial information. What we are morally obligated to do, I argue, depends on what it would be reasonable for us to believe we are morally obligated to do. Moral saints provide us with a special kind of evidence that bears on what we can reasonably believe about our obligations. They do this by modeling for us the level of sacrifice a person can realistically bear. Exposure to moral saints thus 'ratchets-up' our obligations by combating a type of ignorance that would otherwise defeat those obligations.

My argument proceeds in three steps. After clarifying what it means to be a moral saint and giving some examples, I introduce a puzzle about moral saints and obligation. With this puzzle in mind, I then give an account of sacrifice as it relates to moral obligation. Finally, I argue for the existence of the Ratcheting-Up Effect.

1. PRELIMINARIES ABOUT MORAL SAINTS

Let me clarify exactly what I mean by a ‘moral saint.’ I do not mean to refer to a person who has a special religious significance, or whose moral goodness is connected to religious notions of piety, purity, or divinity.¹ Nor do I mean to refer to a person who commits an isolated heroic action or a single large gesture, though these actions are sometimes described as ‘saintly.’² Finally, the moral saints I am interested in are not idealizations. They are not maximizers or obsessive perfectionists, and they are not morally flawless.³ Rather, they are human beings with more-or-less ordinary psychology who have devoted their lives to a moral project, who consistently perform actions that are: (1) good but not required; (2) morally significant⁴; (3) undertaken at some personal cost; and (4) not outweighed by other morally bad or blameworthy actions. These are not meant as necessary and sufficient conditions, but as features of a paradigm case. To make this sort of person more vivid, consider two living examples.

PAUL FARMER. Paul Farmer is doctor who founded Partners in Health, an organization that runs medical clinics serving the world’s poorest and sickest people. Farmer’s clinics treat tuberculosis and HIV regardless of the patients’ ability to pay, and they do it for a fraction of what it costs in developed countries. What began as a single clinic in Haiti is now a worldwide operation with an annual budget in the tens of millions of dollars. Farmer himself treats countless patients directly, sometimes hiking for hours to make house calls. In the early years of the organization, Farmer alternated between living in a hut next to his clinic in Haiti and sleeping in the basement of his office in Boston. He had his entire salary deposited directly into the organization’s budget, leaving the office staff to make sure his modest bills were paid. He didn’t buy new clothes or take

¹ See, e.g., Robert Adams (1984). For Adams, ‘real’ saints like Mother Teresa or Ghandi are ‘people in whom the divine can be seen’ (398), and philosophers who take such people as examples of *moral* saints are misguided. I avoid examples of ‘real’ saints for precisely the reason that such examples encourage us to conflate moral and religious notions.

² See Urmson (1958) for the classic account of supererogatory actions. Urmson calls an agent who performs a single supererogatory action as a ‘saint’ or ‘hero,’ though he also discusses agents who have longer-term moral commitments.

³ See Wolf (1982) for an account that conceives of moral saints as needing to be flawless, obsessive, and perfectionist. Wolf argues that moral saints are not fit to be personal ideals precisely because they must have these characteristics.

⁴ This constraint is necessary because some supererogatory (i.e., good but not required) actions are not morally significant. Michael Stocker (1968) offers a list of supererogatory actions that includes ‘buying an ice cream cone, on a hot day, for a child one does not know’ (56). With Stocker, I shall assume that not all supererogatory actions are saintly, and that there is more to being a moral saint than just many supererogatory actions strung together.

vacations, and he went for long periods without seeing his wife and daughter. Despite all this, Paul Farmer maintains an upbeat disposition marked by a quirky, dark sense of humor, and he constantly feels that he should be doing more.⁵

SUSAN TOM. Susan Tom is a single mother from California who has adopted over a dozen children in addition to the two she had with her ex-husband. Tom adopts children few other parents would be willing to take. All of them suffer from some degree of illness or disability: one is mentally retarded from fetal alcohol syndrome; one is horribly scarred from a crib fire in infancy; two were born without legs; one is wheelchair-bound from spina bifida; another has a rare, fatal disease that causes his skin and cartilage to deteriorate. While Tom may have been motivated in part by loneliness, her devotion to the children is remarkably selfless. She has no life savings, no retirement account, no time to work outside the home, and no time for a personal life. More significantly, she attends to the daily needs of her children with compassion, resolve, and good humor, and she gives them as normal and joyful a life as possible. She cooks, cleans, shops, and does laundry for this massive household with only her daughter as a helper. She also acts as their nurse, managing the children's medicines and administering painful but necessary disinfecting baths to Anthony, whose skin disease requires special care. Despite her exhaustion, and despite facing the inevitable death of her more seriously ill children, Tom creates a household filled with laughter and play. And like Paul Farmer, Tom seems always to be thinking that she should be doing more.⁶

With these examples in mind, my aim in the sections that follow is to better understand what is going on when people like Farmer and Tom hold themselves to such high standards, and what normative consequences their behavior might have for the rest of us. I begin with a puzzle.

2. A PUZZLE ABOUT AGENTS AND OBSERVERS

Paul Farmer gets called a 'saint' quite frequently, but it is not an assessment he agrees with. 'I don't care how often people say, "You're a saint." It's not that I mind it. It's that it's inaccurate. ... People call me a saint and I think, I have to work harder. Because a saint would be a great thing to be' (Kidder 2003, 16). Perhaps Farmer just has a stricter notion of 'saint' in mind than the one I have proposed. Still, it is clear that many of the actions his admirers consider saintly are ones that he himself does not regard as optional: to him, treating the sick is simply a moral obligation, even if it requires a hike through the mountains.

Similar sentiments can be found in accounts of Holocaust rescuers. Take John Weidner, an agent for the Netherlands Security Service who worked for the Resistance, helping Jews escape during World War II. Weidner put himself in grave danger, endured severe beatings, and narrowly escaped execution after being captured by the Gestapo. When asked whether what he did was 'an extraordinarily good deed,' he replied 'No. Absolutely not. I did my duty. That is all' (Monroe 2003,

⁵ See Kidder (2003) for more on Paul Farmer and Partners in Health. See Carbonell (2009) for more on how Farmer stacks up against Wolf's (1982) account of moral saints.

⁶ See Karsh (2003) for more on Susan Tom and her family.

117). Statements like this are ubiquitous in the literature on Holocaust rescuers.⁷ Indeed, when one examines accounts of historical or contemporary moral saints, it is striking how frequently one encounters claims like ‘I was just doing my duty,’ or ‘I would have felt guilty if I hadn’t done it,’ or ‘I’m not special. Anyone else in my position would have done the same thing’.⁸

These statements have an odd ring to them. When a firefighter runs into a burning building, we are not surprised to hear him say, ‘I was only doing my duty’. After all, it *is* the firefighter’s duty, though it may sometimes be a false modesty that prompts him to say so out loud, and though we may still be awed by his bravery even though he was specially trained for dangerous situations. But when an *ordinary* person with no special training or contractual responsibility runs into a burning building, the statement ‘I was only doing my duty’ means something different—it’s no longer trivially true, and in fact there is good reason to think it is not true at all. In some cases, of course, the ordinary rescuer is just being polite and modest, or being insincere, feigning modesty. However, in other cases she may honestly believe that she was bound by some moral obligation to risk her life, while observers think otherwise. These are the cases that pose an interesting puzzle for us—cases of a sincere disagreement between agents and observers about whether an action was required or optional. Let’s call this the ‘Puzzling Data’:

PUZZLING DATA: Moral saints habitually perform actions that are, intuitively, beyond moral obligation. Yet the moral saints sometimes consider these actions to be obligatory, not supererogatory. There is a *persistent agent-observer disparity*.

How might we resolve this puzzle? On the one hand, these moral saints (and other occasional heroes and rescuers) might simply be mistaken: their actions are in fact not required, and morality demands less of them than they thought.⁹ On the other hand, we observers might simply be mistaken, and morality requires more of us than we thought, including sometimes putting our lives in danger or setting aside our personal projects. As I will argue in what follows, perhaps things are more complicated than these two possible explanations suggest. While it may be true that the saints

⁷ See also Fogelman (1994), Gilbert (2003), and Gies (1987). Miep Gies, who helped hide Anne Frank’s family, said in her memoir, ‘I am not a hero...I was only willing to do what was asked of me and what seemed necessary at the time’ (11).

