

5-2016

Contract Ethics and the Problem of Collective Action

Tyler W. Stolt

University of Arkansas, Fayetteville

Follow this and additional works at: <http://scholarworks.uark.edu/philuht>



Part of the [Ethics and Political Philosophy Commons](#), and the [Other Public Affairs, Public Policy and Public Administration Commons](#)

Recommended Citation

Stolt, Tyler W., "Contract Ethics and the Problem of Collective Action" (2016). *Philosophy Undergraduate Honors Theses*. 1.
<http://scholarworks.uark.edu/philuht/1>

This Thesis is brought to you for free and open access by the Philosophy at ScholarWorks@UARK. It has been accepted for inclusion in Philosophy Undergraduate Honors Theses by an authorized administrator of ScholarWorks@UARK. For more information, please contact scholar@uark.edu, ccmiddle@uark.edu.

Contract Ethics and the Problem of Collective Action

An Honors Thesis submitted in partial fulfillment of the requirements
for Honors Studies in Philosophy

by

Tyler Stolt

Spring 2016

Philosophy

J. William Fulbright College of Arts and Sciences

The University of Arkansas

Acknowledgments

First, I owe the idea of this thesis to Bekka Williams, who, by recommending that I read an article on vegetarianism for a term paper, sparked two years of inquiry on the subject of collective action problems. I am grateful to Dr. Richard Lee for agreeing to advise me in this process, and for his comments over the course of three semesters. I am also grateful to Dr. Warren Herold for his comments, recommended readings, and for always having an open door. I would also like to thank Dr. Zola Moon for providing me with new ways to apply my philosophical understandings and arguments to a real-world empirical and sociological context. I would like to extend my thanks to the staff of Mullins Library for processing my requests for articles through inter-library loan with an unbelievable quickness.

Abstract

Collective action problems are a unique family of puzzles in ethics about what people do together. These puzzles pervade our modern and globalized world. In collective harms, individual actions lead to a serious moral wrong; yet, the individual actions themselves do not seem to be moral wrongs. In collective benefits, individual actions fail to bring about some serious moral good; yet, these individual actions do not seem to be morally wrong. In this paper, I will explore three different ways of grounding reasons why we ought not participate in the production of collective harms (or, refrain from producing a collective benefit). I argue that consequentialism fails to ground these reasons because it has problems in moral mathematics. I argue that deontological theories also fail to ground these reasons, because of problems with group actions. Finally, I conclude that contract theories provide the best way to make sense of these problems, and ground reasons why we ought not (or ought) to participate in the creation of collective harms (or benefits).

Table of Contents

Section 1:	Introducing the Problem of Collective Action	5
Section 2:	Building a Case Against Consequentialism	6
Section 3:	Deontology and the Group Agency solution	30
Section 4:	Contractarianism, Contractualism, and Social Systems of Ethics	43
Section 5:	Conclusions	57
	Works Cited	60

Section 1: Introducing the Problem of Collective Action

There is a family of puzzles in ethics that are commonly referred to as collective action problems. In these cases, a group of people together cause some great harm. However, the actions undertaken by individuals in the group do not seem to be morally wrong. Certainly, many moral issues today take this form. For example, imagine a city with a large car-commuter population. Air pollution in the city is a collectively-caused harm; that is, it is a major harm caused by many people's acts in combination. However, individual acts of driving to work, which contribute to air pollution, do not seem to be morally wrong. There are also cases where individuals fail to produce collective benefits. For example, if enough people donated money to Oxfam, we could save the lives of many people suffering from malnutrition. However, individual failures to donate to Oxfam do not seem to be morally reprehensible. The reason these cases are puzzling is that there is a disconnect between individual acts, which seem to be of no moral significance, and the morally significant outcomes they collectively bring about.

Some people have the opinion that contributing to collective harms is not morally wrong. In fact, this seems to be a common viewpoint. For various reasons, these people argue that so long as individual acts are permissible, whatever collective harms they may cause are not relevant. However, I (and others) have the intuition that contributing to collective harms is morally wrong. Attempts have been made within various ethical viewpoints to arrive at this conclusion. While these puzzles are mainly problematic for consequentialists (ethical theorists who hold that the moral rightness or wrongness of an action is based on its consequences), this problem manifests in different ways in other

ethical frameworks as well. In this essay, I will survey attempts from various ethical frameworks to explain why contributing to collective harms (and/or failing to achieve collective benefits) is wrong, and explore what justifications these solutions have. In looking for a solution to this puzzle, I am concerned with whether or not some approach can condemn contributions to collective harms, and whether or not the explanation of this moral wrong is plausible. It is not only important that it be possible to reach this conclusion, but that we must do so in the right way.

