Demandingness, Integrity, and Consequentialism

How Consequentialist Theories Can Avoid the Demandingness Objection

George R. A. Baker

UCL

MPhilStud

I, George Rory Andrew Baker, confirm that the work presented in this thesis is my own.
Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

One of the most prominent arguments against act consequentialism is that it generates extreme demands that agents find intuitively objectionable. This extreme demand is generated by a combination of its requirement that one must always act in whatever way would make things go best, and the fact that there are a great many people with extreme needs whom a well-off person may easily be able to help.

In this thesis I argue that our intuitions regarding demandingness are not enough to motivate this problem, and that a baseline of acceptable demand is necessary in order to be able to determine whether or not act consequentialism is too demanding. Referencing Williams' infamous integrity problem, I argue that the intuitions that agents have regarding high levels of demand are an expression of their attitudes towards their integral projects, and that a suitable baseline for demand ought to be expressed in terms of the effect an action would have on the agent's integrity.

In light of this assertion, I examine Scheffler's use of an agent-centred prerogative to avoid the integrity problem and argue that, while agent-centred prerogatives are promising as a means to avoid the objection, Scheffler fails to adequately limit the scope of his prerogative. After a brief exploration of various possible maximising strategies, I offer an alternative solution which incorporates the importance of an agent's integrity into a consequentialist framework. This is achieved by combining a revised maximisation strategy with an agent-centred prerogative which utilises the process of reflective equilibrium to determine a set of protected integral projects which the agent may choose to protect over making things go best impersonally. I conclude that this solution adequately resolves both the integrity objection and demandingness objection to consequentialism.

Impact Statement

This thesis aims to give a solution to the demandingness objection for those wishing to defend a consequentialist perspective. If this task is successful, then this will have a large impact on the landscape of consequentialist ethics, owing to the prominence of the objection in consequentialist debate over the past few decades. In attempting to provide a solution, I also contribute two other significant arguments. The first is that the demandingness problem is derivative of the integrity problem. If this argument is widely accepted then it will point others in the direction I have taken and narrow the field of discussion. As a result of this argument, I also present an argument to resolve the integrity objection to consequentialism, another objection discussed almost as frequently as the demandingness objection. If my thesis is accepted, then it makes consequentialism, at least in the form I am suggesting, a much more palatable ethical stance. Within academia this will may well pave the way for a more favourable discussion of the theory, and outside of academia it may create the opportunity for consequentialism to be seriously reconsidered as a standard for which to hold people's actions and moral values to account. It could also lead to an uptake in consequentialism as individuals' guiding moral principle, which would likely lead to an increase in charitable donations and selfless actions. In fact, by the very nature of the thesis, if even a few individuals are convinced by the arguments presented herein and adopt my conclusion, then a significant amount of good will be done in the world that might otherwise not have been.

Table of Contents

1. Introduction

"Any plausible moral theory has difficulty in showing how agent's impartial moral commitments and their personal commitments can be harmoniously integrated" (Ashford, 2000, p.434).

While the above is true, no moral theory has been more dogged by its perceived over-demandingness than consequentialism. In fact the "demandingness objection" is frequently associated directly with consequentialism, despite the objection also bearing on deontological and contractualist theories. All moral theories place some level of demand on the agent, and those that are more theoretically rigorous have an even greater sense of demand attached, having more numerous and stronger restrictions and duties than common-sense morality. Consequentialism appears to have the greatest sense of demand attached to it, which is perhaps unsurprising considering the requirement to make things go as well as they possibly can.

In this essay, I attempt to defend consequentialism from the seemingly overwhelming strength of the demandingness objection. In the process it shall be necessary to draw a link with another problem plaguing consequentialism - the integrity objection. I argue that by recognising the link between these strong objections to consequentialism, it is possible to simultaneously solve both, using what the latter teaches us about the nature of the former. In doing so I   develop a consequentialist framework which draws on existing methods utilised by other theories that have a lesser sense of demandingness than consequentialism. The problem, I will argue, is partly axiological, partly methodological. I argue that we ought to revise the way in which we approach the relationship between individuals and what it means to make things go best, as well as the relationship between the agent and her integrity. In doing so I aim to elucidate a solution that refutes the claim that consequentialism is too demanding, while ensuring that one is always in a position to do what would make

things go best.

1.1. Consequentialism

Consequentialism can refer to any ethical theory which judges the rightness or wrongness of actions solely by the consequences they produce[1]. Consequentialism is therefore an act teleological theory, in that an action is permissible if and only if the action is maximally good. However, in order to make this view intelligible, an axiological view is needed; without a specific good or end to be maximised, it makes little sense as a moral theory and says nothing about how one would go about making things go best or which consequences should be produced.

Instead, there must be some end which ought to be maximised in order to bring about the best set of circumstances. Utilitarianism is a classic example of a consequentialist theory, arguing that "actions are right in proportion as they tend to promote happiness [or pleasure], wrong as they tend to produce the reverse"[2]. One issue to be taken with utilitarianism initially is that it is debatable that pleasure is the only true good for human beings. Many philosophers have contested this, offering examples of how pleasure alone cannot be the good that a consequentialist theory must maximise, with many thought experiments - such as Nozick's experience machine[3] -

---

[1] Sinnott-Armstrong, 2019

[2] Mill, 1962, p.257

[3] Nozick, 1974, p.43. Nozick postulates a machine in which one can plug in their brain, and live out a life of total bliss, ignorant to the fact that it was simulated. From the perspective of pure hedonism, this is an ideal situation; pleasure is maximised and pain minimised to the user. From the perspective of some basic utilitarian theories, this would be ideal should we have enough machines for everyone (or at least enough to ensure that the overall pleasure was increased). However, Nozick argues that this solution would not be palatable because we do in fact actually want to experience the world, and have meaningful connections with things external to us, not just revel in base pleasure. The force of this argument is debated but the fact that it has a clear and intuitive argument means that it goes at least some way towards explaining why there is resistance to the idea of a pleasure based approach to consequentialism.

highlighting that this is an oversimplified account of what gives a human life meaning. John Stuart Mill, perhaps the most famous utilitarian, recognised that mere pleasure alone is not to be strived for and that some pleasures, due to their more intellectual nature were to be considered higher pleasures and therefore more valuable[4]. However, rather than deciding on an axiological approach now, in this opening section I shall try to narrow down the definition of the form of consequentialism I wish to defend, without pointing to a specific axiology in order to allow scope for narrowing the definition further as the argument unfolds.

There are several features that make a consequentialist theory appealing which are included in the majority of consequentialist views. Firstly, consequentialist theories take the form of a maximising theory, such that the rightness of an action depends on what makes things go best. If a moral theory is concerned with consequences, then the most right action under this theory will be that which has the best consequences based on the axiology of the theory; if pleasure is the ultimate aim, then a consequentialist theory will argue that the best consequences, will be those that create the most pleasure.

Secondly, consequentialist theories are generally aggregative; the sum of an action's consequences, negative and positive, determine the overall rightness of the action. This is necessary because, without an aggregative approach, it is not clear how one could possibly ascertain whether an action actually had the best consequences, or whether it simply appeared to in isolation. In other words, it is necessary as part of the consequentialist calculations to take stock of how the agent's actions result in terms of the state of the world at large. The decision to purchase an electric car, for example, may save the agent money over time which could be donated to charity, as

---

Later arguments in this essay surrounding the importance of integrity to the individual may go some way to explaining this intuition in more certain terms (see chapter 5 in particular).

[4] Mill, 1962, pp.258-260

well as reducing the use of fossil fuels and therefore reducing harmful emissions, but if the car was made in a factory that did not respect the rights of its workers and was itself a significant polluter, then these facts also need to be considered in the agent's consequentialist calculations.

I shall also take it that consequentialism of the type I am aiming to defend here is universal, equal, and agent-neutral. It is universal because it considers the consequences as they effect all persons involved, equal in that all of those whom the consequences affect are treated with equal weight, and agent-neutral in that the rightness or wrongness of the action cannot be determined by the perspective of the agent, over and above any legitimate and properly weighted concerns that she may have that are affected by the consequences. These principles, as well as being inherent in the kind of reasoning behind having a consequentialist theory over a merely hedonistic theory, are intuitively held to be the best sort of principles for ensuring that we live in a just, fair, and morally decent society, and tend to map very well onto most people's moral intuitions.

1.2. Act Consequentialism

As well as the above criteria, there is one further distinction that needs to be elucidated: this essay specifically defends a form of *act consequentialism*. Act consequentialism, generally conceived, has the above qualities, but it has a further constraint: "an act is right if and only if it results in at least as much overall [good] as any action the agent could have performed"[5] or put another way: *an act is permissible if and only if there is no other act available to the agent that would make things go better*. Both definitions amount to the same constraint at the core of act consequentialism: one must only act in whatever way would make things go best, and no other action, no matter how close to the ideal in terms of making things go best,

---

[5] Eggleston, p.125

is permissible under this view. This constraint is crucial to understanding both the demandingness objection and the integrity objection; it is the impermissibility of failing to meet this rigid constraint which generates the idea that act consequentialism demands too much of the agent. In summary, nothing short of moral perfection on the part of the agent is permissible under this view.

Act consequentialism, like other forms, can function in two ways: as a deliberative tool for the agent to use when faced with a choice of actions and their consequences, or as an evaluative tool to be used to determine whether an act that has already been performed was right or wrong. In the former case, the act consequentialist would weigh up the available options based on what total universal consequences they had and choose the action that makes things go best, all things considered. Similarly, in evaluating whether an act was right or wrong, the options which were available to the agent at the time of choosing are judged based on their consequences and the act is deemed to have been right if and only if the action performed were the one which made things go better than they otherwise could have gone if another action were performed. If, in either case, the agent does not choose the action which makes things go best, then the action, no matter how well things did/do go, is morally wrong. To be clear, in all consequentialist theories, making things go best is done by maximising a chosen value or good. As yet, there is no clear reason to favour any particular value, so for most of this essay I shall refer to maximising the good as meaning the same thing as making things go best.

1.3. The Demandingness Objection to Act Consequentialism

The demandingness objection is not an objection to merely any level of demand being placed on moral agents. Morality requires that there is some demand on the individual, in that a moral agent, if following a moral code, will at some point have to sacrifice something, be it a good or a goal, in order to be acting morally. Therefore,

the fact that a moral theory demands some sacrifice on behalf of the agent is not in itself an objection to that theory, or at least not one that is likely to be found compelling. Such an argument would most likely not only beg the question against said theory, but also against the concept of morality in general.

The demandingness objection specifically takes issue with the extent to which a theory is demanding on the individual's self-interest. Consequentialism, as an impartial moral theory, is particularly susceptible to this objection, as it weighs all wellbeing, happiness, or whatever is the good to be maximised, equally among all concerned parties such that it may require the agent to sacrifice much of their goods or wellbeing in order ensure that the best possible consequences have been brought about, impartially speaking. This can be the case even if the agent has little to give and the net increase in utility is relatively little.[6]

As a result, act consequentialism can be shown to be "unreasonably demanding".[7] This problem is partially caused by there being no cut-off point in act consequentialism where an individual can reasonably reject the obligation to make personal sacrifices, so long as there is net gain in the measured good. Consequentialism, therefore, "demands that I ask how I can make my greatest possible contribution, all things considered—even though this may impose considerable hardship on me—and it forbids me to do anything less. If the claim is correct, most of my actions are immoral, for almost nothing that I do makes optimal use of my time and resources.[8]" This inherent demandingness is exacerbated by the extremely high level of preventable suffering that exists globally. As the world in which we live has a vast wealth disparity between the poorest and the merely well-off (let alone the richest in society), and there are simple and effective ways for us to

---

[6] Hooker, 2002, pp.151-3
[7] *ibid*
[8] Kagan, 1991, p.1

alleviate the suffering of those that are worst off, the requirement to do what would make things go best requires a seemingly enormous sacrifice for most individuals, owing to the readily available actions which can enable us to do so (Unger, 1996, p.133).

Under act consequentialism then, as long as there is sometimes a moral obligation to maximise the good, there is always an obligation to do so, regardless of personal cost; "where duty stops, so stops permissibility".[9] Thus, it seems that act consequentialism is entirely incompatible with the concept of supererogatory acts in general, demanding absolute sacrifice from as many as necessary in an attempt to increase utility.

However, some proponents of act consequentialist theories have indicated that while on an individual level act consequentialist theories seem to demand too much, if everyone were to become an act consequentialist overnight, then the actual demand placed on most individuals would be very low or at least much more reasonable.[10] As there are ample resources available in the world for all existing people to live comfortable lives, poverty exists solely because they are distributed improperly. If the very wealthy efficiently redistributed their wealth, there would be very little cause for those of average or low wealth to have to redistribute theirs. While this could be seen as a compelling argument for adopting act consequentialism as a political ideal, it does little to alleviate the demandingness problem for individuals adopting act consequentialism in a world where very few others do so. Arguments of this kind shall therefore be set aside and I shall continue by looking only at the demands on individuals.

1.4. The Project and Structure of the Thesis

---

[9] *ibid*

[10] See Murphy, 2003, Chapters 5 and 6, for a nuanced argument of this kind

The overall project of this essay is to defend consequentialism from the demandingness objection. I begin by giving strong reasons why we ought to attempt to defend consequentialism from the demandingness objection, highlighting to the reader what is at stake in the argument. I argue that the main reasons for defending consequentialism are its appeal to rationality, its simplicity, its concern for real-world states of affairs, and its impersonal nature. I then give some illustrations of the demandingness objection in order to highlight the extent of the problem.

After these preliminaries, I shall examine Liam Murphy's argument wherein he claims that what we object to when it comes to demandingness is not just that the demands are extreme, but that they are unfair. Most moral theories make extreme demands of us, but beneficence demands seem to be particularly demanding. I agree with Murphy's assertion that what makes consequentialism's demands unpalatable cannot be due solely to the perceived high levels of demand and that it is necessary to establish a baseline of demand if we are to account for our intuitions regarding demandingness. However, I argue that Murphy's suggestion that the cause of these intuitions are rooted in the concept of fairness is ultimately flawed in that it incorrectly makes use of only one particular sense of fairness - the fairness of the distribution of responsibilities.

Having accepted Murphy's argument that the demandingness objection cannot be derived merely the high level of demand in itself, I shall then argue that it is actually a derivation of the more insidious integrity objection to consequentialism. Many argue that the integrity objection is a derivation of the demandingness objection, however I argue that this view is mistaken, and that the reverse is true. I argue that what is intrinsically valuable to us are our projects, partial relationships and life goals, and that if these projects are the source of meaning and worth in our lives, then the resources and wellbeing we are being asked to give up by consequentialist theories

are only valuable in their ability to help us satisfy these more foundational goals.

I then examine Samuel Scheffler's attempt to defend consequentialism from the integrity objection by including an "agent-centred-prerogative" which allows for the agent to dedicate their time and resources to their own personal projects and preferences out of proportion with their impersonal value. The amount of demand that can be placed on the individual is thereby limited to prevent the individual from losing a sense of integrity. I argue that while the use of the prerogative has potential, Scheffler does not offer a satisfactory alternative to more traditional consequentialism owing to his deviation from the fundamental tenants and core ideals of consequentialist theories, and his inability to properly respect the integrity of agent.

Responding to Scheffler's rejection of maximisation strategies, I briefly examine several purely maximising strategies which aim to protect the agent's integrity without appealing to an agent-centred prerogative or restriction. I argue that while there is a form of maximisation strategy that comes close to offering a solution to the integrity objection, that any purely maximising strategy is ultimately incapable of solving the integrity problem owing to its inability to bridge the gap between the agent and her own integrity caused by consequentialism's requirement that she treat her own projects as no more important to her than those of any other individual.

Finally, I return to the idea of agent-centred prerogatives to present a novel solution to the integrity problem, suggesting that the only way for consequentialism to defend itself against the demandingness objection and the integrity objection is to adopt an agent-centred prerogative which utilises reflective equilibrium as the method for determining the baseline of demand for each agent. I argue there is a special set of projects for each agent which make her life meaningful and worth living, and that she can discover what this set of projects is through reflective equilibrium within a

consequentialist framework. By doing so, we can ensure that the agent has a protected set of projects, which she cannot be required by consequentialism to abandon, but also that she limit this set to only those projects which, upon reflection, are integral to her sense of self. By requiring constant engagement in this process of reflection, the agent ought to eventually arrive at a state whereby she has as minimal a set of integral projects as required to both respect her integrity and allow her to act within a consequentialist framework which is still highly demanding.

2. The Plausibility of Consequentialism

A primary appeal of act consequentialism is its rational nature. At the level of the individual, if an agent is given multiple choices of actions to perform, one of which creates a situation that is better for the individual than the other two actions all things considered, then it would clearly be rational for the individual to perform that action over the other two. It would, by contrast, be irrational if the individual chose to perform an action which made them worse off than they otherwise could have been, given no external reason for doing so. For example, given the choice to take a medicine that would cure a terminal illness or one that would only extend a life by months, it would be irrational for a terminally ill patient to take the medicine which would only prolong her life by months (all things considered and with no special reasons for doing so). Given that it is rational for a single agent to act in a way that would make things go best for them, it therefore makes sense that a moral theory ought to require that we act in a way which is rational from the impersonal perspective by choosing the actions available to us which would make things go best impersonally. Put another way, it is rational from a personal point of view to choose to bring about the state of affairs that is best for the individual and therefore rational to bring about the state of affairs which is best impersonally from the impersonal moral point of view.[11]

Another strong reason to defend consequentialism, brought about by its appeal to rationality, is its simplicity. In most moral deliberations, the actions and their consequences will be immediately apparent to us. If I donate £10 to a malaria charity, I will be responsible for saving one life, if I cycle rather than driving to work my carbon footprint will be reduced but I will be more tired, and so on. All we have to do at that point is weigh up the costs and benefits based upon the good that we value (e.g. the

---

[11] Eggleston, pp.133-4

wellbeing of the concerned individuals) and we have an answer ready for us. Of course, some situations may be more complicated than others to discern which action would have the best consequences, all things considered. Other theories have a seemingly much more complex method to moral decision-making. Kantian deontology, for example, requires that we must ensure that we can will that act as a universalised maxim, and that in doing so we are not willing in a way which is internally inconsistent, or inconsistent with other maxims which it is rational to will. Are there any inconsistencies with willing that everyone give £10 to a malaria charity? Might this be inconsistent with a maxim that willed that everyone ought to donate as much money as possible to ending world hunger? To some the answer to this may seem obvious, but I admit that to me it does not.