⁸ A related phenomenon is described by Gregory Trianosky (1986): ‘Sometimes we are challenged to perform acts that are good to do but not required, by individuals who plainly are already committed to performing them’ (27). Trianosky notes that when challenged to do something like join the Peace Corps, we don’t simply decline—we provide an *excuse*. Yet excuses are commonly thought only to be appropriate when we have omitted an *obligatory* act.

⁹ Urmson (1958) takes this position, at least with respect to certain instances of the agent-observer disparity. In considering the case of the paradigm hero—the soldier who smothers a grenade to save his comrades—Urmson claims, ‘[...] if he were to survive only a modesty so excessive as to appear false could make him say, ‘I only did my duty,’ for we know, and he knows, that he has done more than duty requires. Further, though he might say to himself that so to act was a duty, he could not say so even beforehand to anyone else, and no one else could ever say it. Subjectively, we may say, at the time of action, the deed presented itself as a duty, but it was not a duty’ (203).

are mistaken about what morality demands of them, they are not *simply* mistaken. They have a perspective on the action that is not readily available to observers, and we should take their perspective seriously. One of the reasons for the agent-observer disparity is that agents and observers have asymmetric access to facts about *sacrifice*, and moral obligation depends partly on what we can reasonably believe about sacrifice.¹⁰ While we may still conclude that Paul Farmer and Susan Tom are doing more than morality requires, we should remain open to the possibility that they know something we don't know. To make sense of the 'Puzzling Data,' we need to know more about what sacrifice amounts to.

3. SACRIFICE AND MORAL OBLIGATION

The boundary between the obligatory and the supererogatory is sensitive to facts about sacrifice because, just as it is implausible that something one *could not do* could be morally required (obligated implies can¹¹), it is also implausible that something one *could not do without certain sacrifices* could be morally required if those sacrifices are significant relative to what is at stake (obligated implies no unreasonable sacrifice¹²). What we need is an account that explains how a given sacrifice could rise to the level of generating a legitimate claim—that is, a claim to be absolved of a given moral obligation. Is *feeling* something *as a sacrifice* enough to make it so? Surely not, for then agents who give up outlandish luxuries could claim these losses as *sacrifices* so long as giving up those luxuries caused them sufficient distress. Do certain losses count as sacrifices regardless of whether

¹⁰ Alternatively, one might think that it is blameworthiness, and not obligation itself, that is conditioned on what it would be reasonable to believe. I address this worry in section 4.

¹¹ I'm avoiding the phrase 'ought implies can' here because 'ought' is frequently used with a broad meaning that is ambiguous between the ought of obligation and the ought of supererogation. For instance, some might think it makes sense to say that you ought, morally, to perform an act even if it is not obligatory or required. Since I'm interested in precisely the distinction between obligation and supererogation, I avoid the term 'ought' for the sake of clarity. It's also worth noting that the 'can' in 'ought implies can' might be weaker than the 'can' in 'obligated implies can'. As Stephen Darwall (2006) argues, there is a weak sense of 'ought implies can' found in Kant, where to say that an agent 'can' do something is just to say that it is 'an open deliberative alternative, that is, something such that one's abilities and opportunities with respect to it do not preclude intelligible consideration of whether to do it' (240). This weak sense of 'can' is consistent with the agent not even knowing or being able to know that she ought to perform the action. However, if we are interested in *obligation* as something that members of a moral community can authoritatively demand of one another, and hold each other responsible for, then we may need a stronger sense of 'can.' According to Darwall, the very idea of blaming someone or holding her responsible requires that 'we presuppose that she must have been in a position to know [what she should have done]' (241). So for us to hold someone accountable for an action, not only must the action be an 'open deliberative alternative,' but also there must be a 'process of reasoning' by which the agent could come to see herself as obligated. In a later section, I argue for what is perhaps an even *stronger* sense of 'can,' in the form of a 'knowledge constraint' on obligation: if an action is obligatory, then it must be the case that it would be reasonable to believe (or a reasonable person would believe) that it is obligatory.

¹² For a sustained critique of this principle, see Kagan (1989).

the agent feels them as such? If so, then we must be willing to accept that an agent need not be *aware* of a given cost for that cost to count as a sacrifice.

In everyday thought and language, we use the term ‘sacrifice’ rather flexibly, perhaps even sloppily. Sacrifices are sometimes understood *subjectively*, as something the agent *feels* as a loss, and sometimes *objectively*, as something we think it is genuinely bad for the agent to have given up, whether she thinks so or not. Sometimes we use the term ‘sacrifice’ purely *descriptively*, just to report the fact that something has been given up, and sometimes we use it *normatively*, that is, only when we judge that what is given up is valuable or significant in some sense. My task here is not to give an analysis of a generic broad concept that underlies each and every use of this tricky term, but rather to isolate a particular, coherent concept of sacrifice that is properly suited to thinking about moral obligation. I begin with the basic thought that moral obligation seems to be a matter of what we can legitimately demand of one another. Just as there is a limit to what we can demand of others morally, there is a limit to which of our own needs, wants, and preferences we can claim to count as defeaters of our own moral obligations. For instance, just as you cannot demand that I risk my life in order to save you from losing a tooth, I cannot demand that you take my love of milkshakes so seriously as to release me from the obligation to save your life when doing so would mean giving up my milkshake. Whether the demands we make of one another are legitimate and reasonable depends, in part, on whether they are reciprocal and fair. Since moral demands often involve sacrifices, our notion of *sacrifice* must be sufficiently objective that we can understandably ask whether sacrifices meet standards of reciprocity and fairness. In order to be able to make moral demands that others will take seriously—demands that will sometimes involve asking others to give up things that matter to them—I need to be willing to limit the things that I put forth as defeaters of my *own* obligations to just those things whose importance I can justify to others. Thus as mutually accountable members of the moral community, it seems that all of us need to make use of a shared notion of sacrifice, one that is publicly understandable.

This approach to sacrifice borrows somewhat from T.M. Scanlon’s account in ‘Preference and Urgency’ (1975). Scanlon argues for a way of understanding benefits and sacrifices that incorporates justifiability to others. He claims that the ‘urgency’ of moral claims is best understood against the background of an objective criterion of wellbeing. On a subjective criterion, we evaluate well-being from the ‘point of view of the person’s tastes and interests’ (656), whereas on an objective criterion, we evaluate well-being independently of those tastes and interests, so that ‘an appraisal could be correct even though it conflicted with the preferences of the individual in question’ (658). Like the subjective view, the objective view can account for the importance of individual tastes and interests by declaring it objectively important that social institutions be set up in such a way as to allow people to develop such tastes and interests (658). Unlike the subjective view, though, the objective view does not take the strength of those tastes and interests to be the *primary* locus of well-being; on this view, what is good for a person depends not just on what she *does* take an interest in, but also on what it would be good for her to take an interest in.

The objective view also captures the fact that a criterion of well-being which is going to serve as the basis for claims people make against each other must represent a kind of ‘consensus’ about what matters. A subjective view can represent a consensus only in a weak sense: the consensus

that everyone's individual tastes are valued and should count equally, whatever they happen to be (657). The objective view, however, seems to capture a more robust type of consensus: it requires that individual tastes and preferences be at least somewhat *intelligible* to others, and it accords them value *in virtue of* that intelligibility. Thus, when evaluating the 'urgency' of unfamiliar goods, 'we can understand why they are of value to someone else if we can bring the reasons for their desirability under familiar general categories' (660). The sorts of categories that Scanlon has in mind are things like 'material comfort, status, or security' and 'health or protection against injury' (661). Insofar as we can interpret a given preference or interest as falling under one of the general categories, we can understand it, and this seems to make it more viable as a sort of currency in interpersonal exchanges.

Following Scanlon, I think it makes sense to impose what we can call an 'intelligibility constraint' on the sorts of losses that could count as a sacrifice. The harder it is for the rest of us to find your attachment to miniature toy soldiers *intelligible* as a member of one of the claim-generating 'general categories,' the less likely we are to consider it to be urgent enough to defeat your obligations to others. We seek both to understand your attachment as being a member of a publicly justifiable category, and to understand the *strength* of the attachment as being proportional to other items in the same category and to items in other categories. This view is objective in the sense that all potential sacrifices must meet a criterion of adequacy that is independent of the person's interests and preferences. Nothing counts as a sacrifice *simply in virtue of* its being valued; but the fact that it *is* valued gives us reason to step back and examine *why* it is valued.