In this essay, I argue that consequentialism cannot provide a satisfactory explanation of why contributing to collective harms is wrong. I also argue that most deontological theories fail to provide satisfactory explanations. A deontological theory of group agency comes close to resolving this problem, but contract theories provide the best answer. Contract theories provide a straightforward and intuitive answer as to what makes participating in a collectively-caused harm wrong.

Section 2: Building a Case Against Consequentialism

Consequentialism is the most initially intuitive framework to use in order to explain what makes contributions to collective harms wrong. Consequentialism is an ethical framework that holds that the rightness and wrongness of actions is based on the consequences those actions have. Utilitarianism is one branch of consequentialism, which elaborates on this central idea to say that acts which have the consequence of producing happiness or pleasure are the right actions, and those that have the consequences of causing harm or displeasure are the wrong actions. The reason that this looks like a promising ethical framework is because there is a strong intuition that what is wrong

about collective harms is that the harm was caused by agents and their actions. To illustrate this, consider an example created by Shelly Kagan. Imagine a fishing village that depends on a nearby lake for its livelihood. “We can imagine that if each fisher restricts their catch to a given, fixed number of fish, the small but stable fish population in the lake will be able to reproduce and sustain itself, and the village’s lifestyle can continue indefinitely. Furthermore, however many fish are taken, one fish more or less would not make a difference to the ability of the fish to successfully reproduce. But if several dozen people take an extra fish each, the fish population will crash, and the villagers will all suffer” (Kagan, 110). There is a close connection between the few extra fish taken (an act) and the destruction of the fishery (the consequence). Consequentialism should be able to explain why acts of overfishing are wrong.

However, explaining the wrongness of these actions is not as easy a task as it may seem. Consider another of Kagan’s examples. Imagine that we run a factory that releases some toxins into the air through a smokestack as a byproduct of production. Next “imagine that the smokestack is sufficiently tall that the pollutants are swept up into the stratosphere, where they are so scattered by the winds that when the toxins do come back down to the surface of the earth, they are spread over a very wide area— indeed, spread so thin that no single individual ever breathes in more than a single molecule from my plant.” Let us stipulate that one molecule is so small that it makes no difference to a person's health. But, because there are thousands of other such factories, many people do take in enough toxins to become ill (Kagan, 108-111). The factory owners fail an important test of causation introduced by Russ Shafer-Landau called the counterfactual

test of causation. He writes “some event is causally efficacious only if, had it not occurred, an outcome would not have occurred just as it did” (Shafer-Landau, 86). We can run this test on any individual factory owner. If this factory owner had acted differently, that is, not polluted, the outcome would still be the same: many people would have gotten sick. This means that the individual factory owner could not have caused the outcome. Because consequentialism grounds moral wrongness in the consequences of actions, it cannot condemn the action of the individual factory owner. Yet, the factory owners together make many people sick.

I have the intuition that the individual factory owners have done something wrong, and I have this intuition at least in part because it seems that the factory owner did make some sort of difference to the outcome in which many people became sick. After all, if each factory owner's molecule of pollution made literally zero difference to anyone's health, then nobody would become sick, no matter how many molecules accumulated in their bodies. We can call these kinds of claims 'make no difference' claims¹. But there are two senses in which the ‘make no difference’ kind of claims can be true. In what we might call ‘weaker’ make no difference claims, what we mean is that the individual makes a very small, perhaps imperceptible difference to a final outcome. In ‘stronger’ make no difference claims, what is meant is that the action makes literally no difference to the final outcome². Some people claim that consequentialism cannot explain what makes either a strong or weak contribution to a collective harm wrong, and conclude that

1 As far as I can tell, Johnathan Glover coined this phrase in his article “It Makes No Difference Whether Or Not I Do It,” 1975

2 This 'stronger' and 'weaker' distinction is based on Glover's original “no difference” and “insignificant difference” distinction, but uses Shafer-Landau's terminology, as it appears in “Vegetarianism, Causation, and Ethical Theory,” 1994

some sort of non-consequentialist theory would be best. But I do not think this is the right way to proceed. For one, the form of consequentialism I have presented above is a straw man: no consequentialist philosopher would actually endorse such a careless account of consequentialism. So, we will turn our attention to consequentialist philosophers who have attempted to resolve the problem presented above, and to the challenge posed by both kinds of 'make no difference' cases. Most proposals are based on the idea that what has gone wrong in the above cases is that we have failed to do our utilitarian calculation correctly. These accounts will differ with respect to what feature of the moral landscape they argue we are failing to account for. Other proposals look at the expected utility of actions. Still others turn to rule utilitarianism. In the following section, I will explore each of these proposals in turn.