If we want the world to be a better place - and surely a morally minded individual would - then actual results brought about in the real world are important. Consequentialism is concerned with exactly that goal, ensuring that our actions bring about the best possible real-world consequences in order to make the world a better place. Contrary to this, deontological theories are more concerned with ensuring that our actions occur for the right reasons. If one could ensure that everyone is able to live a healthy and moderately long life by enforcing non-discriminatory population control for example, a deontologist might take issue due to its not treating individuals as ends in their own right, no matter how beneficial the outcome might be for the wellbeing of people throughout the world. I am sure most deontologists would likely argue that a global adherence to a deontological theory would in all likelihood bring about the best possible world. However, if this is not the primary concern of the theory, then what faith can we really have that it is tending towards it regardless.

One final and important reason to defend consequentialism is that consequentialism is an impersonal theory. It is impersonal and impartial in that it is not up to the

discretion of the individual making the decision as to whether they ought to do what would make things go best. There is therefore no scope for the individual to refuse to act in accordance with maximising a good purely because they find the methods of doing so distasteful or inconvenient. There is no opposition from a personal level allowed to affect our moral reasoning, nor are we allowed to prioritise certain individuals in our reasoning. That is not to say that the deliberator or their loved ones and acquaintances do not factor into the moral deliberation, but they have no more a stake in the deliberation than any other individual concerned. Thus, consequentialism demands selflessness in moral deliberation, and does not in any way accept or tolerate selfishness and preferential treatment.

Counter to my earlier point about the intuitiveness of consequentialism, to many this will seem counter intuitive. Why not put myself or my loved ones first, why not ensure that I do all I can for the people I care about, rather than concerning myself too heavily with the lives of strangers? We form strong connections with individuals purely because we prefer these people over others, so why not let our actions reflect that? As set out above, this is the primary concern of the demandingness objection which I have taken to be a serious threat to consequentialism, so why now defend the very element of the theory which generates the problem I am trying to address?

Later in this essay I show how the impersonal aspect of consequentialism can be mostly preserved in order to defend against the demandingness objection, but for now, I will explain why I believe impartiality must be preserved as much as possible in consequentialism. Simply put, the best consequences cannot be brought about by favouring those that we cherish over all others, or by putting ourselves first the majority of the time; this has been borne out empirically as the current status quo. What seems like a perfectly reasonable desire to provide the best possible lives for our families and loved ones, motivates and explains gross wealth disparity and abuse of resources on a scale that ought to be considered unreasonable. Billionaire tax

avoiders do not necessarily hoard their money and underpay their employees because they are evil, or believe that the annual deaths of thousands of children in underdeveloped countries from malnourishment and easily preventable diseases is to be expected as just a standard facet of existence. They do so   believing that it is right to protect and provide for their loved ones as a priority over all others and they simply have the capacity and ability to take this reasoning to the extreme. The impartiality of consequentialism prevents this outcome better than any other theory by ensuring that the reasoning that enables greed is cut off before it can begin. The fact that it demands too much of us in the process is an unfortunate side-effect brought about by the fact that not everyone adheres to consequentialist principles[12]. As Liam Murphy notes in one of his discussions surrounding the limits of beneficence: "we have one clear ground for the *suspicion* of the belief that there is a limit to the demands of […] beneficence: such a limit is much more in the interests of the best-off than the worst-off".[13] This thesis then, is primarily an attempt to maintain as much as possible of the impartial nature of consequentialism while solving the demandingness objection.

At this stage it is worth briefly addressing the argument that perhaps pursuing distributive justice through the reform or our political institutions, rather than through the propagation of consequentialism, would be the optimal ethical approach to ensure that poverty and wealth disparity are remedied. The argument might run that - accepting that there are enough resources available globally in order to ensure that no individual suffered from hunger, or preventable ill-heath or premature death - focusing our efforts on creating political institutions to ensure that these resources are distributed properly (i.e. in such a way as to ensure a certain quality of life for all, but not necessarily equally) would bring about a better world than if individuals

---

[12]   Liam Murphy uses this line of reasoning in an attempt to resolve the demandingness problem, as discussed in chapter 4.
[13]   Murphy, 2003, p.15

focused on abiding by consequentialist principles. Another way of looking at it would be to argue that a collective effort to redress the causes of suffering, rather than individual efforts to treat the symptoms, would be more effective and more efficient as an ethical theory.

To this line of argument my response is that consequentialist principles, particularly those that I argue are at the core of the theory, are not mutually exclusive with pursuing distributive justice through political reform. If pursuing distributive justice through political reform would make things go best, then it would fall under the responsibility of the consequentialist to follow these principles (alongside other actions that consequentialism might prescribe). In other words, from the perspective of the individual, if they believed that what would make things go best is ensuring that distributive justice is enacted through political reform, then this ought to figure into their consequentialist calculations informing what actions they ought to take; if it would do more good to campaign for distributive justice than to work in a charity shop, then they ought to do the former. However, it is unlikely that, generally speaking, followers of consequentialism would devote large periods of time to push for distributive justice through political reform. The main reason for this is that in all likelihood the average individual is going to make very little difference to the policies of governments and public bodies that would be able to drive the change needed, and therefore most individuals would be able to do more good by focusing their time on more direct action. This does not mean that they would be prohibited for pushing for large scale social reforms with distributive justice in mind - indeed it is highly likely that they would support political parties with policies that favoured distributive justice. It simply means that, rather than allowing the pursuit of distributive justice through political reform to become their overarching aim, a tacit support for political reform will be manifested by the consequentialist's political interactions.

3. The Demandingness Problem in Action

Susan Wolf argues that there is something problematic about the nature of moral theories such as consequentialism which demand moral perfection, and describes the people who meet, or strive to meet these moral ideals as moral saints.[14] She argues they are deeply flawed individuals, who may well be too good for their own wellbeing, owing to a lack in the kind of qualities that we may see as being crucial to a well-rounded human being.[15] The problem with those seemingly striving for moral sainthood is that they lack elements of human lives that make lives good to begin with,[16] such as friendship, appreciation of the arts or culture, or a genuine love of a hobby, activity, or individual which goes beyond an appeal to the utility value of the act. Moral saints do not live what we might call a good life from the "point of view of individual perfection"[17] as they do not contain or recognise the value of obtaining ideals which we ordinarily see as good from a non-moral perspective, owing to their incompatibility with moral sainthood.[18] Consequentialism's failure to account for this is what drives some of the intuitions we may have regarding its demandingness. If, by recognising the values of ideals, personality traits, and character refinements that are outside of the moral sphere, we have good reason to want people to live lives that are not morally perfect, then we have good reason to believe that any moral theory will make use of some conception of supererogation.[19]

---

[14] Wolf 1982
[15] *Ibid*, p.421
[16] *ibid*, p.436-7
[17] *ibid*
[18] *ibid,* p.426
[19] *ibid*, p.438

Consider the following example given by Roger Crisp[20] wherein an agent, Anna, is given the choice between three different lives:

> "A: A life in which she has flourishing personal relationships, plenty of leisure and a lucrative career, and does nothing for charity.
> B: A life in which she has all of the above goods, though not to the same extent because of the fairly large amounts of time and money she donates to Oxfam.
> C: A life in which she has none of the above apart from the career, which she pursues solely in order to give all her spare money to Oxfam." (Crisp, 1996, p.62)

The consequentialist would clearly point to option C as the correct option. However, as Crisp argues, it is plausible to think that Anna may choose life B and be right to do so. She cannot do so however, by an appeal to an evaluative judgement from an agent neutral position as consequentialism requires because, as we have seen, the consequentialist will retort that she could have done more good overall by choosing option C. If there is plausibility in Anna's choosing B over C, there must be some reason for doing so that is non-evaluative in order for her to fall prey to the evaluative reasoning that generates the high demands of consequentialism.

> "It must be that these goods are instantiated in her life. The fact that the life is hers and that therefore the goods in life B will be hers provides her with a counterbalance to the reason to promote the good which justifies and explains her selection of that life over life C. [...] I would ask those who are not persuaded by this suggestion to ask themselves the following question: is Anna being unreasonable in choosing life B over life C?" (Crisp, 1996, p.63).

---

[20] Crisp, 1996, pp.62-3

That there are aspects of an individual's life that may be worth sustaining for reasons other than an appeal to utility will be the subject of debate from chapter 5 onward wherein I address the importance of the personal point of view in maintaining a life of integrity. However, it is worth quickly addressing some of the arguments that Wolf appeals to here in defence of supererogation within a moral theory. Firstly, it is open to the consequentialist to argue both that consequentialism recognises the value of these seemingly non-moral ideals, but gives them the appropriate weighting when considering where they fit in the grand scheme of making things go best. It might be the case for example, that it does make things go best for someone to become a virtuoso cellist, as their skill may result in more good being done than otherwise would have been. This, I believe, [21] is a weak defence on the part of the consequentialist particularly when faced with a world in which there is extreme poverty and we cannot expect full compliance with consequentialism. However, a more effective response is that while these non-moral ideals have value, they clearly have an inferior value to that of basic wellbeing, or curing disease, or ending starvation, and so on. Once a consequentialist "saint" has done away with these more pressing needs, the next step towards making things go best might well be to focus on maximising flourishing in the arts.

Consequentialism, then, is demanding in several ways. Firstly, it requires that we sacrifice a potentially enormous amount of personal wealth in order to make things go best impersonally. Even for those that do not have an enormous amount of wealth, it still requires that we sacrifice as much wealth as possible and therefore makes high demands of even some of the poorest in society. Also, because it does not allow for agents to place greater weight on their own preferences, it requires that personal projects, luxuries, and indulgences are abandoned by the agent if doing so would maximise the overall good. This could be something as simple as foregoing a holiday, or as trivial as purchasing books from charity; no matter what the project, if it would

---

[21] as does Wolf, 1982, pp.427-30

not make things go best, then it is not morally acceptable under act consequentialism. Another issue is that, owing to the impersonal nature of the theory, it does not allow for extra weighting being given to loved ones; even an entirely selfless person is acting in a morally reprehensible way under consequentialism if they promote the welfare of their loved ones out of proportion with the welfare of others.

A final problem for consequentialism is that it requires not only that we act entirely selflessly, without any bias, and entirely from the impersonal perspective, but also when we do so, we do so as efficiently as possible. Meaning that, even if one dedicated their lives to ending world hunger (generating as much income to donate to anti-hunger charities as possible and using all of their free time feeding the poor), if it were the case that more lives could have been saved by focusing on preventing malaria deaths, then the agent would still be judged from the consequentialist view to have been acting immorally. However, even if one were to live the perfect consequentialist life, they would only ever be able to be deemed to have lived a morally acceptable, or morally good life. There is no scope within an act consequentialist framework for supererogatory acts. In other words, it is impossible for one to have gone above and beyond the call of duty when it comes to consequentialism. Because the only right thing to do is perform the very best action - all other actions being morally deplorable - there is no scope for some actions to be more right than others, as in the case of common-sense morality and other theories. Moral sainthood, as Wolf phrases it, gives too much credit in cases such as this - consequentialism admits of no moral saints, only moral adherents at best.

4. Liam Murphy and the Baseline of Demand

It should be clear by now that we are not discussing any and all demands on the individual by a moral theory; morality makes demands on moral agents by its nature. The demandingness objection refers then to "extreme demands" — those that require us to sacrifice some set of goods that it is either unacceptable or impossible for us to relinquish and still live what we might think of as an acceptable life. One question that remains is by what measure are these demands extreme? Acceptability of either the demands themselves or the life that we would lead should we adhere to them will depend either on the personal views or the intuitions of particular agents, and therefore, there will be some dispute over the limits of moral demands. It could be argued that the limit of demandingness is set by our intuitions regarding acceptable levels of demand. However, as each person has very different ideas regarding the acceptability of demands, which are most likely independent of any theoretical grounding, then to accept that demandingness can be limited by our intuitions it seems that there ought to be some plausible logical grounding for these intuitions in order that we can logically justify them against consequentialism. Therefore, a suitable objective threshold for acceptable demands must be established by those wishing to argue that the demandingness objection poses a significant threat.

4.1. Is the Demandingness Objection Question Begging?

Liam Murphy[22] notes that there are various theories which place what could be seen as extreme demands on the agent, yet these theories tend not to be strongly associated with the demandingness problem. This is in part due to the fact that these demands seem to coincide with our common-sense morality. Kantian deontology, for

---

[22] Murphy, 2003

example, makes the *prima facie* acceptable demand that we not become a professional bank robber, performing daring heists purely for monetary gain. While this seems to be an intuitively reasonable expectation, it could be seen, at least from a consequentialist perspective, as an extreme demand. Our lives might be much better off given the money that robbing banks provides and therefore removing this money and the opportunities it affords can be seen as an extreme demand despite the fact that it appears to make very little demand on the average individual. Equally, supposing an individual were as wealthy as our hypothetical bank robber, it would appear intuitively overly demanding to many to require her to give up the sum of money equal to the difference between the wealth of an average individual and a successful bank robber. This being the case, another justification is necessary for thinking the demands of consequentialism to be extreme in a special sense.

There appears then, to be something about extreme beneficence demands in particular that people find objectionable. This presents a potential counter argument to the demandingness objection to consequentialism: Without an extra, special account of why beneficence demands are too demanding, it seems too much is at stake for us to accept that extreme demands can be a legitimate concern for moral theorists, considering the *prima facie* legitimate claims they make on individuals in the majority of cases.[23] Thus, it appears there is a serious explanatory gap that it is necessary to bridge if we want to say anything meaningful about the demandingness of moral theories, and the particularly extreme levels of demand that act consequentialism implies which other theories seem not to. This gap consists of exactly what it is that makes beneficence and other self-sacrificing theories appear extremely demanding, while other demands which reduce the agent's wellbeing by equally large amounts seem less so.[24]

---

[23] *ibid*, p.39, p.60
[24] *ibid*, pp.45-7

A potential solution would be to point to the fact that consequentialism and other optimising beneficence theories have an active rather than passive demand. This can be seen by contrasting the seemingly common-sense notion that we ought not to rob banks and that it is morally permissible to not donate to charity an equal sum of money as one would have had from robbing the bank; the difference here is that not robbing banks is a passive demand on our morality, whereas a beneficence demand is an active demand. Rather than being prevented from taking certain courses of action that might have made us better off, we are required to actively pursue courses of action which will make us worse off. This distinction seemingly draws out a difference between beneficence demands and most other types of extreme demands. However, this distinction between active and passive demands is not as clear cut as it seems: beneficence demands are both active and passive. Beneficence demands are not just giving up things we already have, but it may well passively prevent us pursuing avenues that we might otherwise have chosen. For example, when a theory requires us to donate to charity we are actively being asked to give up our income, but we are also being passively prevented from accruing any luxuries for ourselves.

4.2. Making Sense of Extreme Beneficence Demands

As a minimum baseline of demand appears to be problematic, Murphy suggests an alternative grounding for the demandingness objection: consequentialism generates an absurdly extreme demand because it violates what he refers to as the "compliance condition".[25] Supposing a real-world wellbeing disparity, it is fair to say that many people in the world are capable of sacrificing some of their wealth in order to address that disparity, without making themselves worse off than those they would seek to help in their sacrifice. In fact, there is so much wealth available amongst those people that could make the sacrifice, that in actual fact there need only be a proportionally

---

[25] Murphy, 2003, p.77

low sacrifice from each individual in order to make things significantly better for the worst off in the world. Suppose then that everyone who was able to make a sacrifice did so, and they gave an amount proportional to their wealth such that we might argue that the necessary sacrifice was distributed *fairly*[26]. This is referred to as a full-compliance scenario; each individual is taking on their fair share of the collective responsibility to help those who need it.

Under a full-compliance situation, while sacrifices might still be seen as extreme given the demands of the optimising principle of beneficence, they are fairly distributed among those who ought to take on the responsibility. Unfortunately, however, we do not live in such a world, and in reality many people do not take on their share of the responsibility and sacrifice. This means that those acting under the optimising principle of beneficence are expected to take on the share of the responsibility that has been shirked by those who are non-compliant. This is intuitively problematic for us, Murphy argues, not because it has surpassed some sort of baseline of demand, but because the situation is inherently *unfair.* The optimising principle of beneficence requires that one ought to do what would make things go best by maximising the overall wellbeing in the world given the real-world circumstances, treating the hypothetical but possible world of full-compliance as irrelevant.

The optimising principle of beneficence seems to generate intuitively unacceptable levels of demand because it violates Murphy's *Compliance Condition*:

> "Agent-neutral principles should not under partial compliance require
> sacrifice where the total compliance effect on her, taking that sacrifice into
> account, would be worse than it would be (all other aspects of her situation
> remaining the same) under full compliance from now on." (Murphy, 2003,

---

[26] I shall not go into how Murphy believes a fair distribution under full compliance may look here, as it is not relevant to the discussion.

p.80)

The compliance condition is a marker for whether a moral theory places too high a demand on individuals by essentially ensuring that the theory does not require one to do more than their fair share. We can measure whether one is having to do more than their fair share by asking whether they would be worse off under partial compliance than with our chosen theory than under full compliance. Act consequentialism violates this condition in real-world circumstances, by requiring that one do all they can to maximise the good, given the actual state of affairs in the world. This therefore requires that we bring our level of expected wellbeing below what it would be in the full-compliance world in order to ensure wellbeing maximisation. The inherent unfairness in having to do more than one's fair share when others fail to do theirs therefore makes consequentialism and the optimising principle of beneficence unpalatably demanding.