While an objective view of sacrifice might initially seem impersonal or inflexible, it actually accommodates a wide variety of attachments. Consider, for example, a case in which someone values something out of proportion to its apparent worth. In the movie *Rain Man*, Dustin Hoffman's character Raymond likes to see certain television shows at certain times. Raymond is autistic and entrenched in his routines; if he cannot see 'The People's Court' or 'Wheel of Fortune,' he gets very upset, so upset that he could hurt himself. It would surely be inhumane to take away Raymond's television. And if he were to give up the TV—even for something else worthwhile—it would clearly be a sacrifice. Under ordinary circumstances, we might find such a strong attachment to a television unintelligible. But in Raymond's case, the attachment can be understood using familiar general categories. He likes order, regularity, ritual, and predictable forms of stimulation. Without them, he feels insecure, perhaps even *unsafe* in a sense. We all understand the need to feel safe, even independently of the need to *be* safe. Thus even an objective view can accommodate a case where something that would not generate a legitimate claim for most people might nevertheless play an important role in a particular person's life. The same can be said for cases where something that *would* play an important role in *most* people's lives does *not* play such a role in a particular person's life.

Of course, thinking of sacrifice in this way requires a theory of well-being. In saying that sacrifices must be intelligible as members of general categories, we are essentially saying that sacrifices are *gross losses of well-being*,¹³ with the categories representing legitimate sources of well-

¹³ Liam Murphy (2000) takes a similar route. He defines the 'demands' imposed by a moral theory as the losses of well-being that would be incurred by complying with the theory. My view differs from Murphy's in at least two respects. First, he seems to interpret 'demands' as being only those losses

being. At the conceptual level, we can make use of an agent-neutral account like Darwall's 'rational care' theory. Darwall argues that a person's well-being (or her welfare, her 'good') is 'constituted, not by what that person values, prefers, or wants (or should value), but by what one (perhaps she) should want *insofar as one cares about her*' (2002, 4). In other words, a person's welfare is whatever it is rational for *us* to desire *for her* insofar as we care about *her* (9). On this view, 'the normativity of welfare is not agent-relative but agent-neutral' (20). That is, if Raymond's television contributes to his wellbeing, it is not because he likes and values it (though he does), but because it would be rational for those who care about Raymond to want him to have his television, for his own sake.

This view is compelling because it accommodates the wide variety of tastes and interests that play meaningful roles people's lives, without making well-being depend essentially on the agent's point-of-view. In other words, the fact that the agent values something does not *make* that thing part of her good; what is good for her can come apart from what she cares about. This is important because it leaves open 'the possibility of pursuing values one cares deeply about at some cost to oneself' (Darwall 2002, 3). It should be immediately clear why this possibility is going to be crucial to an account of sacrifice that bears on the behavior of moral saints. As Darwall puts it, 'if there were no difference between what a person valued and what benefited him, self-sacrifice would be impossible, except through weakness of will. [...] [I]t would be impossible for pursuing one's values ever to cost one *on balance*, since realizing a value would be the same thing as benefiting from it' (3). It is difficult to see how we could talk about moral saints at all if our theory of sacrifice left no room for actions that run counter to the agent's welfare.

Of course, this is a theory of well-being at the conceptual level only. It is not meant to provide a test, as it were, for us to apply in each and every case. It needs to be paired with another theory at the normative level: a theory that tells us, as a substantive matter, what sorts of things actually promote well-being. In keeping with the constraint we borrowed from Scanlon—that benefits and burdens need to be intelligible as members of certain general consensus categories in order to generate legitimate claims—it makes sense to pair Darwall's theory with some variant of an 'objective list' theory at the normative level. As Derek Parfit (1984) explains, 'On *Objective List Theories*, certain things are good or bad for us, whether or not we want to have the good things, or to avoid the bad things' (493). The types of things that appear on the objective list vary from theory to theory. Parfit thinks the good things on the list might include 'rational activity,' 'the development of one's abilities,' and 'the awareness of true beauty,' while the bad things might include 'being betrayed, manipulated, slandered, [or] deceived,' and 'being deprived of liberty or dignity,' among others (1984, 499). James Griffin has a list with just five broad items: (1) accomplishment; (2) the components of human existence; (3) understanding; (4) enjoyment; and (5) deep personal relations (1986, 67). Martha Nussbaum's list of ten 'capabilities' can also be thought of as an objective list

that would be imposed when a moral theory *requires* some course of action, whereas on my view a 'sacrifice' can be incurred not just by a required (i.e., obligatory) act, but also by a supererogatory act. Second, his view seems to count only *net* losses of well-being as demands, whereas on my view it is *gross* losses of well-being that constitute sacrifices. Thus an action can impose a sacrifice even if, in performing it, one thereby accrues certain benefits that are equal to or greater than whatever one is giving up.

theory; its items include life, bodily health and integrity, affiliation, play, and control over one's environment (1999, 41-42). Without endorsing any particular list, we can see that objective list theories of wellbeing meet the special constraints that arise in situations where moral obligation is at stake: they meet the intelligibility constraint; they are normative rather than descriptive; and they are objective insofar as they do not essentially depend on the agent's point-of-view. Other theories of well-being (e.g., hedonism) tend to fail on one or more of these criteria.

With this account of sacrifice in mind, we can return to cases like Paul Farmer and Susan Tom. It is clear that their lives are marked by great sacrifices: they give up things that would surely appear on any plausible objective list theory of well-being. Some of these things are mundane, like a full night of sleep, whereas others are sad, like Susan Tom's lack of meaningful adult relationships and Paul Farmer's lack of time with his family. Some of their sacrifices involve the *absence* of good things, such as hobbies or a savings account, whereas others involve the *presence* of bad things, such as the anguish of watching your child in pain or the frustration of seeing patients die from preventable illness.

Of course, we cannot ignore the fact that Farmer and Tom also reap great rewards from their actions. For Susan Tom, a houseful of fourteen children isn't a moral burden—it's just her family, and thus it is the most important thing in her life. For Paul Farmer, working on behalf of the world's poor sometimes brings great joy, and it prevents the guilt and angst he would feel if he did nothing; in a sense, he might be even worse off without the sacrifices. 'If you're making sacrifices,' he says, 'it stands to reason that you're trying to lessen some psychic discomfort' (Kidder 2003, 24). His own sacrifices, he says, can be seen as an expression of ambivalence. 'I feel ambivalent about selling my services in a world where some can't buy them' (24). Sacrifice keeps him going.

So the lifestyles of people like Paul Farmer and Susan Tom are not filled exclusively with misery and toil; their lives also involve unique goods, or 'gains' in well-being. Indeed, in many cases the losses *make the gains possible*. For example, Susan Tom takes on the challenge of raising several children who use wheelchairs, two of whom have no legs. This entails certain sacrifices: she must drive a special van; she cannot take the whole family on an outing unless the destination is accessible; she must comfort her children when kids at school make fun of them, etc. Nevertheless, parenting these particular children brings unique opportunities for happiness and joy. To take a somewhat dark example, Susan organizes an impressive backyard spectacle for Halloween, during which one of her sons performs the magic trick of sawing his sister in half. The trick is especially convincing given that the sister has no legs to begin with and is simply lying next to two fake stuffed legs. Everyone involved seems to find the stunt hilariously funny, in an ironic but not mean-spirited way. This moment of levity is made possible by the unique circumstances of Susan's family—the very same circumstances that require her to make extensive sacrifices. So one might wonder: might these moments of joy, these 'gains in well-being,' add up and outweigh all the losses? And might we therefore question whether Tom and Farmer really lead lives so filled with sacrifice as to make their actions beyond obligation?