A chapter of Derek Parfit's 1984 book "Reasons and Persons" has sparked much of the discussion on this subject. Chapter 3 of this book is called "Five Mistakes in Moral Mathematics." In this chapter, Parfit explains five mistakes that people commonly make when making consequentialist calculations. He argues that if we adjust our consequentialist reasoning so as not to make these mistakes, we will get the right answer in collective action type cases.

The first of these mistakes is called the "share of the total view." On this view, whenever there is a collective action, each agent "produces his share of the total benefit" (Parfit, 68). So, if four of us together save the lives of 100 people, the share of the total view says we are each responsible for saving 25 lives. Parfit shows that this view is mistaken because it has bad implications. For example, suppose that I could save 100

people as a collective action with 4 other people. Suppose also that, at the same time, I could go elsewhere and save 3 lives. There is someone else who could take my place as part of the group of four, but who could not save the other lives. The share of the total view implies that I should take part in the rescue mission in which I save 25 lives, and let 3 needlessly die, because doing so maximizes my share of the total lives I save. If this argument is convincing, we should agree that the share of the total view is mistaken.

The second mistake in moral mathematics is that it is wrong to think "if some act is right or wrong because of its effects, the only relevant effects are the effects of this particular act" (Parfit, 70). This is to say that it is wrong to consider only "the effects single acts," because this act could be "one of a set of acts that together harm other people" (Parfit, 70). Or, we could be part of some group whose actions together harm or benefit some people. Take the example of watering my lawn in a drought. In order to avoid making the second mistake, I must also consider the other acts of 'lawn-watering' done by others, because our acts together might create a water shortage.

The third mistake is related to the second. It says that it is a mistake to "ignore very small chances when they would either affect very many people, or would be taken very many times" (Parfit, 75). The case of voting in a large scale election is a good way to illustrate this principle. There is a very small chance that a single vote in the election will make a difference- where making a difference means swinging the outcome of the election in favor of one candidate over another. But, we can assume that the outcome of an election will make a large difference to very many people, so the importance of the outcome multiplies the (otherwise small) chance that a vote will matter, so voting

becomes very important.

The fourth and fifth mistakes have to do with the very small consequences of actions. Parfit argues that it is mistaken to think that "if some act has effects on other people that are imperceptible, this act cannot be morally wrong because it has these effects. An act cannot be wrong because of its effects on other people, if none of these people could ever notice any difference. Similarly, if some act would have imperceptible effects on other people, these effects cannot make this act what someone ought to do" (Parfit, 75). The fifth mistake differs from the fourth in that it replaces each instance of 'imperceptible' with 'very small.' There is some debate as to whether there can be such things as 'imperceptible' harms and benefits. But if anyone has a problem with the fourth mistake, they can accept the fifth mistake instead.

The reason Parfit takes time to point out these mistakes is that he believes that if a consequentialist avoids making these mistakes, they will get correct answers about contributing to collective harms. I find Parfit's arguments somewhat convincing, but I also think that his discussion here raises more questions than it solves. For example, we should agree that the share of the total view is a mistaken view: we should not merely divide the total harm done by the number of people who produce it³. Parfit suggests that we should instead adopt some sort of 'whole of the total' view, in which each person counts as producing the whole of the total benefit (Parfit, 69). But this faces a few problems. First, the whole of the total view is not an intuitive way to describe the effects of actions. Using Parfit's example involving four people on a platform, the whole of the

3 For another explanation as to why the 'share of the total view' is wrong, see Shafer-Landau, "Vegetarianism, Causation, and Ethical Theory," 1994, p90-91.

total view would have us say that each person saved 100 lives. In a way, this seems false; what we want to say that each person saved 100 lives with the help of three others. The second problem with this revision is that it has no real justification other than that it gets the right answer in this specific case. Of course, Parfit is trying to argue against the share of the total view, and he is not attempting to provide a full-fledged argument in support of the whole of the total view. But I think we should put some pressure on what seems to be an ad-hoc solution.