Murphy argues that what is necessary is a *"non-ideal* collective principle of beneficence", one which takes stock of the actual world in its calculations towards decision-making.[27] The collective principle of beneficence requires not that the agent acts as if under full-compliance, but that she promotes the maximisation of wellbeing as much as possible without bringing her wellbeing lower than it would otherwise have been under full compliance, or as Murphy phrases it: "a person need never sacrifice so much that he would end up less well-off than he would be under full compliance from now on, but within that constraint he must do as much good as possible"[28]. This formulation manages to effectively incorporate the key aspects of the optimised beneficence principle by requiring some form of maximisation, while placing a limit on the responsibilities and demands placed on the agent. While the full-compliance model is still imagined, because we are acting by taking into account

---

[27] *ibid,* p.86
[28] *ibid,* pp.86-7

real-world circumstances, we are not acting in a way that is entirely out of kilter with the demands of the real world.

Fairness is at the heart of the collective principle not in the way in which the responsibilities are distributed, but in the way in which the effects of the responsibility distribution are shared. I have talked so far about distribution of responsibility and the unfairness this causes to those who adopt an optimised principle of beneficence in a partial compliance situation, but the fairness or unfairness of the situation is to do with the distribution of effects of partial compliance on the agent - the loss of their wellbeing compared to full compliance. The distribution of responsibility under partial compliance is already unfair - some people who ought to be complying are not, while others are - regardless of whether complying agents take on the extra burden of the shirked responsibility. This may be less unfair under the collective principle of beneficence, but it is unfairness that we are concerned with, not the amount. What is an avoidable unfairness in the case of partial compliance is that the wellbeing levels of the compliers suffers compared to full compliance.[29]

Murphy's argument here is about the unfairness of the distribution of demands, not about the increase in demands as such. In a situation wherein under both full and partial compliance the demands were equally exceptionally extreme (Murphy gives the example of a "lucky" and "unlucky" country), this does not seem to elicit an intuitive objection over over-demandingness, as in both cases the demand is the same; there is no unfairness at play. The absurdity of the demandingness problem only comes into play once those acting under partial compliance are required to give up more than they would otherwise have to. A parallel may be this: suppose there is a tsunami that devastates the entirety of North America, leaving millions of survivors,

---

[29]   This focus on the fairness of resource distribution will later become an issue for the collective principle of beneficence.

but no resources at all for them to exist on. The rest of the world's nations agree to harbour refugees and provide a large percentage of their GDP to the relief effort. The distribution of these donations is decided by the wealth of the individual nations so that richer nations pay more, and poorer nations less, and is unanimously agreed. Every citizen of every country around the world feels the brunt of this decision, becoming considerably poorer and losing many if not all of the comforts and luxuries they previously had. However, none disagree to the terms, arguing that it is the right thing to do and they are only doing their fair share. This is the essence that is meant to be captured by the collective principle of beneficence; no matter how high the cost, and to how many people, the demands placed on us by principles of beneficence cannot be absurdly demanding unless they are unfair.

4.3. The Unfairness of Fairness

Murphy outlines a possible objection to the collective principle of beneficence available to those who are unconvinced by his use of fairness as an explanation for the intuitive feeling of being over-burdened by demand:

> "It could be objected that the collective principle of beneficence is even more unfair because it lets the cost of non-compliance fall on the people who should have been benefited by the noncompliers. Those 'victims' of noncompliance are likely to be worse off than the compliers whom the optimising principle of beneficence requires to take up the slack, and we would normally think it fairer to let some of the cost fall on the better off of two people." (Murphy, 2003, p.92)

Murphy dismisses the objection he raises, arguing it "assimilates a concern with the fairness of the way a principle imposes responsibilities on agents to a general concern

about the fairness of the distribution of wellbeing[30]". The confusion here rests on the difference between what is unfair due to distribution of wellbeing, and what is unfair in terms of requirements to sacrifice wellbeing: compliers in the partial compliance situation are being treated unfairly due to the responsibility that is being forced upon them which otherwise would have been burdened by the now non-compliers. The apparent unfairness to the 'victims' of partial compliance occurs only because under full compliance they would have benefited greatly, whereas under partial compliance, they do not.

However, whether intentionally or not, Murphy has played a clever trick here. In describing those that would benefit most from a full compliance scenario as the "victims of noncompliance", he has shifted the blame of their loss specifically to those who do not comply. However, the optimising principle still holding true would mean that even in a partial compliance situation, these people would receive the benefits of full compliance. The difference between their state of wellbeing then, is not the amount of compliance, but the following of the collective principle of beneficence. The optimal situation, from a consequentialist standpoint, is that the level of wellbeing is increased to its maximum levels, given the available options to do so in the real world. Accepting that the full compliance world is not the existent situation and that instead we live a partial compliance world, the thing that ought to be done is that the compliant among us follow the optimising principle. Supposing a situation in which five people needed to be untied from train tracks to avoid an oncoming train, and five people happened to be passing. There is enough time for any one person to untie all five, and yet under the compliance condition each of the agents would only be morally required to save one person, even if the other agents did not act. Under the compliance condition then, what seems like a morally required act under most plausible moral theories, is relegated to a supererogatory one, and we would not be able to place moral blame on an agent that left four people to die unnecessarily. This

---

[30] *ibid*

highlights two levels of unfairness: it is unfair that those shirking their responsibilities, morally requiring others to take it up, and it is also unfair that those who would have benefited from the optimising principle being followed are no longer receiving that benefit. They are therefore, the victims of the collective principle, not of partial compliance. If consequentialism requires that those in the partial compliance situation take up the slack, then they are shirking their duties also in following the collective principle.

Still, Murphy holds that the unfairness in these two situations is of two different kinds. The unfairness that concerns Murphy is that people are being required to take on extra responsibilities: "though the collective principle of beneficence leaves the victims of noncompliance worse off than they would be if the compliers took up (some of) the slack, it cannot be said that the victims have been required to take on (either actively or passively) responsibilities that belong to others." (*ibid*) This seems true enough. It is hard to argue that under partial compliance the victims of the collective principle have been made to take on extra responsibilities.

However, this focus on responsibilities is an odd tactic bearing in mind that Murphy spends some time explaining that what really matters is not that the distribution of responsibilities is fair *per se*, but that it reflects a fair distribution of the effects of compliance:

> "What matters are the effects on people's well-being of compliance with the agent-neutral principle; this is what needs to be fairly distributed [...] a fair distribution of responsibility is a distribution of responsibility that affects the well-being of the members of the group fairly: what we look for, in fact, is the distribution of responsibility that yields a fair distribution of compliance effects." (Murphy, 2003, p.90)

If it is wrong for someone to be worse off under partial compliance than under full compliance, then as an agent-neutral theory this ought to apply to all. If we are concerned with the loss or gain of wellbeing between a hypothetical optimal scenario and a real-world scenario, then arguing that responsibilities increasing is the desiderata for whether the theory is fair or not is irrelevant. The optimising principle recognises that while the best possible outcome is full compliance, in the partial compliance situation the way to make things go best is for remaining compliers to pick up the slack. If the collective principle is motivated by fairness in terms of loss of wellbeing compared to a full compliance model, then the loss of wellbeing of the victims of the collective principle ought to be considered as part of making a decision surrounding fairness. The move to increased responsibility burdens as a way of evoking a special sense of fairness seems to violate any remaining sense of agent neutrality; in effect the move here is to make the wellbeing loss of the agents more important by virtue of them being the decision-making agents. To use Murphy's phrasing here: if we are looking for the distribution of responsibility that yields a fair distribution of compliance effects, then the optimising principle is fair in that it creates the minimal loss of wellbeing between the full compliance situation and the partial compliance situation while maintaining agent-neutrality.

To illustrate this point, imagine a variant on Singer's drowning child analogy: [31] Walking past a pond on your way to work and in the centre of the pond you notice a child marooned on a small island. The child calls for help, as they cannot swim and the water levels are rising rapidly. Wearing an expensive suit, you recognise that due to the state of the pond any attempt to rescue the child alone would ruin your suit, but would certainly save the child's life. You also notice there are a dozens of onlookers whom, coincidentally, are also wearing equally expensive suits. A plan forms in your head: you have just enough time for everyone in your group to remove their ties, tie them together and throw this to the child to grab hold of, allowing you

---

[31] Singer, 1972

to pull the child to safety. Everyone agrees that your plan would work. However, for one reason or another, no one appears willing to risk their ties. Recognising that the water level is rising and there is little time to argue with the crowd, you decide to take the plunge and swim over to rescue the child yourself, ruining your suit in the process.[32]

4.4. Fairness Isn't Everything

The above example demonstrates there are cases where the collective principle under-motivates actions that both common-sense morality and the optimising principle prescribe. I will now turn the other edge of the blade to Murphy's principle. Murphy argues that, all other things being equal, even very extreme demands do not trigger the same intuitions regarding over demandingness if the burden of demand is distributed fairly:

> "Imagine a world where one small country is very well-off, and the rest of the world is very badly off […] Both the optimizing and the collective principles of beneficence would impose extreme demands on all the residents of the lucky country. The difference between this case and our actual case of world poverty is that the extreme demands are not in any part due to agents being required to take on more than their fair share of the demands of beneficence. […] the question is whether such a requirement qualifies as absurd for the reason that its demands are absurdly extreme." (Murphy, 2003, pp.100-101)

Murphy argues that there is no intuitive feeling that the burden of demand is overly extreme because there is no unfairness at play. It is absurd to expect any member of the compliant group under partial compliance to act in accordance with the

---

[32] A similar example can be found in Blincoe, 2017, pp.98-9.

optimising principle, but it is not absurd for everyone to act under the optimising principle under full compliance. Thus, it is not absurd in this case for the compliant under partial compliance to follow the collective principle. In neither the case of full compliance, nor the collective principle under partial compliance are our intuitions regarding overly high levels of demand triggered according to Murphy. This lack of intuitive repugnance leads Murphy to argue that "fairly imposed extreme demands do not stimulate the same confident negative reaction that unfairly imposed extreme demands do.[33]"

I am inclined to agree with Murphy in that I also find the level of unfairness in scenarios such as these to leave a bitter taste in the mouth. However, I believe there is something missing in Murphy's examples, in that he has failed to give a proper account of what has been lost. It is easy to focus on unfairness in the distribution of burdens as being what compels us to reject the optimising principle when we do not know exactly what it is we are losing. When we think of what is at stake merely as "extreme demands", it loses the persuasive force that a more precise description might have.

Consider the following: Surviving a shipwreck on the S.S. Jeremy Bentham, and marooned on an uninhabited island with the ship's navigator and a survival expert, the three of you manage to eke out a somewhat comfortable consequentialist existence for several years while waiting for help to arrive. The ship's navigator believes that the island is on no map, and it is therefore an incredibly small chance that you will ever be rescued. While grateful to be alive, you all miss your rich and fulfilling lives back home, particularly your loved ones. One day while on a routine trip to forage for cliff-bird's eggs, the survivalist loses their footing and falls. Fortunately both you and the navigator manage to grab a hand each, but the survivalist has been knocked unconscious. While struggling to raise yourselves and your unconscious

---

companion up the cliff, you both hear a ship's horn from the opposite side of the island. The only way to gain the ship's attention now is to light a beacon on the shore nearest the ship before the ship passes by. However, to save the survivalist's life, you both need to maintain your grip or else they will certainly plummet to their death. The choice is that of saving the life of the survivalist - who has done at least their fair share of work while marooned and deserves salvation as much as anyone - or to abandon them to their death while the remaining two of you signal for help and return to your rich and fulfilling lives, surrounded by loved ones and home comforts. The burden of responsibility falls on both of you fairly, and given that it was a freak accident that caused the scenario, it could have been any one of you who was now dangling unconscious below. Given that you could all survive on the island indefinitely if you saved the survivalist's life, the choice is now between maintaining three somewhat comfortable but ultimately unfulfilled lives, or sacrificing one life so that two people can lead rich and enjoyable lives, full of the sort of things that give our lives meaning and purpose. As the two of you consider the situation and ultimately decide that saving your friend's life is objectively the right thing to do, neither of you can help but feel as though you have somehow been cheated of the only opportunity to live a meaningful life, and that the consequentialist calculus has been overly demanding in this case.

While an extravagant example, it highlights a simple point: even when the burden is distributed as fairly as possible in a situation, if the demands require us to give up those things which we find most valuable they will still intuitively seem overly demanding. Requiring that we give up some luxuries in order to make other people significantly better off does not seem as objectionable as being required to give up one's chance at a meaningful life to preserve the life of someone who will also lead an impoverished life. My argument here is equally simple: Murphy is plainly wrong to think that fairness always motivates our sense of over-demandingness, even if there are cases where our sense of fairness and our sense of over-demandingness overlap.

Murphy's argument regarding the unfairness of the optimising principle in partial compliance situations is compelling but ultimately overextends its reach. It relies on describing the demand in an impersonal and unquantified manor, such that the only aspect of the example that one can intuitively latch onto is the unfairness, simply because we do not know what we have lost in the arrangement other than our equal regard.

4.5. Still No Baseline

Two conclusions can be drawn here: Firstly, it is not sufficient to say that consequentialism and other beneficence theories are too demanding simply because the demands are extreme. Extreme demands exist in almost all moral theories, and yet there appears to be something in particular about beneficence-based theories that generate strong intuitions regarding overly high levels of demand. If we are to accept that the demandingness problem is a genuine and uniquely difficult problem for consequentialism, then we need to be able to point to a particular feature of consequentialism that makes the demandingness problem particularly virulent.

Secondly, we can reject Murphy's argument that the unfairness of consequentialism under partial compliance is what motivates our intuitions regarding demands. I have argued that fairness is under-motivating when it comes to preventing us from helping others. If one believes that we have a responsibility to do as much good as we can in a given situation, then the apparent indifference of others ought not to override this responsibility. Murphy overemphasises the importance of fairness and incorrectly claims that extreme demands do not seem intuitively unpalatable when the burden of demand is distributed fairly. If one considers the demands as giving up particular assets, goals, ambitions, or interests, then it is easy to imagine situations whereby the unfairness of the situation is trivial compared to the overall loss to the agent.

Murphy has therefore, been able to present us with a difficulty for the demandingness problem, but not a suitable explanation of why that problem exists. There is clearly an intuition shared by many that consequentialism is overly demanding, but there is as yet no obvious baseline for when the theory moves from demanding to over-demanding. What is needed is a baseline of demand as well as a strong motivation for that baseline. In the next section I shall argue that a baseline can be found by appeal to another problem faced by consequentialism: the integrity problem.

5. The Demandingness Objection as a Component of the Integrity Objection

One way to underpin the demandingness objection would be to appeal to an intuitive sense of self that captures what we take to be a fundamental part of being a moral agent. The charge against act consequentialism here would be that it demands that we sacrifice too much of what we consider to be fundamental to our sense of self, which aligns roughly to our integrity. Our sense of self - and hence our integrity - is highly important to each of us, such that it would be unreasonable for any individual to be expected to abandon it. After all, our sense of self grounds us in the world, and helps to shape the way we act, as well as how we interact with others, therefore abandoning or significantly altering this sense of self is likely to have far reaching consequences for the individual, which are most likely to be to her detriment.

In this chapter I outline the integrity objection to consequentialism - using Bernard Williams' seminal 1973 essay, *A Critique of Utilitarianism* - and argue that this objection is the true underlying cause of our intuitions regarding demandingness. I will briefly examine some arguments which could potentially undermine this view and argue that, in the end, the integrity objection is a substantial objection to consequentialism and gives rise to the notion of over-demandingness, whether we consciously recognise this or not.

5.1. The Integrity Objection

In *A Critique of Utilitarianism* Bernard Williams argues that utilitarianism "makes integrity more or less unintelligible" owing to the fact that it "cannot coherently

describe the relations between a man's actions and his projects".[34] While the project of the utilitarian agent must be to maximise utility, there must also be other lower order projects which enable the pursuit of the utilitarian ideal in a non-vacuous way to begin with, and so not everyone can be acting simply in accordance with utility at all times. Some "lower order projects" will include things which the individual takes to be central to her life; commitments with which she identifies at the deepest level. These may be commitments to family or other close partial relationships, or they may be commitments to ideals or intellectual goals. Williams argues that it would be absurd to require an individual to abandon such a commitment purely because the results of the utilitarian calculus demand it, and act utilitarianism is thus seen to be "in the most literal sense, an attack on [one's] integrity".[35] Williams' argument, though referencing utilitarianism directly, clearly has the same result when applied to act consequentialism. Whatever the chosen good is that is meant to be maximised, the individual's major projects would still form only a part of the calculus based on their value in terms of this good.

One frequently levelled objection to consequentialism is that it cannot make a distinction between doing something and allowing something to happen. Because it is only the resultant state of affairs that matter (the amount of good in the world), it is morally equivalent to a consequentialist whether one commits murder or allows someone die through inaction, all things considered. In this regard, not donating life-saving funds to charity is no different than personally executing those that die as a result. Intuitively, this ought to be considered a significant difficulty for consequentialism by all but its most stalwart advocates.