Another way of putting the worry is this: the lifestyles of moral saints might turn out to yield a 'net gain' in well-being, and we should not count losses as sacrifices if they result in, or are part of a lifestyle that results in, an overall net gain of well-being. There are two assumptions

necessary for this worry to make sense: (1) that gains and losses from different sources of well-being are *commensurable*, and (2) that gains and losses from different sources of wellbeing can *compensate* for one another. Are these assumptions correct? The first, commensurability, would require that we somehow translate the wellbeing or illbeing generated from different sources into units on a common scale. Given the diversity of possible sources of wellbeing or illbeing—recall that objective list theories include items as varied as health, freedom from deception, and the awareness of true beauty—such commensuration would in many cases be extraordinarily difficult or perhaps impossible. Presumably its difficulty would vary with the similarity or dissimilarity of the items we need to commensurate. For instance, we might think that *mental health* and *physical health* can be roughly commensurated if we focus on factors that influence both—such as pain (emotional and physical) and dysfunction (say, the inability to handle social situations or the inability to walk). But how would we commensurate physical health with *freedom from deception*, especially if we think the value of not being deceived is independent of one’s knowing or caring that one is not deceived? This would seem to be a more intractable problem, and one that lies beyond the scope of this paper. Fortunately, it’s not necessary to solve this problem, because even if gains and losses in wellbeing from different sources were fully *commensurable*, it would not follow that the gains could *compensate* for the losses, if by ‘compensate’ we mean something like *make up for* or *replace without loss*.¹⁴ In other words, the second assumption underlying the worry—compensation—is false in most cases.

To see why this is the case, consider an example. A gymnastics prodigy gives up the normal pleasures of childhood to train and compete, thus gaining the pleasures of being an elite athlete. She gives up playdates, ice cream, sleeping in on weekends, attending a regular school, and a normal childhood free from extreme pressures and expectations placed on her by adults. In return she gains physical fitness, the thrill of victory, the satisfaction of perfecting a craft, and a college scholarship. Now supposing we could roughly commensurate all of these disparate losses and gains, we might conclude that the magnitude of the gains is equal to or greater than the magnitude of the losses, so that with respect to her wellbeing, being a gymnast is a ‘net gain’ or at least ‘net neutral.’ Nevertheless it would be odd to say that, in light of the net gain, the losses do not count as *sacrifices*. Victory is great, but so is childhood!¹⁵

¹⁴ If by ‘compensate’ we meant something weaker, like ‘counterbalance,’ then—assuming for the sake of argument that commensurability is possible—it would seem to be trivially true that gains from one source could compensate (counterbalance) losses from another source. After all, if we could translate the gains and losses into units on a common scale, ‘counterbalancing’ would simply be the act of rendering a sum in terms of those units, or balancing the metaphorical scale. But I take it that what’s at issue in arguments about whether losses count as sacrifices in someone’s life is not merely whether those losses have been ‘balanced’ by gains, but rather whether the gains render the losses insignificant or irrelevant to how things are going for that person. I grant that we would prefer a life in which losses are balanced by gains to a life in which there are losses but no gains. Nevertheless, even when losses are balanced in this weak sense, they are not ‘made up for’ in the meaningful sense, and so the gains do not render the losses *non-sacrifices*.

¹⁵ Michael Stocker discusses a related phenomenon in his book *Plural and Conflicting Values* (1990). In the chapter ‘Plurality and Choice,’ he gives an account of what it is to lack a good. His account allows for cases in which a life can coherently be said to *lack* a particular kind of value even if the

What this means is that lives with net surpluses of well-being may involve many sacrifices or few, and lives with a net losses of well-being may be morally extraordinary or morally repugnant; these notions can come apart from each other. Obligation is constrained by unreasonable sacrifice even when the sacrifices open doors to new sources of well-being. This is because gains of well-being in one area of one's life cannot be said to directly *compensate* for losses of well-being in another, in the most meaningful sense of 'compensate,' even if the gains are actually made possible by the losses. On this account, *gross* losses of well-being count as sacrifices even if there is no *net* loss. Indeed, given the difficulty of commensurating losses and gains from diverse sources, in most cases it may make little sense to even speak in terms of 'net' losses or gains.¹⁶

Returning now to the Puzzling Data, we can begin to understand agent-observer disparities in appraisals of the actions of moral saints.

PUZZLING DATA: Moral saints habitually perform actions that are, intuitively, beyond moral obligation. Yet the moral saints sometimes consider these actions to be obligatory, not supererogatory. There is a *persistent agent-observer disparity*.

What's going on here, it seems, is that the observers are noticing just how much these moral saints have given up. They are forming an intuitive sense of the gross losses of well-being these actions involve, and judging that no plausible moral theory could *require* someone to incur these losses.

The agents, on the other hand, have a different vantage point. They know exactly what it *feels* like to give up what they are giving up, and perhaps it doesn't feel as bad as we imagine it would. Moreover, they also know what it feels like to reap the *rewards* generated by their good deeds. From an impartial perspective, we may judge that these rewards do not compensate for or mitigate the sacrifices. But experiencing these rewards from the first-person perspective no doubt influences how the moral saints think about their own lifestyles. Perhaps one way to think about it is that the moral saints have calibrated their scales differently; first-hand knowledge of sacrifice has caused them to think that we would all be justified in demanding more of each other morally.

We might still think these moral saints are in some sense mistaken. Their vantage point causes them to overemphasize what it *feels* like to give certain things up and underemphasize what is really in their best interests. On reflection, perhaps they would not demand that others be held to the same standard to which they are holding themselves. Nevertheless, it is important that we take their perspective seriously. By testing the boundaries of sacrifice, they serve to shake us from our

absence of that value makes possible the presence of other important values. Stocker argues that value is essentially plural, and therefore we cannot say that one value could ever really *compensate* for the loss of another value. Thus he claims that 'a life can be lacking in pleasure even if the only way to have additional pleasure would involve losing so much wisdom and honour that the life with greater pleasure would be overall worse' (171). I think this is correct, and I think my notion of sacrifice is similar to Stocker's notion of *lack*. Paul Farmer is *sacrificing* time with his family even his life is overall better on account of this sacrifice, because the goods generated by his lonesome lifestyle do not *replace*, as it were, the time spent with family.

¹⁶ I thank an anonymous referee for encouraging me to clarify the points made in the preceding discussion.

complacency. We might discover that the more we learn about moral saints, the more we think we can reasonably demand of each other. In the next section, I argue that moral saints serve precisely this purpose.

4. THE ARGUMENT FOR RATCHETING

In this section I explore how the boundary between obligation and supererogation might be flexible, and how the behavior of moral saints affects where the boundary lies for the rest of us.¹⁷ We tend to think that there is some amount of hardship or sacrifice that we just cannot expect people to bear for the sake of morality. James Fishkin (1982) calls this the ‘cutoff for heroism’ (5). But what exactly is that cutoff, and could it change? After all, if certain psychological and sociological facts were different, the boundary could easily have ended up much lower or higher than it is. If getting your shoes wet were a grave hardship, then perhaps we would deem someone a ‘moral saint’ just for stepping in a puddle to help an old lady cross the street. Indeed, we might think what passes for sacrifice and hardship among well-off members of wealthy countries is already severely distorted.

This distortion is highlighted in a recent *New Yorker* cartoon titled ‘Latter-Day Saints.’¹⁸ It shows people on the streets of New York City, with descriptions of the ‘sacrifices’ they are making. ‘Eats however many grams of fiber you’re supposed to’ reads the bubble next to a woman, and next to her young son it says, ‘Is putting up with a 56K modem.’ One man ‘does all the grocery shopping’ while another ‘has commuted from New Canaan every day for 32 years.’ The other weighty burdens include: ‘Jogs ten miles a day, no matter what’; ‘Buys her shampoo in an ordinary drugstore’; ‘Doesn’t own a TV’; ‘Forcing self to read ‘Beowulf,’ although she hates it.’

The cartoon makes fun of people for over-rating themselves, for valorizing any old task that they would rather not do. We don’t actually think tolerating a 56K modem involves a particularly significant loss of objective well-being or makes you a saint; but we do find it irksome—it *feels* like more of a sacrifice in light of our high expectations. Perhaps that is why we are so awed by people like Paul Farmer and Susan Tom: for those of us who can’t handle a slow internet connection, what *they* do is unimaginable and shocking. This shock could send us in either of two directions. We might think that they are just fundamentally better people than we are and there is no hope for us. On the other hand, we might see their lifestyles as evidence that we could do *more*, and that it might not be as hard as we thought. This latter thought is the more fitting reaction, and I shall argue that it has implications for what we are required to do. This isn’t just an epistemological claim about how moral saints influence our beliefs about our obligations. It’s a claim that the epistemology of the

¹⁷ Not everyone agrees that it makes sense to think in terms of this boundary. In ‘Above and Below the Line of Duty’ (1986), Susan Wolf claims that the concept of obligation (or duty) has had too much influence on our thinking about morality. ‘There is a line of duty,’ she writes, ‘but it is, necessarily, a dotted line’ (32). In recent work on ‘Deontic Pluralism’ (in progress), Margaret Little argues that we need many more concepts than just the classic obligatory, forbidden, permissible, and supererogatory.