We can put pressure on the second mistake as well by showing that it raises more problems than it solves. Even if we agree that the premise endorsed by the second mistake is, in fact, mistaken, we still have a lot of work to do in order to assess the boundaries of 'sets of acts' and 'groups of actors.' Let us return to the lawn-watering example. Because I want to avoid making the second mistake, in assessing the consequences of the act of watering my lawn, I must also look at whether or not my act would be part of a set of acts which together causes a water shortage. But in doing so, do I begin by looking at acts which share the same content (lawn-watering)? Or, do I look at any act that has the same outcome (water supply usage)? Depending on the details of the water reserves, there might be an important temporal element as well- imagine that the water supply is replenished every five years. If this is the case, the set of acts in question might include my act of lawn-watering today and someone's act of showering five years ago. Parfit does make a suggestion as to how we might be able to more carefully delimit a set of acts, but what I hope to have shown is that doing so will require some strong positive arguments that make sense of an otherwise vague but complicated idea.

Parfit's arguments are closely related to those of Shelly Kagan's and Johnathan Glover's. These three philosophers have been engaged in an academic dialog about collective action problems for some time. Kagan's and Glover's proposals are within the same family as Parfit's because they are attempts to correct the way we do moral mathematics. But Kagan and Glover introduce an important distinction: they argue that cases of this sort can be sorted into two types of structures. In the first type of case, there is an "absolute threshold" or boundary between two different outcomes (Glover, 173). Kagan calls these 'triggering' cases, because while most acts make a very small difference, one action will 'trigger' a different outcome (Kagan, 118). The most common example of this type of case is voting. It is true that one vote can trigger a radically different outcome in an election between candidate A and B. The other type of structure is that of a "discrimination threshold." In these cases, "a single person's act will push a situation slightly further in a certain direction, but where his contribution, although real, may be too small to be detected when its effects are spread throughout the community" (Glover, 173). Kagan calls these 'imperceptible difference' cases instead. In these cases, there is no sharp trigger, and each act represents a push along the underlying dimension. Consider the collapse of a fishery. There is no single fish caught that serves as the trigger point between the fishery being 'healthy' and 'in decline.' Instead, there is some undetermined point at which the decline becomes noticeable to the fishermen. Kagan argues that "these two sorts of cases- imperceptible difference cases and triggering cases- are importantly different in terms of their underlying mechanisms, and thus require different analyses" (Kagan, 119). With this distinction in place, let's look at Kagan's and

Glover's accounts, beginning with Kagan.

Kagan's main line of argument is to "take seriously the idea of looking to see what difference my individual act makes" (Kagan, 114). Kagan's account makes a conscious attempt to avoid making the fourth mistake in moral mathematics, in that it "insists that results need not be perceptible to be real, and they need not be perceptible to count morally" (Kagan, 114). Kagan is arguing that there are no such things as stronger 'make no difference' cases. "Make no difference" claims are, for Kagan, merely colloquial. In reality, all acts make a difference, and if we take small differences seriously, we can condemn participation in collective harms.

First, let's consider what happens for Kagan in 'pure' triggering cases. Kagan admits that there is a very small chance that a given act will be the one that makes a big difference. For example, in the pollution case, assume that I operate one of a thousand other such factories like mine. Because of this fact, there is a very small (1/1000) chance that the molecule from my factory will be the one to make someone sick. But, Kagan argues, if the outcome is important enough, or if the outcome would affect very many people, we should not discount this chance. Using an appeal to expected utility, we should see that in these types of acts, where an act could trigger something very bad, the expected utility is negative. Thus, the straightforward consequentialist condemns these actions (Kagan, 129).

Now, let's look at how Kagan handles incremental difference cases. Kagan eventually concludes that cases of this sort are really just triggering cases in disguise (Nefsky 2011, 365). Kagan believes that there must always be some sort of triggering

action, otherwise the outcome would not change. He also concludes that "in some cases, there are many triggering acts, each making only a very small difference to the overall outcome," and in others, "only a few triggering acts, each making a large difference" (Kagan, 140). In order to illustrate this point, he develops a case involving purchasing chickens from the store. For brevity, I will summarize the important details from the final form it takes. Suppose that chickens are sold in cases of 25. For every 25 chickens sold, the butcher makes a note. At the end of the day, he calls the farmer and tells him how many more chickens to begin raising, in order to keep up with demand. Each consumer makes an incremental difference to the number of cases ordered, and the person who orders the 25th chicken triggers the next case order. However, if someone purchases the third or 24th chicken in the box, they do not trigger the purchase of another 25 chickens. But of course, the consumers are unaware of what number their chicken is: "I may not know what the actual triggering number, T , is, but I do know that I have a 1 in T chance (more or less) of being part of a cohort that triggers an increased order, and that if I am part of such a cohort then another T chickens (more or less) will suffer. The net expected utility of my act of purchasing a chicken is negative, and so the consequentialist condemns it" (Kagan, 127). Once again, we use the 'familiar consequentialist tool' of expected utility in order to explain why the acts of chicken-purchasing are wrong.