Williams also argues that consequentialism does not recognise the difference between acting and allowing someone else to act. This fact is illustrated by his two

---

[34] Williams, 1973, pp.99-100
[35] *ibid*, p.115-7

thought experiments of Jim and the natives and George the scientist. In the former, Jim is undertaking research in South America when he emerges from the jungle to find a firing squad poised to shoot twenty native tribes-people for political protests, while their families watch. Upon seeing Jim and discovering his identity, the captain of the firing squad offers to let Jim shoot one native as a mark of honour and in return the rest of the natives will be freed. If Jim does not take up the offer, all twenty will be executed and nothing more will be said of the matter.

From the consequentialist perspective, all things being equal, Jim ought to kill a single person to save the lives of nineteen others. Here consequentialism fails to accept that there is a difference between Jim acting to bring about the worse consequences and acting in such a way that enables another to bring them about as a consequence. It is only making things go best which matters, and it is wrong to perform (or not prevent) any action that fails make things go best, whether the agent is directly responsible or not.[36] Williams refers to this peculiarity as "*negative responsibility*: that if I am ever responsible for anything, then I must be just as much responsible for things that I allow or fail to prevent, as I am for things that I myself, in the more everyday restricted sense, bring about",[37] including the actions of others.

Arguing that the negative consequences for the agent can be accounted for by the consequentialist calculus by including their strong misgivings as part of the calculus fails to solve the issue. As consequentialism requires that we treat the feelings of the agent, however strong they are, as if they were anyone else's, we ought not to give special weight to them to avoid acting: "'it's me' can never in itself be a morally comprehensible reason.[38]" However, requiring the agent to act in a way that contradicts their most heartfelt moral beliefs renders these beliefs nearly valueless

---

[36] *ibid*, p.93
[37] *ibid*, p.95
[38] *ibid*, p.96

and disconnects the agent from them in such a way as to alienate her from her integrity and from her sense of self:

> "Because our moral relation to the world is partly given by such feelings … to come to regard those feelings from a purely utilitarian point of view … as happenings outside one's moral self, is to lose a sense of one's moral identity; to lose, in the most literal way, one's integrity." (Williams, 1973, pp.103-4)

That we have moral limits or absolutes, or special attachments is essential to how we consider our relationship with others and the world around us to be confined and regulated. Should we not have a conception of what we would and would not do in given situations, it is likely that our conceptions of ourselves may well become inconsistent or at least blurred.

The negative responsibility generated by consequentialism also has grave consequences for our projects and goals. In allowing the projects of others to become just as much our responsibility as our own, the projects of others have a distinct and unreasonable effect on our decision-making:

> "While the deaths, and the killing, may be the outcome of Jim's refusal, it is misleading to think … of Jim having an effect on the world through the medium … of Pedro's acts; for this is to leave Pedro out of the picture in his essential role of one who has intentions and projects…. Instead of thinking in terms of supposed effects of Jim's projects on Pedro, it is more revealing to think in terms of the effects of Pedro's projects on Jim's decision. " (Williams, 1973, p.109)

Jim does not directly bring about the deaths of the natives and furthermore, Jim's

action or inaction does not force the captain to act in a specific way; as an agent he can choose his own actions. It therefore seems an odd feature of consequentialism that it holds Jim just as responsible as Pedro - in being necessarily directed by the projects of another agent, his actions are separated from his own projects by consequentialism. In forcing Jim to take responsibility for the projects of the captain, consequentialism has alienated him from his own projects and disregards the intuitive argument "that each of us is specially responsible for what he does, rather than for what other people do.[39]"

The consequentialist may highlight in response to the above, that the agent is not the only individual with personal projects, and that their projects must be balanced against those of other individuals. Most consequentialists take this impartiality to be a critical feature of the theory, in that it always ensures that the best possible outcome is brought about, rather than the one most preferable to the agent. However, this argument ignores the fact that the agent is not merely someone who has certain projects or goals which can be dismissed or changed to suit the consequentialist calculus; she *is* those projects and goals, they are inseparable from who she is at that moment in time, and to have them brushed aside in order to fulfil the projects of another is to devalue her as a moral entity in her own right:

> "The point is that he is identified with his actions as flowing from projects
> and attitudes which in some cases he takes seriously at the deepest level, as
> what his life is about. It is absurd to demand of such a man [...] that he
> should just step aside from his own project and decision and acknowledge
> the decision which utilitarian calculation requires. It is to alienate him in a
> real sense from his actions and the source of his action in his own
> convictions. [...] It is thus, in the most literal sense, an attack on his integrity"
> (Williams, 1973, p.116)

---

[39] *ibid*, p.99

Another aspect of the alienating nature of consequentialism is that it not only alienates individuals from their own projects, but it also alienates them from even their most important partial relationships. It does this in two distinct ways: The first and most obvious is that, in requiring an impartial view, there will necessarily be situations in which the consequentialist is required to act towards the benefit of a stranger while neglecting the needs of a loved one. For example, suppose a consequentialist is forced to choose between saving the life of his wife, or two strangers. Consequentialism, as a strictly impartial view, would require that he save the lives of the two strangers regardless of his feelings towards his wife - something that many would view as an appalling outcome.

However, this outcome is not unique to utilitarianism, and many other theories, including deontological theories, might require the same action. What is peculiar to consequentialism is the mental state of the agent in such a situation. Suppose that in the above example there is only a single stranger. Now the consequentialist calculates that, on balance and all other things being equal, it is permissible for him to save his wife and he acts upon this reasoning. In this situation he has acted as we might expect almost all people to act with regards to their loved ones. However, the difficulty for consequentialism lies not in the action but in the thought process that led him to act as he did:

> "This construction provides the agent with one thought too many: it might have been hoped by some (for instance, by his wife) that his motivating thought, fully spelled out, would be the thought that it was his wife, not that it was his wife and that in situations of this kind it is permissible to save one's wife." (Williams, 1981, p.19)

The problem here is that in being required to adopt an impartial perspective on the situation, he has had to treat his wife as if she were any other person in existence, favouring her merely because the consequentialist calculus permits it. In this way, he is alienated from the deeply held feelings he has towards his wife, and thus from her as well. This is an attack on the integrity of the agent because, much like our projects, our strong partial relationships are a fundamental part of us. In strongly limiting the value of these relationships in moral decision-making, consequentialism renders a large part of our lives far less meaningful than we would want to consider it, and in doing so makes itself a deeply unpalatable moral theory.

Before moving on, it is worth briefly investigating an alternative explanation of the intuitions surrounding the above cases. In response to William's example of Jim, Victor Tadros argues that integrity is unlikely to be significant in forming our intuitions surrounding the requirements of consequentialism. [40] The reason given is that violating one's integrity only occurs when the individual violates a duty to herself, not when she merely has a self-regarding reason.[41]   Not wanting to act against the conviction that he ought not to kill is certainly a *reason* for Jim not to kill one of the natives to save the rest, but it does not generate a sufficiently strong duty to outweigh his duty to save lives in this instance. Instead, Tadros argues, the intuition at play in cases such as this (when the agent is under duress) is that the act of killing one as a means to save nineteen is problematic precisely because of the restrictions we have on harming others as a means. Given a different situation, wherein Jim could save nineteen terminally ill people's lives by murdering another terminally ill person and redistributing her viable organs, while consequentialism both condones and encourages the act, it generates a strong intuition in most people that the act is wrong because we have a *pro tanto* duty not to harm others as a means to some further end. A similar intuition is brought out when examining another variation on the Jim

---

[40] Tadros, 2017, p.104

[41] *Ibid,* pp.99-100

case:

> "Suppose that there is only one Indian who is doomed to die. Jim can kill this Indian or let the captain do so. If Jim kills the Indian, he will receive a reward of some jellybeans that will otherwise go to waste. The Indian is no worse off than he would be were the captain to kill him, and Jim is better off in virtue of having received jellybeans. Yet it is surely wrong for Jim to kill." (Tadros, 2017, p.104)

That the native would have been killed anyway and Jim now receives a benefit justifies the act under consequentialism, but raises an intuition not regarding a duty to oneself, but about justifying the harm of others as a means to a positive end. In this variation the case is clearly wrong in violating this intuition, whereas in Williams' example the case is not so clear cut and most likely many would feel that Jim ought to kill the one to save the others, despite the conflict with the intuition at work. In this case, the outcome is the same as in Williams' argument, but importantly the motivation for hesitation is entirely different.

This argument is appealing - many of us hold the moral intuition that it is generally wrong to harm others to further some other end. However, it is clearly possible for us to appeal to multiple intuitions in a single case, and Tadros is wrong in arguing that the intuition regarding a respect for integrity does not apply. Tadros' error lies in his argument that in order for one to violate one's integrity, one must violate a duty to oneself. The notion of a moral duty is complex and unnecessary to explain the relationship between an individual and her integrity. If a person's integrity consists in a set of projects or goals which make a person's life worth living, then an attack on that integrity generates a reason for acting. That one has a duty to keep oneself alive seems a strange argument to make, but to argue that a threat to one's life generates strong rational reasons to act against that threat is perfectly natural. If a life has any

value at all, then that life has value specifically in the things that make that life worth living, such as the individual's integrity. Thus, if we can accept that one need not have a duty to act in a way that protects their life, and their life is valuable by the merit of their integral projects, etc., then it follows that they need not have a duty to protect those projects in order to justifiably act to do so.

A further problem for Tadros' is that in appealing to a specific duty regarding not causing harm to others a means to an end, his explanation has limited scope. Imagine Williams' George example is different, such that he is not required to work on biological weapons but instead is required to work for an oil company while he holds strong positive beliefs regarding environmentalism. Williams' appeal to integrity still holds while Tadros' appeal to not harming others does not (or at the very least does not have the same strength of appeal, due to the more nuanced subject matter). We could even present an alternative scenario wherein a farmer, Kate, has just finished harvesting her annual crop when an aid agency arrives and requests that she donate the entire harvest in order to save hundreds of lives in a neighbouring country which has been struck by famine. Kate's farm, which has been in the family for generations and is more important to her sense of self than anything else in the world, would be instantly bankrupted and she would have to give it up forever. Tadros' appeal to not harming others has no bearing on this case at all, whereas we can see how Williams' appeal to integrity can capture the intuition that consequentialism requiring her to give up her farm, while it may end up being the right thing to do objectively, is deeply problematic.

5.2. The Link Between Demandingness and Integrity

Some similarities with the demandingness objection should be apparent at this point. Firstly, both objections make the case that consequentialism requires that we regularly make a sacrifice that a moral theory ought to only require us to do in rare

circumstances. The demandingness objection takes issue with the fact that we are required to give up a lot of our resources and the integrity objection holds that we are required to give up projects and relationships crucial to our sense of self in order to promote the good. What they have in common is a seemingly overwhelming sacrifice on behalf of the individual for the promotion of the overall good. Another similarity is the lack of a right or power of the individual to be able to determine how they redistribute their resources or energies and how much of it they redistribute. In the demandingness objection there is no scope for holding back a little bit of wealth for oneself for luxuries, and in the integrity objection there is no possibility of pursuing even a relatively minor project for oneself if it hinders your ability to maximise the good, and in both cases the form or direction of the redistribution of either the money or efforts cannot be chosen based on the agent's preferences.

There is also a generally perceived notion that the two are linked causally as well. Specifically, the causal link is usually considered to be that the integrity objection is secondary to, derivative of, or simply a rephrasing of the demandingness problem.[42] The argument here is somewhat straightforward: if we are required to give up most of our resources and produce yet more resources to pass on for the maximisation of the good, then it follows that this will have an adverse effect on our ability to pursue our own projects and have mutually satisfactory partial relationships. There are also reasons to believe that the demandingness objection is pragmatically prior to the integrity objection. Before personally-fulfilling projects can be undertaken, it is necessary to have the resources required to ensure survival and thus, pragmatically, resources are more important than personal projects. Therefore, the demandingness objection comes in at every level, whereas the integrity objection is only prevalent amongst those for whom survival is not a pressing concern. Resources are essential to the existence and perpetuation of human life in a physical and immediate way,

---

[42] Chappell (2007) for instance, refers to the integrity objection as Williams' "version of the demandingness objection".

whereas personal projects and relationships are important in a less immediate and primarily psychological sense. With the health of the mind strongly contingent on the health of the body, so it appears that the integrity problem is contingent on the demandingness problem to a certain extent.

Furthermore, it is clear that many would, at a strong push, sacrifice their integrity for resources; the individual who would ordinarily refuse to utter even a white lie, might well be willing to tell a large lie if she were starving to death and the lie would result in a desperately needed meal. Many people also recognise that there are many things they would not ordinarily do owing to principles of one sort or another, that they would do given the right financial persuasion. It appears then, that the urge to gather and protect resources can trump the need to maintain one's integrity, given certain situations.

A final argument as to why the demandingness objection brings about the integrity objection, is that consequentialism more explicitly asks us to sacrifice resources for the maximisation of the good than it asks us to abandon personal projects. In the non-ideal world we live in, resources are more immediately crucial to raising the overall good in the world, owing to the prevalence of famine, disease and premature deaths in the world. Those currently dying of hunger and disease in Yemen, for example, are not in need of having personal projects fulfilled and partial relationships cultivated, their need is for resources (money, food, medicine, etc.). Their plight is not perpetuated or made worse by the rest of the world maintaining a network of close friends or striving to the be the best player in their local badminton club, but by the hoarding of resources and the reluctance to redistribute to those who could do more with it that we can. That having less resources is what may hinder our projects or relationships is clear, but it is only the importance of resources that brings out the integrity objection at all on this view.

5.3. The Demandingness Objection as a Derivation of the Integrity Objection

Some believe that the demandingness and integrity objections are not related at all, or at least, one does not beget the other. One objection to the idea of a link between the two is that the integrity objection takes aim only at the way in which consequentialism damages the idea of the self, and has nothing to say about demandingness or even the distribution of resources at all.[43] Here the only factor in the integrity objection is this effect on the self, and all other personally negative effects such as a large loss of resources are purely incidental. One's integrity, is and ought to be considered separate from one's material wealth it may be argued and, therefore, the integrity objection ought to be separated from the demandingness objection. To equate the psychological importance of maintaining integrity to the material importance of resources is a comparison that we cannot or ought not make. It is a mistake to equate them in terms of overall importance to the individual, or to connect them together as anything more than incidentally linked by the way in which resources are often necessary for maintaining integrity.

A second objection is that it is not clear that the integrity objection necessarily leads to higher levels of demand, as there are clear cases wherein integrity can be violated without excessive demands being placed on the agent (such as Williams' Jim and the natives scenario). If the link is not always present, then the two objections must be distinct and unrelated - "any theory, *whether demanding or not*, violates agential integrity if it does not respect the way agents relate to their projects and commitments.[44]" Thus, it is argued that this lack of a necessary connection between the two also signifies their separateness and distinctness.

However, rather than do away with the connection between the two, the notion that

---

[43] Tanyi, 2015, p.505
[44] *ibid,* emphasis added

they are not related only serves to highlight the inverse relationship between the ease of acting in accordance with each moral requirement; in times of plenty it is easy to maintain one's integrity, but in times of need what seemed integral to one's notion of the self seems less important compared with the more pressing issue of survival. It is, nevertheless, a clear connection. Integrity, and one's ability to preserve it, relies on resources as much as any of our other actions do. Clearly then, integrity is tied up with resources, and hence demandingness - at least in so far as limiting one's resources - limits their potential to maintain their integrity.

However, the argument that the two objections aim at different things overall might still stand. We might ask, what does this matter? Even if a person's integrity is only incidentally affected by the kind of demands that the demandingness problem takes aim at, it is nonetheless affected. Just because these sorts of demands do not seem to aim at the integrity of the individual by preventing their possibilities to pursue personal projects or relationships, it does not mean that the limiting effect it has on integrity is any less relevant. What may seem incidental, is only so if we imagine the two as distinct problems in their own right.

It is clear then that there is a meaningful connection between these two objections to consequentialism, though I believe that thus far most authors have been looking at this connection from the wrong direction. Rather than the demandingness objection informing or even causing the integrity objection, I believe that the reverse is true.

Instead, we ought to look why resources are valuable to us at all - surely, the fact that we find a sacrifice of resources demanding is because they have some use to us. In the first and most basic instance resources are useful as they are necessary for survival, though past this immediate need it becomes clear that resources are necessary only to meet some further goal. Money, for example, has absolutely no

value aside from its exchange value and therefore simply having money is of no use to anyone that does not intend to use it to some further end. This end could be purchasing a new house, investing in a private business, or buying artwork. Because all of these are then used to fulfil some other goal, when we are being asked to sacrifice or redistribute our resources what is problematic is not the reduction in resources, but the reduction in the possibility for goal fulfilment. Only someone with the obtuse and megalomaniacal desire to simply have as much money as possible would lament the loss of money qua money. This is not to say that all resources are viewed in entirely this way. A farmer may refuse to give up her land for even ludicrous sums of money because of a familial tie to the land stretching back generations. Here the resource is important in its own right, not simply for its ability to fulfil a goal. But this example illustrates my point rather than counting against it. The refusal to do away with the farm for money is simply because its value is not in what it can be used for, but in that it is integral to the sense of self of the farmer; without the family farm, she may feel separated from her past, and hence a part of her identity. All resources then, are either useful only for some further goal, or are deeply connected to a person's integrity owing to some peculiar trait of the resource itself.