¹⁸ Roz Chast, ‘Latter Day Saints.’ *The New Yorker*, October 16, 2006.

moral boundary could actually change the *metaphysics* of the boundary. Our obligations depend on what we reasonably *believe* about our obligations. Learning about moral saints can change what beliefs would be reasonable.

We commonly think that in most cases, whether others are (or are not) fulfilling their moral obligations does not bear on what *I* ought to do. According to this thought, our obligations are what they are, regardless of whether other people are behaving well or badly.¹⁹ This thought serves us well most of the time. For instance, if it is wrong to keep the extra \$10 a cashier gives you by mistake, then surely it is wrong regardless of whether 87% of people would keep the money. And if you are morally obligated to get a bike helmet for your child, surely you are obligated even if only some of the other parents are doing the same. In cases like these, we consider it a lousy excuse to say ‘but everyone else was doing it.’ Indeed, we point out the lousiness of this excuse to children when trying to teach them right from wrong.

Nevertheless, not all cases work this way. The cashier and bike helmet examples are perhaps misleadingly simple. If I don’t return the extra \$10 to the cashier, the register will come up short, and no less short on account of the fact that I’m doing something fairly common. The cashier could be blamed for the shortage and could even lose his job. In keeping the money, I not only risk harming an innocent person, but in some sense I’m also lying and stealing. My action has these morally troublesome features regardless of what other people are doing.

But there are many situations where what other people are doing might directly affect what morality demands of each person. In aid scenarios, the number of people helping might determine how much each person needs to contribute, and in collective action problems like whether to walk on the pathway or on the grass, the noncompliance of some might render meaningless the compliance of others. Indeed, philosophers have taken very seriously the idea that the compliance or noncompliance of other people bears on our moral obligations. The difficulty is in figuring out whether others’ noncompliance means that I ought to do *more* or that I am permitted to do *less*. Both positions have some intuitive plausibility. Some might think, ‘Since no one else is donating to famine relief, surely *I* don’t have to. At most I ought to give only what my fair share would be if everyone were donating.’ However, others might think, ‘Since no one else is donating to famine relief, surely I ought to donate *as much as I possibly can*, because everyone else has left me with a greater need to fill. The less they give, the more I must.’²⁰

¹⁹ David Estlund (2007) thinks moral philosophers differ from political philosophers with respect to this point. ‘Moral philosophers know that people are likely to lie more than they morally should,’ he writes, ‘but this doesn’t move many theorists to revise their views about when lying is wrong. Things are often different in political philosophy’ (12).

²⁰ See Singer (1972) and Unger (1996) for examples of the view that one should give as much as one can. See Schapiro (2003) for the view that others’ noncompliance can render our own compliance meaningless, and Murphy (2000) for the view that we should only do what would be our fair share under full compliance. Recently, Michael Ridge has articulated a compromise position in ‘Fairness and Non-Compliance’ (2010). Ridge claims that we overemphasize the unfairness to the compliers in situations of partial compliance and neglect the unfairness to the persons who are on the *receiving* end of the aid, who are forced to bear an unfair share of the burden. He argues that the burdens should be distributed fairly amongst the compliers, and between the compliers and the recipients,

Philosophers have traditionally approached these questions from the perspective of noncompliance or partial compliance—that is, situations where most people, or at least a significant number of people, are *not* discharging a particular moral obligation. My argument, however, focuses not on *noncompliance* but on *overcompliance*—the overcompliance of moral saints. For we are no longer talking about people failing or succeeding in doing their duty, but rather about people *exceeding* their duties. And I am not concerned with particular obligations (say, to famine relief) but with obligations and moral commitments more generally. Just as some philosophers think that our obligations are different in situations of partial compliance than they would be in situations of full compliance, I will argue that our obligations are different when there are *overcompliers* than they would be in the absence of such people.

The argument begins once again with the notion that moral obligation is fundamentally a matter of what we can legitimately demand of one another. Moral obligation in this sense is subject to certain constraints. The first constraint is what I call the Sacrifice Condition, and it was mentioned earlier as the principle ‘obligated implies no unreasonable sacrifice.’

Sacrifice Condition (S). Someone can be obligated to ϕ only if ϕ -ing does not involve an unreasonable sacrifice.

Another, perhaps more controversial constraint is what I will call the Knowledge Condition:

Knowledge Condition (K). Someone can be obligated to ϕ only if she could know (or it would be reasonable for her to believe) that she is obligated to ϕ .

If the Knowledge Condition is correct, then moral obligation is dependent in an important way on our evidence. This suggests a refinement of the Sacrifice Condition. Call this refinement K-S, the Knowledge of Sacrifice Condition.

Knowledge of Sacrifice Condition (K-S). Someone can be obligated to ϕ only if she could know (or reasonably believe) that ϕ -ing does not involve an unreasonable sacrifice.

We can begin to see how the Ratcheting-Up Effect would work: if K-S is true, our obligations are constrained by what we could reasonably believe about how much of a sacrifice it would be to perform a given action. And what we can reasonably believe about this level of sacrifice is surely a

which means the compliers may have to give more than they would in a situation of full compliance, but less than they would if their obligation was completely unconstrained. Thus Ridge demands less than Singer and Unger but more than Murphy.

function of what sort of *evidence* is available to us. When the lives of moral saints are publicized, our pool of evidence about sacrifice changes. To be sure, the behavior of *everyone*—not just moral saints—contributes to our stock of evidence. But moral saints play a unique role, both because they are rare and because they behave at the margins of obligation, taking on unusual sacrifices. In short, the argument goes like this:

- (1) Moral obligation is at its core about what we can reasonably demand of each other.
- (2) We cannot reasonably demand that someone do something that it would be reasonable for her to believe to be too much of a sacrifice (the K-S Condition).
- (3) Exposure to moral saints can change what it would be reasonable to believe about how much of a sacrifice it would be to take on certain actions or patterns of behavior.²¹
- (4) So, exposure to moral saints could change what a person is obligated to do by way of removing a defeater of obligations.

Much of the argument rests on the Knowledge Condition (and its derivative K-S). We ought to look more closely at this condition, since it is not uncontroversial.

5. OBJECTIONS TO THE KNOWLEDGE CONDITION

The Knowledge Condition says that, if it is reasonable for someone to believe that she is not obligated to perform an action, then she is not obligated to perform the action. At first glance, this may seem too permissive. Ignorance about morally relevant facts is fairly common, so one might worry that the knowledge condition releases people from too many obligations. Yet this worry is misguided. For the knowledge condition only governs cases of what we might call *justified ignorance*; it does not govern just *any* case of ignorance. It trades on what it would be reasonable for us to believe, not on what we actually believe. Some people might believe that beating their children ‘builds character,’ and that parents are not obligated to refrain from beating their kids. But surely it is not *reasonable* to believe either of these things, at least not in the modern-day epistemic environment.

According to this way of thinking about moral obligation, *what I am obligated to do* simply amounts to the same thing as *what I am obligated to do in light of my justified beliefs*.²² We can see the

²¹ As I will explain below, it makes no difference if the moral saint’s action is in fact supererogatory, because there will be some lesser action that involves a proportionately lesser amount of sacrifice, which itself is not supererogatory, and which the observer will come to be obligated to perform.

²² In ‘Justified Wrongdoing’ (1997), Sarah Buss argues that ‘an agent’s blameworthiness is a function of what he can reasonably be expected to know’ (338). According to Buss’s account, justifiable

motivation for this principle if we look at a puzzle about utilitarianism. According to act utilitarianism, the right action is the one with the highest utility compared to the alternatives. But is the agent morally obligated to perform the action with the highest *actual* utility, or the action with the highest *expected* utility? Suppose that John can donate to Charity A or Charity B, and the best available evidence indicates Charity A will produce more utility. It turns out, however, that Charity A is actually in going through an unpublicized crisis and John's money will go further at Charity B. Despite this fact, it seems that the most plausible theory of moral obligation is the one that says John should maximize *expected* utility, and thus he was obligated to donate to A. Indeed, it would not even make sense for the moral community to demand that John maximize *actual* utility, since the only way to do so would be to aim for it via *expected* utility. If obligation were determined by the actual utility of the consequences regardless of the agent's beliefs and expectations at the time of action, then we would only be able to determine the agent's obligations either retroactively or from an omniscient perspective. Of course, we can still say that given how the things turned out, John 'ought' to have chosen B rather than A, meaning that it would have been better if he had. But the 'ought' here is not the same sense of moral obligation that is grounded in the demands we justifiably make of one another. For how could we be justified in making demands on others that they will be systematically incapable of living up to on account of their limited information?

Indeed, this would seem to be so unfair that some philosophers have proposed simply separating the 'criterion of rightness' in utilitarianism from the 'decision procedure.'²³ The idea here is that an action's rightness or wrongness is a matter of the *actual* consequences, but what an agent 'ought' to do is follow the most rational decision procedure—such as maximize expected utility. So John did the 'wrong' thing, but since he was morally obligated to act using a particular decision procedure, he lived up to his moral obligation. Notice, however, that when we choose to separate what's 'right' from what's obligatory in this way, 'rightness' is no longer particularly connected to what we can justifiably demand of one another. And that notion—what we can justifiably demand of one another—seems to remain the key moral concept governing how we can live together in a moral community. So even if we cordon off a concept of 'rightness' that is not subject the Knowledge Condition, it seems that the concept of moral *obligation* still is.

One worry about the Knowledge Condition is that ignorance never alters what we are *obligated* to do, only what sort of behavior we can justify, or what behavior we deem praiseworthy or blameworthy. The idea here is that our evidence has *practical* consequences about what we can hold others responsible for, but it has no *metaphysical* consequences about what our obligations are. In science, we might distinguish the true theory from the theory we are justified in believing given our

ignorance can result in blameless wrongdoing. Whereas Buss's account focuses on *wrongdoing*, my account focuses on *obligation*. I argue that a certain type of justifiable ignorance does not merely free a person from blame, but in fact frees him from obligation. His action may remain 'wrong' in some sense. But it is *obligation*, and not wrongness, that best captures the type of normative expectation of behavior for which members of the moral community can hold each other accountable.

²³ See, e.g., Bayles (1971).

evidence, and here we might distinguish what is genuinely the obligatory action from what is the justifiable action given our evidence.²⁴

But how can what a person is ‘justified’ in doing come apart from what she is obligated to do, if moral obligation really is a matter of what we can reasonably demand of one another? It seems to me that what we can reasonably demand of you *just is* whatever you can be justified in doing in light of your information. A stronger interpretation of reasonable demands, according to which we could reasonably demand things of you even if you could not know that you ought to do them, would seem to strain the credibility of the term ‘reasonable.’ Consider an example: Presumably there was a time—if not in the 20th century, then surely in, say, the 17th century—when it was entirely reasonable to believe that second-hand smoke had no negative health consequences. There simply was no available evidence to the contrary. Surely it would be incorrect to say that people in 1650 were *obligated* not to smoke around their kids, but not blameworthy for failing to meet this obligation. Instead, we ought to say that there was no obligation in the first place.

Another way to frame the ‘no metaphysical consequences’ worry is to make a distinction between a *justification* and an *excuse*.²⁵ As before, I will use ‘justification’ to mean a consideration that actually *releases* an agent from a given obligation. In this sense, justifications genuinely affect the metaphysics of obligation, whether by eliminating or nullifying an obligation. An ‘excuse,’ on the other hand, is simply a consideration in light of which the agent is not blameworthy for failing to heed the obligation. When an excuse is present, the metaphysics of the obligation remain unchanged; all that changes is whether or not we can blame the agent. For example, if I’m obligated to meet you for lunch, but on the way to meet you I encounter a severely injured person who needs my help, we might say that I am *justified* in missing lunch—I am no longer obligated to be there, or at least my obligation does not have its original force. I have done nothing wrong. However, if I simply forget about our lunch, not from carelessness but because my mind has been preoccupied with something important lately, we might say I am just as obligated to be there as ever, but since it was an innocent mistake, I have an *excuse*—I am not blameworthy. With this terminology in place, we might formulate the objection to the Knowledge Condition this way: ignorance of our moral obligations is never a justification, but it may sometimes be an excuse. In other words, there is no Knowledge Condition on moral obligation, only on moral blameworthiness. Agents living in 1650 *were* morally obligated to protect their kids from second-hand smoke; they just didn’t realize it. Their behavior was not *justified*, but they did have an *excuse*. While the obligation was there all along, the agents were not blameworthy for failing to heed it.

Yet I don’t think that this language of ‘*justifications*’ and ‘*excuses*’ makes the objection any more compelling. Intuitions may differ on whether reasonable ignorance of one’s purported obligations is more like the case of the preoccupied person who simply forgets her lunch date, or more like the person who encounters an emergency along the way. But on the account I’ve given, moral obligation is fundamentally about what we can reasonably demand of one another. Just as it would be unreasonable for us to demand that you show up for lunch dates when an emergency

²⁴ I’m grateful to Elizabeth Anderson for voicing this worry to me and presenting it via this analogy with science.

²⁵ Thanks to Stephen Darwall for suggesting this terminological distinction.

arises, and just as it would have been unreasonable to demand that 17th-century parents worry about second-hand smoke, it seems that it would be unreasonable to demand that members of the moral community perform actions even when their evidence suggests that such actions are not obligatory.

Yet while I do not find the ‘No Metaphysical Consequences’ objection compelling, it turns out that this objection would not be devastating to the Ratcheting-Up argument even if it were true. After all, the Ratcheting-Up Effect could exist even if only *blameworthiness* were subject to the knowledge condition. What we can be blamed for would be *ratcheted-up* by the behavior of moral saints. This in itself would be a normatively significant result, even if it did not constitute a change in the underlying metaphysics of obligation. Still, the notion of moral obligation I’ve put forward is closely tied to blameworthiness; on my account, we could not ratchet-up what we could *reasonably blame* people for without also ratcheting-up what we can reasonably *demand of* people, that is, without ratcheting up the obligations themselves.

6. TWO MECHANISMS OF MORAL RATCHETING

More needs to be said about the role of *evidence* in the Ratcheting-Up argument. How exactly does the ratcheting happen? In other words, how can information about the sacrifices involved in *saintly* action constitute evidence that bears on *obligatory* action, since presumably we are not obligated to be moral saints? Let me suggest two possible mechanisms: the Cascading Ratchet and the High-to-Low Ratchet²⁶. Both mechanisms depend on the following general principle: an ordinary person learns about Paul Farmer or Susan Tom and thinks, ‘Adopting a dozen ill and disabled children sounds terribly burdensome. But Susan’s life is not as miserable as I would have thought. If *her* level of moral action involves only *that* level of sacrifice, then surely a much more modest level of action would involve even *less* sacrifice. Maybe I can do more than I realized.’²⁷

CASCADING RATCHET. Suppose that I am giving \$100 a year to global public health organizations, while Paul Farmer is working twenty-hour days and sleeping in a hut to treat as many patients as he can. It might seem that Farmer’s actions are too different than my own, and too alien to my lifestyle, to be particularly good evidence about what it would be like for me to do a little more. If this is right, then one way to understand ratcheting is as a cascading effect: Farmer ratchets up the people who are doing just a little bit less than him, and they ratchet up the people just below them, and so on, all the way down, so that Farmer’s efforts have an *indirect* influence on me and the other \$100 donors. The appeal of the cascading ratchet is that it seems to require fewer leaps of faith when it comes to seeing the sacrifices someone else makes as bearing on the sacrifices one might

²⁶ I’m grateful to Aaron Bronfman, Tim Sundell, and Dustin Locke for helpful discussion about evidence and ratcheting.