At this point it may be asked, if resources are crucial for human survival, then surely consequentialism's demand on resources is more important and prior to any demands regarding integrity? In a situation where one ought to sacrifice resources at the cost of survival, this would be a more immediate and problematic demand than any placed on our integrity. However, the obvious response here would be to ask what is the point in survival at all? Surely not just the mere desire to be alive for the sake of it, as there are many horrifically painful situations we might be able to consider to which death might be preferable. Perhaps then, we might add the caveat that it is enough to simply be alive and pain free. But then what value does simply being alive and pain free offer us, if no further goals exist or can possibly be met? A pain free life, locked in a sensory deprivation chamber seems no life at all. So then

again, let us add the caveat that this pain free life must be one of experiences, but then we must ask again - which experiences? Herein lies the heart of the matter: a life worth living is one in which people have experiences guided by the pursuit and achievement of personal projects beyond mere survival. I am not claiming here to have touched upon some abstract meaning of life, but merely what many would find to be an intuitive baseline for a life worth living.

However, an objection might be raised that there are many millions around the world for whom survival is a daily issue, for whom the possibility of having attainable goals beyond mere survival is next to nothing, and for whom the integrity objection will mean very little. For these people, their primary concern is resource gathering for themselves and their loved ones and any hopes of having personal projects beyond mere survival will, for now at least, be abandoned. However, human beings' thoughts are concerned primarily with survival only when survival is an issue for them, otherwise our thoughts and concerns are directed primarily at our projects and personal relationships. A full human life requires survival first, but only as a means to the pursuit of projects which give life meaning. Thus, it seems strange to say that the integrity objection is derivative of the demandingness objection, and not the other way around. When survival is likely, alienation or conflict with integrity seem a far more pressing concern.

If the demandingness objection is a derivative of, or a secondary objection to the integrity objection then it follows that *the baseline of demand for a moral theory is linked to the negative effect that the redistribution of resources has on the integrity of the agent.* In other words, the baseline of demand, and therefore the maximum amount that any theory can demand of an agent, is dependent on a further baseline for the maximum negative effect a theory can have on the integrity of that same agent. We have therefore, answered Murphy's concerns regarding the need for a baseline to underpin a demandingness objection to consequentialism.

As we have seen above, the demandingness objection is concerned primarily with resource distribution and we have also seen that resources, in so much as they have a value, are valuable for two things: survival and the fulfilment of goals and personal projects. I should not, therefore, care much if consequentialism demands that I give £1000 to a certain cause, unless in parting with this money it impacts either my chances of survival or my ability to fulfil goals. If I were to object to this sum of money being redistributed from me without being able to point to any effect it may have on either my survival or goal fulfilment, then it would be clear that I was acting in a miserly or even irrational fashion. If one imagines the multi-billionaires of the world refusing to redistribute £1000 of their fortune to save a hundred lives, under almost all moral theories this appears to be an atrociously selfish and inhumane act.

However, should the same be asked of an average, reasonably well-off person, they may be able to point to any number of meaningful goals and personal projects that would be seriously hindered by the loss of this money. The fact that intuitions may vary regarding whether it is moral to retain the money for the projects in question is not crucial at this stage - what is crucial however is that intuitively a demand on resources can only reasonably be objected to when the fulfilment of this demand prevents some goal or aim of the individual further than pure selfishness. The fact that some goals or aims may seem to justify the denial more than others - for example buying an expensive hat seems like less of a justification for denial than replacing a broken car - only serves to strengthen the link between the level of demand felt and the level to which our personal projects, goals, and relationships are affected. If this intuitive connection is correct, finding a baseline for the effect a moral theory can have on our personal projects and relationships should shed some light on where the baseline of demand lies.

Before I attempt to demonstrate where exactly this baseline lies however, some brief

explanatory points are necessary. Thus far, I have talked somewhat loosely about the concept of integrity. I have supposed that integrity is tied up with our sense of self, and that it is comprised of our personal projects, in particular those which make life worth living. While I will elaborate further on this definition in chapter 8, I should like to make some further remarks here. It seems at first glance as though this conception of integrity is somewhat different both to how Williams conceives of it in the cases of George and Jim, but of also of how the term is used colloquially. When someone talks of their integrity in an everyday sense, they often intend it to mean something equivalent to following a moral code; such that when someone violates their own integrity, they are acting contrary to whatever moral code they prescribe to. This definition seems to fit very well in the case of Jim: Jim's moral code does not permit him to kill, and thus by following consequentialism, Jim is forced to act contrary to his own convictions in order to make things go best.

However, clearly there is an issue with using this definition of integrity. Under this definition, the fact that a person's integrity is violated when she is forced to act under consequentialism is trivial because any moral theory that did not align perfectly with her own moral code would also bring about situations whereby her integrity is violated (excepting moral relativism). That consequentialism prescribes that Jim kill when his own moral code forbids it is no different to deontology requiring that he not lie, when his own moral code requires him to occasionally lie to spare other's feelings. Furthermore, if the following of an individual's moral code were integral to her sense of self, then either she would adopt consequentialism as her moral code and there would no longer be a conflict of interest, or she would not, and she would not be compelled to act according to its rules in the first place. In this way then, if a person's integrity were to do with following a personal moral code, then either her integrity would always be violated by any moral code that conflicted with her personal morality, or it would never violate it because she simply would not adopt it.

So, it seems that this could not possibly be why Jim's integrity is violated by adopting consequentialism in this instance (at least if we want to talk about the concept meaningfully). Instead, I believe that the issue here is that Jim's integrity is violated because the act that consequentialism requires him to perform conflicts with Jim's conception of himself. Jim, as someone who has a desire to act morally, views himself as a person who does not kill innocent people, and as a result he commits himself to a project to not kill innocent people. When he is morally required to kill innocents, his sense of self is threatened, as is the project he has formed around this. This is not to say that he is acting in conflict with his moral code - as we can assume that if he is acting in accordance with consequentialism that his moral code *is* consequentialist in nature - but instead that he is acting in conflict with a project that informs his sense of himself, a project that he has taken on in order to give his life meaning. Integrity, then, is not to do with following one's moral code (otherwise those that adopted consequentialism would have no issue) but is about being able to continue to commit oneself to the projects that make one's life meaningful or worth living, including have a consistent sense of self. Consequentialism violates this kind of integrity, while other moral views do not, because other moral views can give integrity special importance.

6. Agent-Centred Prerogatives

In this chapter I will examine a prominent argument from Samuel Scheffler that makes the case for a hybrid theory which avoids the integrity objection in a way that classic consequentialist arguments do not. I shall begin by outlining Scheffler's argument as it is stated in *The Rejection of Consequentialism* and proceed to critique both his argument for supporting a hybrid theory, as well as his various arguments pertaining to consequentialist responses to the problem of integrity, wherein he intimates that consequentialism cannot properly accommodate the notion of integrity.

Scheffler's hybrid theory is comprised of two distinct parts: firstly, an agent-centred prerogative (henceforth referred to as ACP), which is incompatible with consequentialist principles and integral to any hybrid theory, and secondly a distribution-sensitive ranking principle which is not necessarily integral to hybrid theories in general, and which could be incorporated into consequentialist theories generally. As the latter is not relevant to the current discussion, my critique will focus on Scheffler's belief that there is a plausible rationale for including an ACP, such that hybrid theories are therefore more desirable than a standard consequentialist model.

Scheffler's aim is not necessarily to argue against consequentialism, but instead to perform a comparative examination of two different types of non-consequentialist moral conceptions.[45] The two kinds of non-consequentialist arguments he considers in the book are the standard deontological accounts (which Scheffler refers to as 'fully agent-centred' accounts) and 'hybrid theories'.[46] Both accounts shift the focus towards the agent and therefore "agent-centred" accounts, signalling a distinct difference from the standard act consequentialist model. I shall not be investigating

---

[45] Scheffler, 1982, p.4
[46] *Ibid,* p.4-5

what Scheffler has to say on the 'fully agent-centered' accounts in this essay for two reasons: firstly, Scheffler himself finds them to be an unsatisfactory alternative to his own hybrid theory, and secondly, as I want to be able to salvage as much of the central tenants of consequentialism as possible in response to the integrity objection, it makes sense to focus on a theory which aims in roughly the same direction, with Scheffler believing his agent-centred prerogative capable of doing so.

The main intuitive appeal of his hybrid theory is that it denies that agents *must* always do whatever would produce the best outcome, and therefore fits with our intuitions regarding our integrity remaining important in moral decision-making. However, it also allows that one *may* always do what would makes things go best, a plausible sounding argument that deontological theories reject by the nature of their agent-centred restrictions. Thus, if Scheffler's project succeeds, we have a viable alternative to act consequentialism which maintains one of its most plausible sounding and intuitive ideas, while circumventing the difficulties posed by the integrity objection.

6.1. The Agent-Centred Prerogative

In the early pages of outlining his hybrid theory, Scheffler claims that a hybrid theory differs from consequentialism only in that it includes an ACP would be capable of dealing with the integrity objection.[47] By including an ACP, the hybrid theory allows individuals to devote time and energy to their own personal projects and partial relationships out of strict proportion with their weight in a ranking from an impersonal standpoint.[48] This contravenes the consequentialist condition that all preferences, projects and partial relationships must be weighted purely from the impersonal perspective in order to achieve the best possible outcome. Thus, the adherent to the hybrid theory is not *always* required to do what would make things

---

[47] *Ibid*, p.14-5
[48] *ibid,* p.17

go best and may act in a way that does not make things go best from the impersonal perspective in order or to pursue projects and relationships of their own choosing.

Scheffler places two early restrictions on the ACP. Firstly, it should allow an agent to devote out of proportion time and energy to her projects "in such a way as to permit an agent the coherent integration of the agent's values and actions within the structure of a unified personality".[49] By this Scheffler intends to avoid legitimising any ACP models which require that the agent take conflicting attitudes towards her integrity at various different times, whether this be dependent on other moral factors arising or not. In other words, if an agent is to have any attitude towards her personal projects and partial relationships, then this attitude must be consistent. This requirement for consistency seems not only appropriate for dealing with the integrity problem, but also for maintaining integrity at all. Any theory which required one to view their integrity with inconsistent weight would, by its very nature, undermine integrity to begin with. The second requirement for an ACP is that it "place appropriate restrictions on the values and actions whose coherent integration and development it will protect"[50]; in other words, there must be some limit to the weight that one can place on certain individual values, in order to avoid the ACP simply allowing agents to do whatever they preferred. Again, the reasoning behind this is clear enough - a moral theory which allows agents to value their own preferences above all else and act on this egocentric valuation would not be much of a moral theory at all.

Considering these restrictions, Scheffler endorses what he refers to as a "plausible agent-centred prerogative", wherein whether an agent is required to promote the best outcome would depend on the amount of good she can produce, and the sacrifice she would have to incur, after she has assigned a proportionally greater

---

[49] *ibid*, p.19
[50] *ibid*

weight to her own interests, thereby allowing her to "promote the non-optimal outcome" of her choice, provided the gain in utility would still not exceed these revised agent-centred weightings.[51] This ACP satisfies the two desiderata outlined above. In the first case, it has the values of the agent "woven throughout the fabric of one's life", and thus allows for the coherent integration of their values by not consigning them to a "special sphere" of moral valuation.[52] Secondly, in requiring that the agent give their own proportional weight to their personal values, it is meant to ensure that there is no possibility of the agent being permitted to pursue their own projects at all costs and would "enable a normative view to accommodate personal integrity without collapsing into egoism".[53] So on this view, the fine line between respecting personal integrity and allowing for egoism to take hold has been trodden, and the integrity objection is accommodated well by this ACP, while still allowing for agents to always be permitted to bring about the best state of affairs (unlike in deontological theories).

Having set out the ACP as a necessary part of a hybrid theory which can deal with the integrity problem, Scheffler also intends to show that the integrity objection to consequentialism is not the only reason which provides an appropriate rationale for the ACP, but that the underlying concerns which give rise to the integrity objection can motivate the use of an ACP without the need for the objection to deal a decisive blow to consequentialism, or for the consequentialist to recognise this as an important argument.[54] These underlying concerns are brought about by the problem of reconciling an impersonal point of view (such as that of consequentialism), and the personal point of view of the agent. The impersonal perspective, which casts its eye over the whole of human affairs to decide which acts are right and wrong based on their bringing about certain consequences, ignores the fact that the agent cares

---

[51] *ibid*, p.20
[52] *ibid*, p.21
[53] *ibid*
[54] *ibid*, p.54

differently about her own projects just because they are her own. While consequentialism and other impersonal view theories can attempt to fold in the preferences of the individual to have her own projects fulfilled, it does not do so by seriously considering the personal, or independent view in its own right. The motivation for an ACP then, which is not dependent on the integrity objection actually taking hold of consequentialism, is that "to have an independent point of view is part of the nature of a person if anything is" and that moral theories, whether they violate the integrity objection or not, ought to reflect this fact by allowing the individual to give weight to their own projects and goals out of proportion with the status or value of these goals from an impersonal point of view.[55] In other words, our underlying concerns which bring about the integrity problem are that it is natural for human beings to have an independent point of view, and we ought not to try to have this independent view subsumed by an impersonal theory, but respect it and mirror it in whatever moral theory we adopt by allowing agents to give disproportionate weight to their own projects and goals. This underlying desire to keep the independent point of view independent of an impersonal ranking system is what motivates the need for an ACP.

Accounts which utilise ACP employ what is referred to here as the "liberation strategy", wherein the use of the ACP allows agents to devote resources to their projects out of proportion with their impersonal point of view by granting the personal point of view "moral independence".[56] The personal point of view is important not just because it allows us to discover and determine the personal projects that make a life seem worth living, but "because of what it tells us about the character of personal agency and motivation": people do not view the world from the impersonal perspective, nor do their concerns derive from it.[57] On Scheffler's view,

---

[55] *ibid*, p.58
[56] *ibid*, p.62
[57] *ibid*

a moral theory gives sufficient weight to a fact only if it reflects it by ignoring the demands that an impersonal view may generate. In other words, the world is seen most readily through the personal view, and hence our concerns and preferences are regularly generated from this perspective. It is only in special moral circumstances where we tend to view the world from the impersonal standpoint, and hence personal agency tends to derive from a personal point of view. Paying the appropriate amount of respect to personal agency therefore requires that we first give the appropriate weight to the personal point of view. The ACP then, is "a structural feature whose incorporation into a moral conception embodies a rational strategy for taking account of personal independence, given one construal of the importance of that aspect of persons".[58]

In practice then, an agent acting under Scheffler's ACP might have a set of relationships that they wish to hold onto, some more important than others, and also a set of personal projects which she wishes to pursue. For example, let us suppose that our agent is married with children, has multiple friends, values her career as an architect for her family's small construction business and wishes to one day design and build an impressive house for her and her family. Under standard act consequentialism, it is highly unlikely that she would be permitted to pursue this last goal actively and, as we saw in the opening chapters, it is unlikely she would be able to maintain the sort of meaningful relationships we would expect one to have with their partner and children. Furthermore, it is possible that act consequentialism might require her to choose a different career or work for a different employer that pays more money in order for her to be able to donate more to altruistic causes later. Under the ACP however, in order to be able to maintain the independence of her personal point of view by adding a greater weighting to her projects and relationships, we might say that her relationship with her immediate family members is weighted ten times higher than it otherwise would have been, with her friends, five times, and

---

[58] *ibid*, p.67

with work colleagues twice as much as usual. Bringing about her personal project of staying in the family business might be weighted ten times more than it otherwise would have been, and to achieve her lifelong dream of designing her own house, twenty times more. These numbers would not be arbitrary for the agent presumably; they would be picked based on the priority of each of the projects, and to how they affect the scope of her integral life. Once she has attributed these modifiers, it is a then simply a case of carrying out the usual act consequentialist calculations, but ensuring that the correct weighting has been applied. In some cases she will find that her preferences take precedent, and in others she will find that they do not. In all cases however, she is at least permitted to make things go best from the impersonal perspective, whether she is required to by the calculations or not.

6.2. Some Objections to Scheffler's Agent Centred Prerogative

Is Scheffler's shift towards the personal view using ACP necessary though? Some move to incorporate or take account of individuals' integrity is clearly required, but must this shift allow each to legislate for themselves under the personal view, as Scheffler suggests? While he makes some caveats to avoid egoism and maintain a roughly consequentialist framework for his hybrid theory, it does not seem to me as though his argument remains within the broad spirit of consequentialism. A criticism that the consequentialist could level at Scheffler here is the relegation of the impersonal view and the appeal to the importance of the personal standpoint. For many advocates of consequentialism, one of its greatest strengths is its legislating from the impersonal view, which prevents instances of individuals from avoiding acting from a duty to make things go best by arguing that it would not go best *for them*. However, as we have previously noted, this legislating from the impersonal view in classical consequentialism does have the unfortunate side effect of leaving the consequentialist position open to the integrity objection.

In allowing for agents to give extra weight to their own personal projects, Scheffler is aiming to maintain an impersonal view which is weighed against the personal view of the deciding agent. The initial response that this could lead to egoism is countered by arguing that the weight given to personal concerns must be proportional to the impersonal concerns to some extent. In other words, there must be some consistent limit to the weight which can be placed on personal projects. There are two major issues here: Firstly, Scheffler never describes how this might be done. This leaves open the option that someone might assign a very high value to a concern which they believe to be crucial to their integrity and that without any further caveats, this could be any concern. Suppose then that someone considered being immensely wealthy to be absolutely critical to their integrity. How would the hybrid ACP prevent the individual from attaching such a weight to this concern that they were not required to redistribute their wealth at all? Without a priority to the impersonal view, or to strict deontological restriction, it cannot. Thus, the protection that Scheffler intends against egoism is distinctly limited under ACP:

> "if one justifies an agent-centered prerogative independently of teleological considerations, but nests it within a more general consequentialism, justified, presumably, with a teleological rationale, where does one look for a justification for fixing the limits of prerogative? What is needed is a justificatory standpoint from which one can balance the competing rationales, but it is far from clear what that would be. In particular, it seems clear that appealing to teleological considerations to balance threatens to dissolve any nonteleological justification for an agent-centered prerogative." (Darwall, 1984, p.224)

Murphy also notes that the only basis we have for deciding on the ratio under normal circumstances is our current set of beliefs about what is an acceptable level of

demand, which are not only vague, but as far as beneficence theories are concerned, very much incorrect. Any beliefs that we can form about acceptable levels of demand would come from our current moral convictions, and our current convictions are evidently not sufficient for setting this limit, or else this discussion would have been over long ago. Our intuitions on the matter seem at best vague, and more often than not fluctuate depending on personal circumstances. It seems we are therefore incapable of making an independent judgement about the limits of demandingness, as we have nothing to fall back on other than our own intuitions formed by our current belief set. As Murphy notes, our inability to point to an acceptable level of demand, clearly brings ideas about unacceptable levels into doubt; "if extreme demands are not acceptable, we ought to have some sense of what levels of demand *would* be acceptable".[59] So if it is decided by the agent acting under an ACP that they need to give a certain extra weighting to a personal project or relationship, then the agent must already have in her mind an idea about what is an acceptable level of demand in normal circumstances. But as we have already seen, our intuitions regarding acceptable levels of demand are clearly inaccurate, owing to their inability to allow us to concede that there is more that we ought to be doing to make things go best. This being the case, it seems that without further teleological justification the agent's chosen weighting is likely not to be weighted in the optimum way. Scheffler's ACP has no response to this criticism.