²⁷ In ‘Moral Exemplars’ (1988), Lawrence Blum proposes an idea that is similar in spirit, if not in mechanism. He writes, ‘While we cannot all be moral exemplars, what we all can be is *better* than we are. For each of us there are surely *some* virtues which we could come to possess in greater degree than we do now, no matter what circumstances we are in. The existence of moral exemplars can help to keep this possibility alive for us, and in some cases, to point the way to particular direction we might take in our own moral improvement’ (216-217).

make oneself. Knowing that daylong hikes in the mountains to make housecalls can actually be quite refreshing, or that sleeping in a hut is not so bad after all—knowing these things would seem not to be of much use to the person who is thinking of going from a \$100 per year contribution to \$500 per year, or thinking of volunteering at a local blood drive. However, this knowledge might be relevant to someone who already works in Farmer’s field and is considering moving to a remote location to take on more work. Then *that* person’s sacrifices could be relevant to people one level further down, who want to spend their two-week vacation at a rural clinic. And so on.

HIGH-TO-LOW RATCHET. I think that the Ratcheting-Up Effect actually works more directly than the Cascading Ratchet suggests. In fact, I would propose that evidence about sacrifice is actually *better* when it comes from someone *further* away on the moral continuum. Call this High-to-Low Ratchet. On this view, it is not particularly useful to look at the sacrifices being made by your nearest moral neighbor. So, for example, it would not be useful for the \$100 per year person to study the sacrifices made by someone who gives \$500 each year. She should already have a pretty good idea of this herself, after all. She knows what that \$500 could buy her and what she would have to give up. She is unlikely to gain any remarkable new justified beliefs about sacrifice by looking at the \$500 person, and so it is unlikely that any defeaters of moral obligation will have been removed by exposing her to this new information.

Indeed, perhaps the force of the ratcheting-up effect lies precisely in the *distance* between agents and observers.²⁸ A lifestyle like Paul Farmer’s is utterly foreign to the ordinary agent, the one who gives just \$100 per year. Until she is exposed to someone like Paul Farmer or Susan Tom, she has every reason to believe that lifestyles like theirs would be unbearable, perhaps even impossible. No one, the ordinary person might think, could single-handedly raise over a dozen ill and disabled children. No one, she might think, could first watch a child die from a devastating and painful illness, and then volunteer to adopt the child’s sibling who is dying from the same horrible illness. At least, no one who is otherwise normal could do these things without some self-serving incentive, and without becoming emotionally deadened, clinically depressed, or less caring toward children with less extreme needs.

This is what one might think, until one looks at Susan Tom’s life as a case study in sacrifice. It turns out one woman *can* do all of those things, and *without* all of those horrible effects, and *despite* being otherwise normal. What might have been outlandish or unreasonable to believe is now as clear as day. Thus the ordinary person, in learning about her counterpart who lives an extraordinary life, now has a meaningfully different stock of evidence on which to base beliefs about the range of lifestyles she might live. However—and this is an important subtlety—lives like Paul Farmer’s and Susan Tom’s may *still* be more self-sacrificial than could be reasonably thought to be *obligatory*. The idea here is *not* that, in observing the moral saints, we discover that their lives don’t involve as much

²⁸ Here I am taking a position that is essentially the opposite of someone like Melden (1984), who thinks saints are so fundamentally different from the rest of us that we cannot ever be like them—indeed we cannot even *aspire* to be like them! For more discussion of the role of psychological realism in thinking about moral saints, see Flanagan (1991).

sacrifice as we would have thought.²⁹ Nor is the idea that these moral saints don't seem horribly burdened by their sacrifices, and what really determines our obligations is how burdensome the losses *feel*. (As I argued in Section 3, it is not *felt burden* but rather gross loss of objective well-being that constitutes sacrifice.) Rather, the idea is simply that the moral saints model for us that these sacrifices are possible. Indeed it is possible to integrate these sacrifices into a full life, a life that is not miserable or shallow or phony, though it may be difficult and tiring. But they are still *sacrifices*, and it is still in a sense *bad for* Susan Tom or Paul Farmer to have made them. All else held equal, they would be better off if they didn't have to incur these losses, and indeed the losses are large enough that we judge their actions to be supererogatory. But while we judge that these particular lives go beyond the call of duty, we can extrapolate that a life involving a bit *less* of a commitment than Tom's or Farmer's would involve a bit *less* sacrifice, until a point is reached where the sacrifice is no longer unreasonable, and then we reach an obligation. In addition to extrapolating downward from the top in this way, the ordinary person can also extrapolate upward from the bottom, by way of a kind of shaming. Compared to Paul Farmer's lifestyle, giving \$100 is a joke, and even \$200 would be rather laughable, and \$1000 might not even be so bad, and even \$5000 might be bearable. Ordinary people can thus extrapolate in two ways—top-down and bottom-up—to triangulate toward a middle ground. As we incorporate new reasonable beliefs into our stock of evidence, we find that morality demands more of us than before.

Of course, the preceding discussion has made it sound like ratcheting is some sort of conscious effort of self-improvement on the part of the observers. But nothing like that is necessary for the Ratcheting-Up Effect to take place. It is enough that people be exposed to the lifestyles of extraordinary moral agents, and through that exposure gain knowledge about burdens and benefits. One need not do any conscious deliberation about where exactly in between one's current level of commitment and the level of a moral saint one should aim to insert oneself. That can come unconsciously or through the gradual shift in social norms governing self-sacrificial behavior. What is most interesting about ratcheting is not that it makes an empirical prediction that moral saints will cause others to *do* more, but that it has a normative entailment: that exposure to moral saints will have the effect of increasing what people can be *required* to do.³⁰

²⁹ Consider by analogy an athlete who performs at the boundaries of human achievement, perhaps an ultramarathoner. Before learning about ultramarathons, it might be reasonable for the ordinary person to believe that a human simply could not run 100 or 1000 miles without dying, or at least without irrevocable injury (whether physical or psychological!). In learning about the athletes who do these things, it wouldn't make sense to conclude, 'that's not so hard after all, and I should really expect myself to run 100 miles instead of a mere 26.' Instead one ought to conclude, 'that's *extremely* difficult, and involves untold hard work, pain, and mental suffering of the sort it wouldn't be reasonable to expect of myself—but it *can* be done, and by mere mortals, and without irrevocable injury. So maybe I can do a bit more than I would have thought.'

³⁰ Indeed, I do not intend to argue for the empirical claim that exposure to moral saints will cause agents to *do* more. I think it probably would, but I'm not in a position to test the claim. Interestingly, Kant seemed to think that exposure to a certain kind of moral saint could have the opposite effect. In 'The Doctrine of the Method of Pure Practical Reason' in his *Critique of Practical Reason* (1997), he argues that we ought to use case studies and vignettes to teach morality to children, but that we

The Ratcheting-Up Effect illustrates one role moral saints play in the moral ecosystem. To be sure, moral saints are exceptional people living highly unusual lives. But one way of seeing them is simply as otherwise ordinary people acting as test cases for the rest of us, like marathon runners breaking from the pack. They push the boundaries of moral behavior, discovering what level of sacrifice it is possible to bear and what bearing it entails for the agent and her life. Moral saints may find that their extraordinary deeds are easier (relatively speaking) than even they themselves would have thought, and then push themselves even further, all the while claiming that they are doing nothing special. Even if they are mistaken, and in fact we cannot *demand* that they push themselves this far, their behavior still contributes to our pool of evidence, removing excuses that may have otherwise been valid. Through a feedback-loop involving perceptions and predictions about sacrifice, this information serves to change what we can reasonably demand of each other. Thus we can think of moral demands, like legal systems, as flexible over time and influenced by precedent. Moral saints set precedents and landmarks that affect the rest of us.

This brings us back to the ‘Puzzling Data’. Recall that it seemed puzzling that moral saints so frequently declare their extraordinary actions to be unremarkable, even obligatory. I argued that part of what is going on in these cases is that the moral saints and the observers have different perspectives on how much sacrifice the actions involve. Sacrifice, I argued, is properly measured by

should avoid using examples of ‘so-called *noble* (supermeritorious) actions’ (127). Such examples will cause children to become ‘heroes of romance’ (128) or ‘fantasizers’ (129). Lofty ideals will distort their thinking so much that they will take themselves to be exempt from mundane duties. Of course, it’s not clear why exactly a pupil would, on the basis of having adopted certain lofty ideals about morality, decide that he was no longer interested in, or in fact no longer *bound* by, everyday moral duties. If we could get children to be committed to ‘greatness’ and to admire great moral figures, wouldn’t this raise morality to salience in a way that would make them more cognizant of even the most banal duties of daily life? The problem here is that in condemning the use of ‘noble’ figures in education, Kant is not just making a point about the *magnitude* of their moral worth—that they are *too good* for emulation; rather, he is making a point about the *nature* of their moral worth—they are not good in the right *way*. For Kant is using ‘noble,’ ‘meritorious’ (or ‘supermeritorious’), and ‘magnanimous’ to refer to a kind of touchy-feely disposition that is precisely *not* an orientation toward *duty*. Since moral education should focus primarily on inculcating a sense of duty, such education should ignore these touchy-feely noble agents found in the ‘sentimental’ literature. Indeed, in a footnote Kant concedes that while we should praise ‘actions in which a great, unselfish, sympathetic disposition or humanity is manifested,’ in the course of educating children we should emphasize the ‘*subjection of the heart to duty*’ rather than the ‘*elevation of soul*’ (128). He seems to be worried about the type of moral saint Susan Wolf called the ‘loving saint’ (1982, 420). The loving saint helps others because his happiness ‘would truly lie in the happiness of others, and so he would devote himself to others gladly, with a whole and open heart’ (420). Kant thinks that to use a loving saint as an exemplar is to send a mixed message: to preach respect for the moral law for its own sake while drawing attention to agents who seem to be driven by happiness instead. He claims that morality will ‘have more power over the human heart the more purely it is presented’ (129). Presenting morality ‘purely’ means not allowing pupils to see any incentive in acting morally, especially not the incentive of happiness. In fact, Kant thinks the best way to present ‘the law of morals and the image of holiness’ is ‘*in suffering*’ (129, emphasis added). For more on interpreting Kant’s remarks here, see Baron (1987).

gross losses of well-being, where well-being is understood objectively and agent-neutrally. The moral saints may be giving up more than they realize, because we can be mistaken about how bad things are for us. And yet we do not immediately dismiss the moral saints' claims that they are 'only doing their duty.' We take these claims seriously. We find it disconcerting that the moral saint is living up to such a high moral standard; we are shaken from complacency. In learning more about the lives of moral saints, we might begin to see them as hypothetical interlocutors in a moral conversation, as a potential audience for our excuses and complaints. As G.A. Cohen has argued, normative arguments can have more or less force depending on the speaker and the audience (1995, 341). For instance, my excuses for not giving more to charity sound 'feeble' when my interlocutor already gives far more than I do (342). With moral saints in the mix, we may find that the standards of what we can reasonably demand from each other are much more flexible than we thought.³¹

Department of Philosophy
University of Cincinnati

³¹ I am grateful to Stephen Darwall, Elizabeth Anderson, and Peter Railton for their comments on an earlier draft of this paper. I also wish to thank the many philosophers who offered feedback when I presented earlier versions of this work at the University of Michigan, especially Sarah Buss, Victor Caston, Eric Lormand, Jamie Tappenden, Louis Loeb, Howard Nye, Dustin Locke, Aaron Bronfman, Tim Sundell, David Dick, and Ivan Mayerhofer. Thanks also to an anonymous referee for this journal. Despite all of this help, many confusions undoubtedly remain, and these are my responsibility alone. Work on this paper was supported by grants from the Charlotte Newcombe Foundation and the University of Cincinnati's Charles Phelps Taft Research Center.

REFERENCES

- Adams, Robert M. (1984) 'Saints.' *Journal of Philosophy* 81, 392-400.
- Baron, Marcia. (1987) 'Kantian Ethics and Supererogation.' *The Journal of Philosophy* 84:5, 237-262.
- Bayles, Eugene. (1971) 'Act-Utilitarianism: Account of Right-Making Characteristics or Decision-Making Procedure?' *American Philosophical Quarterly*, 8:3, 257-265.
- Blum, Lawrence. (1988) 'Moral Exemplars.' *Midwest Studies in Philosophy* 13:1, 196-221.
- Buss, Sarah. (1997) 'Justified Wrongdoing.' *Noûs* 31:3, 337-369.
- Carbonell, Vanessa. (2009) 'What Moral Saints Look Like.' *Canadian Journal of Philosophy* 39:3, 371-398.
- Chast, Roz. (2006) 'Latter-Day Saints.' (Cartoon. *The New Yorker*, 10/16/06.)
- Cohen, G.A. (1995) 'Incentives, Inequality, and Community.' *Equal Freedom: Selected Tanner Lectures on Human Values*. Ed. Stephen Darwall. (Ann Arbor: University of Michigan Press.)
- Darwall, Stephen. (2002) *Welfare and Rational Care*. (Princeton: Princeton University Press.)
- Darwall, Stephen. (2006) *The Second-Person Standpoint*. (Cambridge: Harvard University Press.)
- Estlund, David. (2008) *Democratic Authority*. (Princeton: Princeton University Press.)
- Fishkin, James. (1982) *The Limits of Obligation*. (New Haven: Yale University Press.)
- Flanagan, Owen. (1991) *Varieties of Moral Personality*. (Cambridge: Harvard University Press.)
- Fogelman, Eva. (1994) *Conscience and Courage: Rescuers of Jews During the Holocaust*. (New York: Anchor Books.)
- Gilbert, Martin. (2003) *The Righteous: The Unsung Heroes of the Holocaust*. (New York: Henry Holt.)
- Griffin, James. (1986) *Well-being: Its Meaning, Measurement and Moral Importance*. (Oxford: Clarendon Press.)
- Kagan, Shelly. (1989) *The Limits of Morality*. (Oxford: Oxford University Press.)
- Kant, Immanuel. (1997) *Critique of Practical Reason*. Trans. Mary Gregor. Cambridge: Cambridge University Press.
- Karsh, Jonathan. Director. (2003) *My Flesh and Blood*. (DVD. HBO Documentary Films.)
- Kidder, Tracy. (2003) *Mountains Beyond Mountains: The Quest of Dr. Paul Farmer, A Man Who Would Cure the World*. (New York: Random House.)
- Levinson, Barry. Director. (1988) *Rain Man*. (DVD. MGM Studios.)
- Melden, A.I. (1984) 'Saints and Supererogation.' *Philosophy and Life: Essays on John Wisdom*, ed. Ilham Dilman. (Boston: Martinus Nijoff), 61-81.
- Monroe, Kristen. (2004) *The Hand of Compassion: Portraits of Moral Choice During the Holocaust*. (Princeton: Princeton University Press.)
- Murphy, Liam B. (2000) *Moral Demands in Nonideal Theory*. (Oxford: Oxford University Press.)
- Nussbaum, Martha. (1999) *Sex and Social Justice*. (Oxford: Oxford University Press.)
- Parfit, Derek. (1984) *Reasons and Persons*. (Oxford: Oxford University Press.)
- Ridge, Michael. (2010) 'Fairness and Non-Compliance.' *Partiality and Impartiality*, ed. Brian Feltham and John Cottingham. (Oxford: Oxford University Press, pp. 194-222.)
- Scanlon, T.M. (1975) 'Preference and Urgency.' *The Journal of Philosophy* 72:19, 655-669.

- Schapiro, Tamar. (2003) 'Compliance, Complicity, and the Nature of Nonideal Conditions.' *The Journal of Philosophy* 100:7, 329-355.
- Singer, Peter. (1972) 'Famine, Affluence, and Morality.' *Philosophy and Public Affairs* 1:1, 229-243.
- Stocker, Michael. (1968) 'Supererogation and Duties.' In Nicholas Rescher, Ed., *Studies in Moral Philosophy*. (Oxford: Oxford University Press, pp. 53-63.)
- Stocker, Michael. (1990) *Plural and Conflicting Values*. (Oxford: Oxford University Press).
- Trianosky, Gregory. (1986) 'Supererogation, Wrongdoing, and Vice: On the Autonomy of the Ethics of Virtue.' *The Journal of Philosophy* 83:1, 26-40.
- Unger, Peter. (1996) *Living High and Letting Die*. (New York: Oxford University Press.)
- Urmson, J.O. (1958) 'Saints and Heroes.' In A. I. Melden, Ed. *Essays in Moral Philosophy*. (Seattle: University of Washington Press, pp. 198-216.)
- Wolf, Susan. (1982) 'Moral Saints.' *The Journal of Philosophy* 79:8, 419-439.
- Wolf, Susan. (1986) 'Above and Below the Line of Duty.' *Philosophical Topics* 14:2, 131-148.