A further difficulty, as Shelly Kagan highlights, is that there is no room within Scheffler's ACP model for differentiating between allowing harm to others to occur by increasing the importance of our own wellbeing and actually causing harm to others in a way that is acceptable under ACP.[60] For example, Scheffler would presumably want to say that, given a plausible proportionately greater weight to our own happiness and wellbeing, it is acceptable for an individual to not give £10 to

---

[59] Murphy, 2003, p.69
[60] Kagan, 1984, pp.250-1

charity but instead spend it on a drinks with a loved one. What it seems neither Scheffler nor the consequentialist would want to allow for is that ACP permit an individual to steal £10 from charity and spend it on drinks with a loved one. Supposing the charity were so large, that no one from the charity would notice that the money had been stolen, and that having a drink with a loved one was fairly important to us, whereas charitable donations were not personally important to us. In this situation, because ACP does not differentiate between doing and allowing to happen, it seems that ACP not only allows for, but permits acts that most would argue are completely morally unacceptable. Standard consequentialism would not allow for this sort of situation to arise, nor would most variations, and agent-centred restrictions would almost certainly prevent this from occurring. So the ACP model actually produces an entirely new issue by which the two approaches it tries to cut between avoid.

A second problem that Kagan raises is that there does not seem to be a good enough motivation for turning to ACPs in the first place.[61] Scheffler claims that the fact that people do not naturally think in accordance with an independent point of view goes at least some way to motivating the idea that they do not have to. This seems a rather unpalatable motivation for the inclusion of an ACP. Supposing the sudden discovery of a new tribe somewhere in the world, whose society was so patriarchal that women were passed around and bought and sold as objects, and even the women of the tribe found this so normal that they assumed it to be the only way of thinking about women. Moral questions about whether or not to interfere with the customs of isolated tribes aside, it seems strange to say that because these people simply do not think in terms of gender equality, or from the women's perspective, that this gives a reasonable motivation for including oppressive patriarchal views in their morality. It should be clear from this extreme example that simply not thinking in accordance with some moral perspective while naturally affirming the alternative, is not enough to motivate a rationale for accepting or including the alternative moral perspective.

---

[61] *ibid,* pp.252-4

Thus, simply not naturally thinking in accordance with the impersonal point of view and naturally affirming the alternative (the independent point of view) is not in itself enough to motivate including the independent view in our moral theory.

Another objection to the ACP is that hybrid theories that utilise only ACPs do not fully protect integrity in the first place, and that to do so agent-centred restrictions (ACR) would also be necessary.[62] ACPs do not fully protect integrity because they do not protect individuals against the desires and duties of others to maximise the overall good. Thus, ACPs (as Scheffler admits) allows for individuals to be forced or coerced into giving up central projects by others, so long as the other agents were acting in order to maximise the good by the standard of the ACP. Scheffler pre-empts this line of argument, but does not expect that it could be carried as far as this objection suggests, and supposes that this would occur rarely in a society governed by a hybrid theory.

Furthermore, ACPs give agents the opportunity to thwart others' projects in favour of their own, because they no longer have to justify their actions as increasing the overall good. The personal weighting allowed by the ACP allows for situations wherein one can rationalise committing acts that we would generally want to consider immoral. Consider the following case:

> "Suppose a stranger's car is worth twice as much to her as it would be to me if I stole it. In whatever units we measure the good, she gets two units if she uses the car for her own projects. I forgo one unit if I don't steal it and use it for mine. Not stealing is thus impersonally optimal, but compare the inferiority of stealing to the sacrifice involved in not stealing. Stealing is inferior by one net unit. My sacrifice if I don't steal is also one unit. Therefore, the degree of inferiority doesn't exceed my sacrifice at all, let

---

[62] Schmidtz, 1990, p.625

alone exceed it by a specified proportion. So even letting the proportion go to zero fails to help. The theory still lets me steal the car." (Schmidtz, 1990, p.625)

So ACPs, in utilising a personal weighting, leave open the option for agents to bring about intuitively wrong circumstances which seem like the sort of situations a moral theory is supposed to prevent against. For ACPs to be rational then, there needs to be something that makes up for the loss of third-party protection.[63] Adding purely consequentialist restrictions to an ACP cannot do this because they cannot distinguish between causing harms and failing to prevent them. This leaves only ACRs as an alternative. However, once ACRs are motivated by a hybrid theory, we have then moved exceptionally far away from consequentialism, as we now have a theory in which there are *at least some* situations in which one is not permitted to maximise the good. There is then very little that needs to be done to motivate a move to fully-agent-centred theories such as deontological theories, and the battle to maintain a consequentialist framework in light of the integrity problem has been lost.

6.3. A More Developed Agent-Centred Prerogative is Needed

What each of the above objections serve to demonstrate is that the ACP as Scheffler puts it forward is lacking something. Scheffler himself recognises the theory as incomplete and even goes as far as to suggest that an ACP that took a different, non-cost-based form, might be more effective.[64] The distinction between doing and allowing harm could potentially be accommodated by an ACP that prohibited doing harm in non-optimific decision-making for example. Regardless, what any ACP needs to capture is only that there is a personal point of view, which is central to the way in which agents interact with the world, and which needs to be accounted for by any

---

[63] *ibid*, p.626
[64] Scheffler, 1982, pp.167-191

plausible moral theory. That the ACP that Scheffler has outlined fails to do this without running into difficulties elsewhere is not a decisive blow against the use of an ACP in general to modify consequentialism in such a way as to make it robust against the integrity objection. What the above objections do suggest, however, are restrictions and suggestions for any ACP that we might try to utilise: it must not allow the weighting to become trivial, or to become too heavily biased in favour of the agent, but it must at the same time prevent the personal perspective from being obscured in any restrictions placed in the weighting process. Furthermore, it seems as though it must accommodate the distinction between doing harm and allowing harm to happen, in order to prevent situations whereby one's weighting can permit obviously immoral actions, but it must make this distinction without imposing an agent-centred restriction such that in some cases we are not permitted to perform the most optimific act. I shall attempt to outline a version of an ACP which manages to accommodate these two major concerns in chapter 8.

7. Maximising Integrity

In contrast to the above "liberation strategy" that Scheffler argues for, there is alternative solution to the integrity problem posed by some consequentialists which he refers to as the "maximisation strategy". Proponents of this kind of argument claim that a more complicated form of consequentialism, that takes integrity as one good among others to be maximised, can accommodate the integrity objection and may have similar results to the ACP hybrid.[65] The main intuition behind such a view is that the importance of the personal point of view "is an important fact for morality […] because it fundamentally affects the character of human fulfilment and hence the constitution of the overall good"[66] and therefore ought to be accommodated into the framework of the consequentialist calculus. However, this view, according to Scheffler, would still contain a strict proportionality requirement in order to satisfy the consequentialist appeal to the impersonal point of view.[67] This, he argues, would itself violate integrity, being vulnerable to "the charge that it systematically undermines the integrity of *all* agents", owing to the fact that agents may be required to sacrifice most of their integrity in order to increase others' integrity or wellbeing.[68] In other words, while maximisation can take account of the fact that people have concerns generated outside of the impersonal view, it does so by denying the moral independence of these concerns; "it denies that personal projects and commitments can have any moral weight for an agent […] independently of the weight those projects and commitments may have in the impersonal calculus".[69]

I shall now explore whether there are other ways in which one can accommodate

---

[65] Scheffler, 1982, p.42
[66] *ibid,* p.60
[67] *ibid,* p.42, p.61
[68] *ibid,* p.42
[69] *ibid,* p.61

integrity into a maximisation strategy, that avoid the issues that Scheffler highlights. Clearly the aim here is to ensure that the theory remains broadly act consequentialist in nature. This therefore narrows the scope of our maximisation theory to two options: either we adopt a pluralistic consequentialist theory in which the difference in deliberative strategies to be employed is purely axiological (and therefore all deliberation is a mode of act consequentialism), or we have a single value consequentialism wherein the axiology allows the theory to outmanoeuvre any integrity-based objections.

As the latter is clearly a simpler matter, I shall begin by positing some potential axiological positions which may allow integrity to be sufficiently incorporated into act consequentialism. Remembering the value focused definition of act consequentialism given earlier[70], it may be that simply marking out integrity as the value to maximised is sufficient for giving us our first integrity-considered maximisation strategy (ICMS for brevity from hereon):

> ICMS 1: an act is right if and only if it results in *at least as much overall integrity* as any action the agent could have performed

Here though, there is an immediate and obvious issue to contend with: the definition of integrity is broad enough that it is very unclear how one might go about enacting this strategy. Furthermore, while it is sometimes argued that it is difficult to know how one might maximise happiness when happiness is so agent relative, integrity is clearly even more so. It would be near impossible to consider how each and every concerned individual's integrity may be affected by a decision, as what is integral to an individual is, more often than not, a private matter. We can, however, assume that in most cases, a person's integrity will be heavily defined by their personal projects

---

[70] "an act is right if and only if it results in at least as much overall [good] as any action the agent could have performed" (Eggleston, p.125)

and relationships[71] and therefore a maximisation of personal projects being fulfilled or carried out will have the effect of maximising integrity:

> ICMS 2: an act is right if and only if it results in *at least as many personal projects being fulfilled overall* as any action the agent could have performed

We have then narrowed down what we are aiming to achieve as a consequentialist now; rather than the somewhat nebulous concept of maximising integrity *qua* integrity, this strategy recognises that, while there may be more to integrity than the fulfilment of personal projects and relationships, these are the most easily promoted components of individuals' conceptions of their integrity and therefore a pragmatic maximisation strategy ought to aim to maximise the fulfilment of personal projects. However, while it is now clear what it is we are aiming to maximise, it is still not clear how best to do this. People's personal projects are still going to be largely unknown to the agent and many personal projects may be closely guarded private affairs for individuals (imagine the teenager whose personal project is to reveal their hidden sexuality to their parents, or the son of a coal miner who wants to be a ballet dancer, for instance). This being the case a more pragmatic approach still may be to work not towards actually fulfilling other individual's projects on their behalf (in many cases this will be impossible to do anyway), but instead to work towards the promotion of an environment whereby there is as much possibility for individuals to fulfil their projects as possible:

> ICMS 3: an act is right if and only if it results in *at least as much overall possibility for personal projects being fulfilled* as any action the agent could have performed

---

[71] As relationships could be considered projects of a social nature I shall subsume these into projects for the purpose of further strategy definitions

This, I believe is almost as close as a single value maximisation strategy can come to being able to actually promote integrity as far as possible. It also, by the virtue of the requirement to meet certain basic needs before one can even begin to have personal projects, has the effect of taking a proactive approach to improving the welfare of individuals as well, such that it will have as positive effect on the poorest in the world as act utilitarianism would, and most likely an even greater positive impact by virtue of recognising values beyond base happiness and welfare. However, there is one last refinement that I wish to make to this maximising strategy in order to avoid a possible problem:

> ICMS 4: an act is right if and only if it results in *at least as much overall possibility for meaningful personal projects being fulfilled* as any action the agent could have performed

Meaningful is enlisted here as a necessary criteria for the promotion of a project in order to avoid the possibility of the strategy requiring that we devote ourselves to bringing about situations wherein we have the personal project equivalent of Nozick's utility monster[72] or situations whereby we have people with many personal projects, but all of which contribute very little to their sense of integrity (we might imagine here an individual who has many small projects such as buying a new oven, completing a model aeroplane, cleaning the garage, playing golf more regularly, and so on). It may be that if we are maximising the possibility to fulfil any projects at all, then there will be such a plenitude of nearly meaningless projects that it will be very easy to fulfil, that we may never be in situation whereby we can dedicate resources or time to promoting the projects that genuinely matter to people. It might also create a situation wherein we might be required to let those in extreme poverty die in order to fulfil the minor projects of the well-off, due to the cost required to bring someone up to the level where they can have personal projects that go beyond mere

---

[72] Nozick, 1974

survival and the relatively minimal cost of fulfilling a teenager's project to finish a video game, for instance. By requiring that we maximise the potential for meaningful projects to be fulfilled, we are thereby ensuring that we are creating a society and situation in which people can pursue the projects that genuinely make life worth living, rather than those that merely give a marginal increase in life satisfaction and therefore contribute very little to individual's sense of integrity.

However, while this strategy may serve to increase the overall amount of individuals whose lives are lived with as much integrity as possible, it is readily apparent that the agent is still separated from her own integrity. In still being required to act from an agent-neutral perspective, she is still required to sacrifice the fulfilment of her own meaningful personal projects in order to promote the possibility for meaningful personal projects to be fulfilled by others, should the calculus indicate that this is possible. Defenders of ICMS 4 might claim however, along similar lines to Sidgwick's defence of utilitarianism[73], that owing to the agent's privileged position with regards to knowing her own personal projects more intimately, that she is far more likely to bring about the possibility of fulfilling meaningful projects by dedicating at least some of her time to her own projects. This, I believe, is an empirical question which will most likely be borne out to be false, however, disregarding this assumption, it is still the case that the agent is separated from her integrity by virtue of the fact that, under ICMS 4, her personal projects and relationships - and therefore her integrity - hold no privileged place and count for no more than those of any other individual. The importance of her personal projects is once again dissolved into nothing more than their weight from an agent-neutral perspective, and the special link between the agent and her projects is undermined, thus undermining her integrity and her sense

---

[73] "For the practical application of this theoretical impartiality of Utilitarianism is limited by several important considerations. In the first place, generally speaking, each man is better able to provide for his own happiness than for that of other persons, from his more intimate knowledge of his own desires and needs, and his greater opportunities of gratifying them." (Sidgwick, 1907, p.431)

of self.

I stated above that there was also the possibility of pluralistic consequentialist strategies that take integrity into account. A basic example might be the following:

> an act is right if and only if it results in *at least as much overall possibility for meaningful personal projects being fulfilled or as much utility* as any action the agent could have performed

However, this does little to resolve the problem of the agent being separated from their integrity. If a pluralistic maximisation strategy is to take account of this problem, then it must include an agent-centred maximisation as part of its criteria of the right. This leads to:

> ICMS 5: an act is right if and only if it results in *at least as much possibility for meaningful personal projects being fulfilled, either for the agent or overall,* as any action the agent could have performed

This formulation is clearly flawed however, owing to the fact that the agent can choose to act in a way that maximises the possibility of her own personal projects being fulfilled over those of others, without any further caveats of when this is acceptable or not. Under this strategy then, it would be entirely possible for the agent to lead an fully egocentric life, only ever tending to her own projects and relationships (so long as they were meaningful) and ignoring those of others. This highlights what I believe to be a general problem for pluralistic maximisation strategies, which is that they fail to demonstrate when a certain value is to be prioritised over another. If, for example, we were to amend the above strategy to include a clause which required that the agent forgo maximising her own projects when there is the possibility to help those in an emergency situation, then we run into familiar problems of what

constitutes an emergency[74]. (Does proximity count? How in need does the person have to be? Etc.)

Furthermore, it is not clear that including a provision allowing the agent to maximise either personally or impersonally would actually resolve the problem of separating her from her integrity. Because both values are constantly in play, it is not clear that there is a case wherein the agent's integrity *ought* to take priority. It does not seem enough merely to state that there is always the possibility that it can, as this means that her own projects are only ever as important to her as maximising the projects of others. If we are to capture the notion that there is something special about the character of one's own projects, then our strategy must contain some notion that we ought to respect the agent's integrity, not that it is merely one choice among others. For this reason, it is impossible that a maximisation strategy, pluralistic or otherwise can capture the intuition that led to the integrity problem in the first place. To resolve this issue, we must return to the notion of agent-centred prerogatives.

---

[74]   See Peter Singer's drowning child thought experiment for a more in depth analysis of this problem. (Singer, 1972)

8. A Solution to the Integrity Problem; Reflective Agent-Centred Consequentialism

We have seen then, that maximisation strategies, no matter the axiological grounding, will always fail to account for the special relationship between the agent and her projects which characterises integrity and lends credence to the integrity objection. If we are to find a plausible solution to the integrity problem, it is necessary to return to agent-centred prerogatives. However, we have already seen from the discussion of Scheffler's ACP in chapter 6 that there are several problems that may plague an ACP. Firstly, it is necessary that the prerogative must have a fixed limit, or at least one that is decided independently of the agent's personal views about how much more weight her personal projects have than those of others. Secondly, there needs to be some strong justification for allowing our intuitions regarding integrity and the personal point of view to dictate the limits of morality. Thirdly, there needs to be some way in which the prerogative protects the projects of the agent from being overridden by other agents' needs, without utilising an agent-centred restriction. Finally, the view needs to be able to accommodate the notion of increased demand in times of great emergency. If these issues can be redressed, then it may well be possible to utilise an ACP to avoid both the integrity and demandingness objections.

8.1. A New Agent-Centred Prerogative

I suggest that a plausible agent-centred prerogative take the following form:

> **An agent is permitted to act in a way that would not make things go best overall, iff making things go best would violate her integrity**

The first difference between this ACP and Scheffler's is that there is no weighting

system employed within this new ACP (NACP for brevity from here on out). There are several strengths to not employing a weighting system like Scheffler's. Firstly, the problem of how much weight to give to our agent's integrity is no longer an issue. Scheffler's ACP had the twin issues of how much weight the agent should give to her own projects and relationships, and how this should be effected as a ratio of the weight of what one might consider general consequentialist demands, whereas the NACP has no such issue. In permitting the agent to flatly refuse to do what would make things go best if doing so would violate her integrity, then there is no need for her to have a weighted valuation of her integral projects; if her integrity is violated, then by how much or to what degree or in what circumstances is irrelevant under NACP - it is enough that the agent would be deprived of her ability to operate with integrity to warrant permitting her to act in a way that would not make things go best.

Part of the problem that Scheffler's ACP had with how to balance the agent's weighted concerns with global demands was due to the fact that his ACP and his maximisation principle had competing values that were not of the same type. For Scheffler, the balance was between the integral projects of the agent, and the overall welfare of rest of humanity. This creates the problem of how to balance two sets of values, where the quality of each was not comparable. It also allowed for the distinct possibility that, given the high demands of consequentialism, the agent would regularly find that she was required to act in a way that undermined her integrity. The NACP has no such problem, as the aim for both the maximisation principle and the prerogative are both to maximise integrity by respecting and promoting the central projects of individuals. In this way we do not have two competing axiological premises, but one single premise that maximising the possibility to fulfil individual's integral projects is what would make things go best overall. The rational personal expression of this axiology is that one ought to perform the actions which would maximise her possibility to fulfil her personal projects. This is expressed morally by ICMS4 and protected at the rational personal level by the NACP. ICMS4 and the NACP

are therefore two ways of expressing the same rational concern for integral projects.

Similarly to Scheffler's ACP, the NACP is strictly not an agent-centred restriction, as it does not require that the agent never make things go best when her integrity might be violated, instead it merely permits her to not make things go best. This means that, should the agent choose to, she would be perfectly able to act in a way that would undermine her own integrity by performing an action that would make things go best. For example, should an agent view maintaining an honest and open relationship with her child to be essential to her integrity, she might choose to act contrary to this by telling her child that she cannot afford to buy a certain desperately wanted toy, when in fact she could but intends to use the money for charitable purposes instead. Thus, the NACP also introduces the notion of supererogatory actions into the consequentialist framework, and admits of no situation where it would not be possible for the agent to make things go best.

8.2. The Set of Integral Projects and Relationships

It may now be asked how we can know when the agent is in a situation where her integrity may be violated by performing an action. In order to understand how the agent can justify a refusal to act in the way that make things go best overall, I shall introduce two further criteria to the NACP which define the limits of the agents protected integrity:

> **Criterion 1: An agent's integrity is deemed to have been violated iff she is unable to pursue some personal project or relationship which she considers to be integral to her sense of self.**

The above criterion provides a clear way of defining the agent's integrity under the NACP. The agent's integrity has been violated only when one or more of the personal

projects or relationships which she believes to be part of her integral sense of self have been undermined. Under the NACP then, the agent's integrity is comprised of a set of projects and relationships which are protected from being necessarily overridden by consequentialist demands. However, trivial personal projects and relationships are excluded from this set, owing to the criterion's demand that the agent must consider those included to be integral to her sense of self. Previously, I have discussed the idea of integral projects as being those that make life worth living; those that, when absent, cause life to feel devoid of meaning or purpose. The set of projects protected by the NACP must reflect this fact by being those which are integral to the agent's sense of self - those which characterise her life and make it worth living. This notion provides a clear limitation that the agent cannot treat all her personal projects and relationships in the same way, as might be the case in an ACP which employed a weighted comparison system to protect the agent's projects. In refusing to act in the way that would make things go best, the agent must be doing so in order protect one or more projects or relationships without which the character of her life would be severely altered and her sense of self damaged.

For example, under the NACP, if the agent were to place an extremely high value in honesty, to the point where it was essential to her sense of self that she always acted honestly, then it would be permissible for her to refuse an act which required her to lie in order to make things go best. Should she examine the situation and find that she nevertheless felt compelled out of respect for consequentialism to make things go best regardless, then this too would be acceptable, and in fact it would be morally commendable as a supererogatory act. Nevertheless, despite how praiseworthy violating her own strongly held beliefs might be, under NACP we cannot admonish her for not acting thusly, if she did so to protect her integrity.

Relating this back to Williams' example of Jim and the natives, if not killing innocents was a value that Jim held to be integral to his sense of self, then it seems that the

NACP would permit him to not shoot the native in order to protect his own integrity. However, herein lies a problem for the NACP, so far as it stands, as there is still no clear distinction being doing and allowing to happen. As a result, if not killing is integral to Jim's sense of self, then he is still faced with the dilemma caused by being in a situation whereby he is either negatively responsible for the deaths of 20 natives, or positively responsible for the death of one. As the NACP has so far made no distinction between doing and allowing to happen (positive and negative responsibility) Jim's integrity will be violated no matter what action he performs.

Another problem for the NACP is that, while we have determined that the set of protected projects must be comprised only of those that the agent takes to be integral to her sense of self, without further restrictions the variation in what is integral to the agent will be likely to produce some interesting and highly undesirable results. Take for example, the Wall Street businessman who's main drive in life is to become as powerful as possible. He may well believe that being powerful is the only thing that truly matters. Under the NACP this agent would be able to justify any number of abhorrent actions or not performing any actions that would make things go best, as these actions would be to the detriment of his integral project of gaining and consolidating power. Furthermore, his integral project would likely have the interesting side effect of putting him in a position of power whereby he might be able to do an exceedingly high amount of good, but refuses to do so in order to maintain and increase his power. That the NACP allows for instances such as this shows that there is need for a further criterion to determine the set of protected projects.

8.3. Utilising Reflective Equilibrium

In order to eliminate as much as is feasible the possibility of agents being able to lead a life of nearly unchecked personal project fulfilment to the detriment of the overall

good, the agent must be required to actively reflect on her intuitions regarding her integrity and revise these as necessary, within the framework of consequentialism by using the method of reflective equilibrium:

Criterion 2: The set of projects and relationships which the agent considers to be integral to her sense of self must be determined through a continual method of reflective equilibrium, from the perspective of consequentialism

Reflective equilibrium, in this context, is the term given to the method (and its resulting state) of working back and forth between our moral intuitions or considered judgements, and our governing moral principles, adjusting or revising elements of our thinking as we do so in order to arrive at a state of coherence between our moral intuitions and our governing moral principles.[75] In the context of the NACP, this means reflecting specifically on our intuitions surrounding demandingness and integrity, with the governing principle of ICMS4. The method of reflective equilibrium in this context therefore requires that the agent continually reflect on her beliefs regarding what constitutes her integrity - that is, which of her projects and relationships are integral to her sense of self and which are not - and the demands placed on her by ICMS4. In this way, the agent ought to be in a continual process of revising which beliefs she deems to be integral to her sense of self, based on which she feels she can justify protecting against the demands of consequentialism[76].

It is clear that this process will not be instantly successful; it is highly unlikely that the agent will be able to realise a state of true reflective equilibrium on her first attempt[77].

---

[75] Norman, 2020

[76] Rawls (1971) defines reflective equilibrium as the process of weighing various proposed conceptions and either revising judgements to accord with one of them or holding fast to the initial convictions (p.48). Within the context of the NACP, the conceptions to be weighed are the integrity of the agent and ICMS4, in an attempt to revise her judgements about what it is that constitutes leading an integral or worthwhile life.

[77] In fact Rawls notes that "it is doubtful whether one can ever reach this state" (1971,

The point however, is that whatever the starting point of the agent, whether they are naturally highly altruistic or very selfish, each agent is morally required to justify holding the set of integral projects that she does at all times. This means that while an agent may start out with a large set of protected integral projects, over time this set ought to be reduced upon reflection until she arrives at the set of beliefs that are truly integral to her sense of self (rather than merely appearing to be when unexamined or misconstrued as more important than they actually are, due to other factors either external or internal to the agent).

Utilising reflective equilibrium as a tool to revise the agent's protected set of projects has several advantages. Firstly, it allows the NACP to be a highly adaptive and revisionary prerogative, unlike Scheffler's ACP which is rigid and unresponsive to the state of the agent and the world. The agent may revise her beliefs based on new psychological circumstances (such as including being a trustworthy person to be an integral project after falling in love) or she may revise them based on the state of the world (perhaps by including a strong commitment to environmental issues after learning more about the impending climate catastrophe). She may also lose as well as gain integral projects by failing to be able to justify them in light of the demands of ICMS4; perhaps her commitment to being a thoroughly fashionable person no longer seems appropriate given that the fashion industry seems to engage in activities that make many people's live far worse off.

The most crucial benefit of engaging in this method of reflective equilibrium is the effect it has on the perceived demands of consequentialism. Suppose that ICMS4 alone demanded that the agent donate 50% of her wages to charity each month in order to make things go best by raising the world's poorest people up to a state whereby they can begin to pursue personal projects of their own. For most, this would be considered an intuitively extreme demand and might have devastating effects on

---

p.49) when discussing the theory in regards to conceptions of justice.

their ability to maintain a sense of integrity. Suppose that our agent has an incredibly active social life that she deems to be integral to her sense of self; she is in some respects defined by her character as a socialite and feels as though maintaining this lifestyle is, for now at least, essential to her leading a life worth living. In order to maintain this lifestyle she meets with groups of friends every night, at cafes, bars, the cinema, and so on. Maintaining this lifestyle, and hence her sense of integrity is therefore incompatible with the demand that ICMS4 places on her to donate 50% of her wages. After reflecting on which projects are really essential to her sense of self however, she decides that she can reduce seeing her friends from nightly to five times per week. This still feels like a high demand, but she feels as though it does not undermine her integrity and is now able to donate 15% of her wages to charity each week, by giving up this and some other, non-integral projects (whatever these may be). This feels like a high but not extreme demand initially and then after some time it hardly seems like a demand at all; it becomes just a part of her (meaningful) life. At this point, the agent may reflect again on her integral projects and realise that she may be able to reduce the amount of socialising she does further from five to three nights a week, as well as eliminating another project which used to seem integral, but no longer does owing to her newly defined sense of self[78].

In this way the agent can keep revising her integral projects and as time passes she may find that demands that once seemed extreme feel less so and she is able to come closer and closer to being able to act fully under ICMS4. The aim here, however, is not for the agent to abandon all integral projects over time, but to arrive at a point of reflective equilibrium - the point where her integral projects can be reduced no further and she has arrived at the set which is truly reflective of her sense of self and

---

[78] Miller (2004, p.362) describes this process as "an exasperating trick", similar to that of a child who keeps pushing the limits of her parents' rules. This, I believe, is a mistake brought about by relying too heavily on the intuition that a moral theory ought to be moderate (see Berkey, 2016 for a discussion of the dangers of allowing this intuition to creep into one's theory).

allows her to lead a predominantly act consequentialist life, whilst pursuing as many personal projects as is actually necessary to lead a life with integrity. Of course, it is unlikely that any agent will ever truly reach a permanent state of reflective equilibrium. The psychology and circumstances of the agent alter over time, as does the state of the world, meaning that it will be constantly necessary for the agent to be examining and revising her set of integral projects. It may be possible for an agent to reach permanent reflective equilibrium, but whether she does so or not is irrelevant to the project of the NACP. What is important is that, in engaging in the process of reflective equilibrium, she is able to adapt to levels of demand in a way that enables her to lead a broadly consequentialist life with integrity. Crucially, however, it is important to remember that at no point in the process of reflective consequentialism is the agent restricted from acting entirely in accordance with act consequentialism. As a result, even in the early stages of the process when she might have what seems to be a counter-intuitively vast list of integral projects, she is still able to react to sudden changes in the world which result in a sudden increase in demand, even if these are temporary measures to prevent disaster.

8.4. Reflective Agent-Centred Consequentialism

With the prerogative and extra criteria set out, combining these with ICMS4 we arrive at *Reflective Agent-Centred Consequentialism*:

> **The agent is only permitted to perform an action, φ, iff φ results in at least as much overall possibility for meaningful personal projects being fulfilled as any action the agent could have performed, unless performing φ would violate her integrity by forcing her to undermine one or more projects or relationships that she has determined to be integral to her sense of self through an ongoing process of reflective equilibrium.**

What remains to be seen is how this new consequentialist theory deals with some concrete examples of the integrity and demandingness problems. Let us return to Williams' example of Jim and the natives. It seems as though Jim now has three options: make things go best regardless of his integral projects by shooting a native, refuse to violate his integrity by allowing all twenty natives to die, or revise his integral projects given the information available to him and decide that he ought to exclude his rule against killing from his set of integral projects. By reflecting on his integral projects, he could, in theory, decide in either direction. That is, he could discover through reflective equilibrium that he no longer needs to hold on to his integral project of not killing, or he might decide that this project is as important as ever. Under RACC, it is quite literally his prerogative. It is highly likely that over time all agents will tend more and more towards having less integral projects and therefore acting more and more in line with act consequentialism, but it is not guaranteed that any particular project is removed from the agent's set of integral projects.

That is, however, excepting those which are internally incoherent. Remember that one of the core issues for Jim was that even if RACC permits him not to shoot the native and therefore to allow all twenty to die, owing to consequentialism's lack of ability to distinguish between doing and allowing he is still negatively causing a death (nineteen extra deaths in fact). However, under RACC Jim is required to enter into the process of reflective equilibrium which requires as part of its method that he adopt a coherent set of projects. As this needs to include act consequentialism as part of RACC, Jim's set is no longer coherent - he cannot have a project to never kill included in his set of integral projects because his overall moral principle, act consequentialism, does not distinguish between positive and negative responsibility, meaning that in some cases (such as the present one) his project is one that it is impossible to protect. So, in actual fact, if it were simply a prohibition from killing that Jim cites as a legitimate reason for not shooting the native, then under RACC he is not

acting morally at all.

This highlights what some may consider a flaw of RACC - that it is still quite demanding; certainly far more so than common-sense morality. Common-sense morality, for instance, might well argue that either outcome for Jim would have been acceptable, or at least neither action would cause him to be considered morally blameworthy given the situation. RACC rejects this by requiring that we hold an entirely consistent and coherent set of integral projects. If it is the case that an agent has an incoherent set of integral projects then this is simply a failure of their utilising reflective equilibrium, but if they refuse to act in whatever way would make things go best because their set of integral projects was incoherent, then they are acting wrongly and are morally blameworthy for their actions. RACC might also be seen as still being very demanding owing to the limitations of what qualifies as a supererogatory act. Common-sense morality might want to argue that any act of charity or self-sacrifice was supererogatory, but RACC acknowledges an act as supererogatory only when an act violates the agent's integrity; the common-sense moralist would most likely want this threshold to be considerably lower.

However, while it is almost certainly the case that RACC is considerably more demanding than common-sense morality, it is definitely not the case that this in itself is problematic for the theory. Any normative moral theory ought to be more demanding than common-sense morality, if only because common-sense morality leads to a world which is highly undesirable - the world we inhabit.[79] Furthermore, we have established already that the baseline for demand ought to be that a moral theory must not demand that we violate our own integrity. If the common-sense moralist wishes to argue that RACC is still too demanding, then the burden is on them to establish an alternate baseline of demand which is more compelling than maintaining one's integrity. If the argument is made that there are more things that

---

[79] Ashford, p.435

are important to the agent in their lives than their core integral projects, then the counter argument is that while these might be important to the agent, it is the agent themselves that has come to the decision through reflective equilibrium that these things do not necessarily outweigh the demands placed on them by a consequentialist morality. These other values that the agent may have, external to their integral projects, still have value as part of the overall consequentialist calculus, but they cannot be given special weight simply because they are preferred by the agent. To give more value to a set of consequences simply because they are preferred by the agent is nothing more than selfishness and egoism, whereas respecting the integrity of the agent ensures that they have the opportunity to lead a meaningful life that is worth living, while enabling them to operate within a substantial moral framework.

It is now worth reviewing whether RACC manages to avoid the problems plaguing Scheffler's ACP noted earlier. It avoids the first and most problematic of the issues raised against Scheffler, in that while it may allow agents to place too high a weighting on their projects initially (owing to an erroneous belief about which projects are integral and which are not), the process of reflective equilibrium forces the agent to revise this weighting over time, such that the agent is never allowed to hold an inappropriately high amount of integral projects for long. Secondly, the objection regarding resolving the conflict between teleological and non-teleological rationales is   avoided by the RACC owing to the fact that the NACP justification is teleological and has the same telos as the maximising principle: individual integrity. In this way, RACC also avoids the problem of motivation for an ACP raised by Kagan.[80]

The issue of negative responsibility might seem to be alive and well for RACC however. It is still the case after all, that there is no distinction between doing and allowing to happen. The problem is partly dissolved by requiring that the agent hold an internally consistent set of integral beliefs, as we saw in the case of Jim above. If

---

[80]   Kagan 1984, pp.252-4

not killing might lead to more death, then not causing death is an internally incoherent belief. It could be argued that it is only in cases such as this where the distinction really creates a problem for the agent's integrity, otherwise the issue is simply a case of high demand; if the agent is allowed to not act in accordance with her negative responsibility when this conflicts with her sense of integrity, then negative responsibility has no effect on the agent's integrity. Negative responsibility cannot even have the effect of increasing the perceived levels of demand, as this too is limited by the NACP.

However, there a further issue could be raised which may actually be unique to RACC[81]: there are instances wherein the agent is required by RACC to force others to make great sacrifices that ordinary consequentialism would require the agent to perform instead. Supposing a situation where an agent following RACC is able to give up her life to save two strangers. Under act utilitarianism the agent would be required to sacrifice herself, but under RACC she can choose not to in order to protect her integrity. Suppose also that there is another individual present who is also able to sacrifice themselves to save two lives, but for ordinary reasons of self-preservation refuses to do so. Under RACC, because the agent must always do what would make things go best unless she is protecting her own integrity, then in this case she is actually required to force the bystander to sacrifice themselves in order save the two lives (assuming she has the power to do so). This outcome would not be required by act consequentialism alone, as it would be impersonally worse all things considered to force someone to sacrifice themselves than to sacrifice oneself willingly, and as such, it is only found in theories containing an agent-centred prerogative, and might possibly be unique to RACC alone.

This problem could easily be avoided by a theory with an agent-centred restriction which prevents the agent from causing harm to others, but I have already ruled out

---

[81] I am grateful to Dr Joe Horton for raising this concern.

the use of agent-centred restrictions owing to them preventing the agent from always being able to make things go best. Another solution might be to include a restriction against causing harm to others which is only placed on the agent when they are protecting their integrity. However, the clause cannot simply prevent the agent from harming others when protecting their integrity, otherwise the protections on integrity would be very limited owing to likelihood of the agent allowing harm to befall others by favouring their own projects; thus bringing the integrity objection back into play. Therefore, the clause must require that the agent not cause harm specifically as a direct result of their actions. Not only does this raise a distinction that consequentialism does not generally recognise in order to make things go best (the distinction between doing and allowing to happen), but it still means that there are situations wherein the agent might not be able to make things go best. Suppose another situation wherein by thwarting a bomb threat the agent could either save ten lives by sacrificing themselves, or nine by killing the bomber. In this case a clause which prevented the agent from causing harm to others when protecting their own integrity would mean that the agent would be paralysed by contradiction and forced to save no lives at all (unless she chose to override her integrity to make things go best - an option which is always available to the agent).

An answer to the problem therefore must appeal only to the agent's prerogative to protect their own integrity, or making things go best. Fortunately, when we consider what the agent would be required to do if situations like this were acceptable under RACC, then it is clear that this is not as problematic as originally thought. The agent, when using her prerogative not to make things go best, is always required to perform the next best action which does not violate her integrity. The current problem is that occasionally the agent may then be required to force another person to perform the action they should have performed under standard act consequentialism, as in the example above. However, if this is sometimes the case, then it is always the case in any situation where it is possible to do so. So for example, if the agent were to be able

to justify not donating large sums of her own money to charity under RACC, but could steal the an equal amount of money from someone else and donate it instead, then RACC would demand that she do so (assuming this would not make the person she stole from worse off than those she would be helping). This could be true in a great many situations, enough in fact, that it would have a significant impact on the shape of her life going forwards. In fact, the constant need to find alternative solutions involving the sacrificing of other's goods would dominate her life in such a way as to be a significant barrier to her leading a life of integrity. Thus, we can argue that RACC, when requiring that an agent perform the next best action that does not violate her integrity, would never (or at least extremely rarely) require her to force others to perform the action in her stead, owing to the fact that this would in itself violate her integrity through the demands involved in doing so.

Nevertheless, unlike some other attempts to circumvent the demandingness objection,[82] RACC does not deny that there are extreme demands placed on the agent in moral deliberation. It is clear from the fact that there is always the possibility of acting in line with ICMS4 and ignoring the NACP that extreme demands still exist; if the demands were not extreme, then there would be no need for the NACP to begin with. Instead, RACC recognises that the demands that are placed on the agent are extreme, but that despite those demands there ought to be a safe place reserved for the integrity of the agent, if our maximising project is to be rationally motivated. That is, it is the importance of agent's integral projects at the personal level that motivate the rational move to the maximising strategy at the impersonal level, and therefore integrity ought to be treated with due respect at both levels; at the level of maximisation it ought to be the case that we impersonally bring about the conditions for integrity to flourish, and similarly at the personal level. That RACC recognises the extremity of the demands at the impersonal level is reflected in the requirement that the agent engage in the process of reflective equilibrium to effectively minimise the

---

[82]   such as Miller, 2004

hindering effect that respecting integrity at the personal level has on the ability of the agent to maximise at the impersonal level. There is, therefore, no intuition at play within RACC that the demands placed on the agent ought not to be extreme (as there are in moderate views about demand where the conjunct is seen to be between moderate principles of beneficence and moderate demands)[83] - instead it is fully accepted that there are extreme demands that ought to be met, but that this cannot be at the expense of the integrity of the agent if the motivation for the project is to be taken seriously. There is, therefore, no intuition regarding demand at play in justifying the RACC and it is not subject to the sort of objections that we have seen used against these intuitions.

One might suggest, that if there are still extreme demands placed on the agent , perhaps RACC solves the integrity problem, but not the demandingness problem. However, as we noted in chapter 4 when discussing Murphy, the demandingness objection cannot be motivated by the extremity of the demands alone (almost all moral theories make extreme demands in some form or another) and there must be some further justification for why these demands are *too* extreme. I have argued in chapter 5 that this further justification is the effect that the demands have on our integrity, and this is where the baseline of acceptable demand ends - where the agent's integrity is under threat. This being the case, RACC, by ensuring that the integrity of the agent is protected, also ensures that the baseline of demand is never crossed and prevents the demandingness problem from taking hold. Furthermore, because RACC allows the agent to fairly determine what they perceive to be their integral projects which are to be protected, then it is the agent that sets the baseline of demand for themselves. If the agent sets the limits of demand for themselves, using their own internal justification, then it is hard to see how they would be able to find it to be extremely demanding in the first place, or at least not without having an internally incoherent set of beliefs about what constitutes their integrity.

---

[83] Berkey, 2016

The argument might be made that RACC does not differ significantly from Scheffler's ACP, and that rather than being a theory on its own, it might be best thought of as an extension of Scheffler's line of argument. Certainly RACC goes beyond Scheffler's own argument, filling a practical gap that Scheffler himself acknowledges, by developing a strategy for how the agent ought to determine the weighting of her own projects against those of others. Of course, RACC goes further than this by also ensuring that there are some projects which the agent need never be obligated to sacrifice, but it could be argued that Scheffler's ACP could also be adapted to ensure this by weighting integral personal projects so highly that only the most possible pressing concerns could ever override them (e.g. saving hundreds of thousands of lives).

Nevertheless, the crucial difference between the Scheffler's ACP and RACC which can never be reconciled is that RACC has true respect for the integrity of the individual, whereas Scheffler's ACP does not. Under RACC, because the agent's integral projects are protected in such a way that she is never forced to abandon them by the theory, no matter how high the demand, there is no danger of these projects being treated as just one set of projects among many, and as a result, RACC recognises the inalienable nature of these projects. However, by design, Scheffler's ACP cannot make the same distinction. No matter how high the weighting of these integral projects, one can always imagine a situation whereby the agent may be required to abandon them in order to make things go best; even if the agent were to give her own projects a weighting one hundred million times higher than that of any other individual's projects, we can imagine a situation where she may be required to sacrifice her most dearly held project in order to save a billion lives. Perhaps, one might argue, she ought to be required to do so and that Scheffler has the correct approach. I object to this response for three reasons: Firstly, RACC permits her to save the billion lives regardless of her projects, and I expect that almost all morally minded individuals would do so, such that we ought to be able to give people the opportunity to respect

their own integrity, while trusting that in extreme circumstances that they would go beyond their moral duties and perform supererogatory acts. Secondly, if Scheffler's ACP were to be used in the above way, such that only improbably extreme demands could overturn them, then we might ask why this is so? Surely, utilising Scheffler's ACP in such a way would be to admit that there are some projects that we wish to protect against all reasonably probable demands on the agent, and if this is the case, what benefit would there be to choosing a theory such as Scheffler's which appears not to be designed to do so, over a theory such as RACC which clearly is?

Thirdly, at the risk of re-treading old ground, Scheffler's ACP cannot avoid the difficulty of Williams' "one thought too many" objection. By ensuring the possibility that the agent could be forced by some extreme demand to abandon her most integral projects, she is constantly required to weigh her projects against every demand to ensure that she is making things go best. Regardless of the weighting given, Scheffler's ACP requires that the agent think of her integral project as one among many, albeit one with huge importance. RACC on the other hand has no such issue. It asks only that the agent consider whether the project is integral to her, and once she has determined it is so she need no longer consider whether it ought to be considered more important than this or that project. Scheffler, crucially, makes the agent think of her projects in the abstract; they are considered outside of herself, placed among the projects of others, and measured by comparison to their value. Whereas, RACC allows the agent to keep her considerations of her projects internal; they matter because they are integral to her, and not in comparison with those of others. If the integrity of the individual is found in her sense of self, and personified by her projects and relationships, then to consider these outside of herself is to render them devalued immediately, by the very fact that they are seen without the context that gives them their importance to begin with. Picasso's *Guernica* is valuable not because of the paint on the canvas or the form that the paint takes (one could make the argument that it is not a particularly aesthetic painting), but because of the

context in which it was painted and the artist who applied the paint. Similarly, an agent's integral projects are valuable to her in a way that is inseparable from the wider context of her life, and who she believes herself to be, and to abstract them from this holistic view is not necessarily to render them valueless (after all a project which is desirable to be achieved has value in itself), but it certainly diminishes the value of the project in such a way as to render the point of giving it a disproportionate value meaningless to begin with. It is for this reason that one ought not to see RACC as an extension of, or improvement upon, Scheffler's ACP, but a consequentialist theory in its own right, which avoids the integrity objection in a way that it is simply not possible for Scheffler's ACP to do.

9. Conclusion

"And now that you don't have to be perfect, you can be good." - John

Steinbeck, East of Eden[84]

If Reflective Agent-Centred Consequentialism solves both the integrity and demandingness problems, then consequentialism ought to be considered a far more palatable theory by many, by having a space for both the integrity of the agent to flourish and for the agent to perform supererogatory acts. Whether it does or not relies on several arguments I have made throughout this essay: that a baseline of demandingness is required to justify the demandingness problem, that the demandingness problem is a reformulation of the integrity problem and therefore the baseline for demand is where the demand level damages the agent's integrity, and that the integrity problem can be resolved by Reflective Agent-Centred Consequentialism.

I have argued along with Murphy that the demandingness problem requires a baseline of demand, owing to the fact that there is no clear discernible difference between the extremity of the demands placed on the agent by consequentialism and by other moral theories. Without being able to indicate a feature of consequentialism that makes its demands particularly unpalatable, the demandingness problem renders almost all theories overly demanding and fails to accurately reflect our intuitions regarding demand.

I then argued that the demandingness problem is particularly problematic for consequentialism because it is derivative of the integrity problem. That consequentialism demands resource redistribution to a particularly high degree is

---

84 Steinbeck, 1991, p.646

problematic only because those resources would otherwise have been used to pursue and fulfil personal projects without which the character of our lives would be significantly altered and integrity under threat.

Finally I have posited that RACC can resolve the integrity problem (and hence the demandingness problem) by requiring the user to enter into a process of reflective equilibrium in order to determine which projects it is essential that they be able to pursue in order to maintain their integrity and therefore establish a limit to what consequentialism can demand the individual sacrifice in order to make things go best. Because this is set by the agent as a prerogative, under RACC they are able to perform supererogatory acts by sacrificing more than what is demanded, in order to make things go best impersonally.

Despite all this there may be some who would argue that consequentialism is still too demanding and that RACC, while alleviating this demand to some extent, still requires the agent to sacrifice a considerable amount for the sake of making things go best impersonally. To this concern I can say only that that if an agent considers RACC to be highly demanding, then this is only because they recognise the large amount of non-integral projects they allow to dictate the direction of their life. If this were not the case, then the perceived demand on the agent would be relatively low. In other words, if reflective agent-centred consequentialism fails to alleviate the concerns of an agent, then this may say more about the agent's disposition than the theory itself.

Bibliography

Ashford, E. 2000. Utilitarianism, Integrity, and Partiality. *The Journal of Philosophy*. 97(8), pp. 421-439

Bennet, J. 1989. Two Departures from Consequentialism. *Ethics*. 100(1), pp.54-66.

Berkey, B. 2016. The demandingness of morality: toward a reflective equilibrium. *Philosophical Studies,* 173, pp.3015-35

Blincoe, A. 2017. Rescue, Beneficence, and Contempt for Humanity. *Journal of Philosophical Research,* 42, pp.95-114

Braddock, M. 2013. Defusing the Demandingness Objection: Unreliable Intuitions. *Journal of Social Philosophy,* 44(2), pp.169-191

Buss, S. 2006. Needs (Someone Else's), Projects (My Own), and Reasons. *The Journal of Philosophy,* 103(8), pp.373-402

Chappell, T. 2007. Integrity and Demandingness. *Ethical Theory and Moral Practice*. 10(3), pp.255-265

Cox, D. 2005. Integrity, Commitment and Indirect Consequentialism. *The Journal of Value Inquiry*. 39, pp.61-73

Crisp, R. 1996. The Duality of Practical Reason. *Proceedings of the Aristotelian Society*. 96, pp.53-73

Daniels, N. *Reflective Equilibrium.* The Stanford Encyclopedia of Philosophy. [online] Available at: <https://plato.stanford.edu/archives/sum2020/entries/reflective-equilibrium/> [Accessed 20 September 2020]

Darwall, S. 1984. The Rejection of Consequentialism by Samuel Scheffler (Review). In: *The Journal of Philosophy*. 81:4, pp.220-226

Dougherty, T. 2016. The Burdens of Morality. *Thought,* 5, pp.82-5

Eggleston, B. 2014. Act Utilitarianism. *The Cambridge Companion to Utilitarianism*. Cambridge: Cambridge University Press

Gabriel, I. 2017. Effective Altruism and its Critics. *Journal of Applied Philosophy,* 34(4), pp.457-473

Hooker, B. 2002. *Ideal Code, Real World*. Oxford: Oxford University Press

Howard-Snyder, F. 1994. The Heart of Consequentialism. *Philosophical Studies*, 71(1), pp.107-129

Hurley, P. 2003. Fairness and Beneficence. *Ethics,* 113(4), pp.841-864

Jackson, F. 1991.Decision Theoretic Consequentialism and the Nearest and Dearest Objection. *Ethics,* 101(3), pp.461-482

Kagan, S. 1984. Does Consequentialism Demand Too Much? Recent Works on the Limits of Obligation. *Philosophy and Public Affairs,* 13(3), pp.239-254

Kagan, S. 1991. *The Limits of Morality*. Oxford: Oxford University Press

Kamm, F. 1989. Harming Some to Save Others. *Philosophical Studies,* 57(3), pp.227–60.

Metz, T. 2001. Moral Demands in Nonideal Theory by Liam B. Murphy (Review). *The Philosophical Review,* 110(4), pp.614-617

Mill, J.S. 1962. *Utilitarianism*. London: Collins

Murphy, L. 1993. The Demands of Beneficence. *Philosophy and Public Affairs,* 22(4), pp.267-292

Murphy, L. 2003. *Moral Demands in Nonideal Theory*. Oxford: Oxford University Press

Nozick, R. 1974. Anarchy, State and Utopia. Oxford: Blackwell.

Unger, P. 1996. *Living High and Letting Die: Our Illusion of Innocence*. Oxford: Oxford University Press

Railton, P. 1989. Alienation, Consequentialism and the Demands of Morality. *Philosophy and Public Affairs,* 13(2), pp.134-171

Rajczi, A. 2007. Consequentialism, Integrity, and Ordinary Morality. *Utilitas,* 21(3), pp.377-392

Scheffler, S. 1982. *The Rejection of Consequentialism.* Oxford: Oxford University Press

Schmidtz, D. 1990. The Rejection of Consequentialism by Samuel Scheffler (Review). *Nous*, 24(4), pp.622-7

Sidgwick, H. 1907. The Methods of Ethics. London: Macmillan and Co

Singer, P. 1972. Famine, affluence, and morality. *Philosophy & Public Affairs*, 1, 29–43.

Singer, P. 2015. *The Most Good You Can Do: How Effective Altruism is Changing Ideas about Living Ethically*. Conn: Yale University Press

Sinnott-Armstrong, W. 2020. *Consequentialism.* The Stanford Encyclopedia of Philosophy [online] Available at:

<https://plato.stanford.edu/archives/sum2019/entries/consequentialism/>

[Accessed 10 March 2020]

Sobel, D. 2016. *From Valuing to Value: Towards a Defense of Subjectivism.* Oxford: Oxford University Press

Sosa, D. 1993. Consequences of Consequentialism. *Mind*, 102(405), pp.101-122

Tadros, V. 2017. Duress and Duty. In: Bazargan, S. and Rickless, S. 2017. *The Ethics of War.* Oxford: Oxford University Press.

Tanyi, A. 2014. Moral Demands and Ethical Theory: The Case of Consequentialism. *The Bloomsbury Companion to Analytic Philosophy*, 2014, London: Bloomsbury

Tedesco, M. 2011. Intuitions and the Demands of Consequentialism. *Utilitas*, 23(1), pp.94-104

Williams, B. 1973. A critique of utilitarianism. In *Utilitarianism: For and Against* (pp. 75-150). Cambridge: Cambridge University Press.

Wolf, S. 1982. Moral Saints. *The Journal of Philosophy*, 79(8), pp.419-439