The most important take-away from Kagan is that he thinks all collective action problems have some kind of triggering structure. Since triggering cases are already well-handled by appealing to expected utility, collective action problems are really not a puzzle for consequentialists after all (Kagan, 140). What is wrong with participating in

the causation of collective harms, like the harm that results from very many people purchasing chickens, is that the expected utility of this action is negative. The actual utility of the act may not be negative, but because we are unaware of the probability that our act will be the triggering act, we should act with caution, and say that the expected utility is still negative.

But as Julia Nefsky points out, we have reason to push back on the idea that expected utility can always solve these types of problems. Nefsky argues that "there is no guarantee that the expected utility will come out negative in every triggering case. Whether it does or not depends on the probabilities and on the goodness and badness of the relevant consequences" (Nefsky 2011, 369). The expected utility only comes out negative in cases where there is a small chance of making a big (as in morally bad) difference. The expected utility would not be so negative if, say, the difference to be made was smaller, or if the chance of making such a difference was much lower. One way we can make the chance of making a difference smaller is to increase the size of the cohort. If, for example, the cohort size is 1/25, then there is a 4% chance that I will be the person who makes the triggering action. But in cases where the cohort exceeds 100 people, I have a less than 1% chance of being the triggering person. At this point, despite the risk of making a harmful difference, I may expect the utility to be positive, given the small chance that I will do any harm, and the near certainty that I will derive some pleasure or satisfy some interest by purchasing a chicken.

Second, we have reason to push back on the conclusion that everything is some sort of triggering case at all. As an example, let us return to the initial worry I had about how

smog is created in a city. There is no single triggering point at which the amount of smog in the air becomes especially dangerous or bad. It seems that every car's contribution to air pollution makes the air slightly worse, and every act of refraining from driving makes it slightly less bad. Nefsky relates this problem to 'other familiar sorites problems,' like baldness. The triggering point at which someone is 'bald' is nuanced and unspecific. "Presumably, if you have enough hairs to count as not bald, then you have more than enough hairs: taking one away won't make you bald (nor will it make you partially bald or 'balding')" (Nefsky 2011, 378). From his arguments in other section of the paper, we can infer that Kagan would argue that the fact still remains that there must be some triggering point at which we say someone is 'bald,' regardless of whether or not we can reliably figure out where this point is. But if there is a difficulty in figuring out where the triggering point is, we are also going to have problems figuring out the general size of our cohort of other agents, and so on- leading to a very vague range of potential expected utility values. It seems that Kagan's appeal to negative expected utility in triggering type cases cannot provide a plausible reason why contributing to a collective harm is wrong.

Johnathan Glover's reasoning is very similar to Kagan's, but he takes more seriously the distinction between triggering and incremental difference cases. He does not think that incremental difference cases, or, in his words, discrimination threshold cases, are merely triggering cases in disguise. He is particularly more worried about discrimination cases where the outcome fails to come about, and the 'harm' is never noticed. To illustrate this point, he creates a scenario involving 100 bandits and 100 villagers:

"Suppose a village contains 100 unarmed tribesmen eating their lunch. 100 hungry armed bandits descend on the village and each bandit at gunpoint takes one tribesman's lunch and eats it. The bandits then go off, each one having done a discriminable amount of harm to a single tribesman. Next week, the bandits are tempted to do the same thing again, but are troubled by new-found doubts about the morality of such a raid. Their doubts are put to rest by one of their number who does not believe in the principle of divisibility. They then raid the village, tie up the tribesman, and look at their lunch. As expected, each bowl of food contains 100 baked beans. The pleasure derived from one baked bean is below the discrimination threshold. Instead of each bandit eating a single plateful as last week, each takes one bean from each plate. They leave after eating all the beans, pleased to have done no harm, as each has done no more than a sub-threshold harm to each person. Those who reject the principle of divisibility have to agree."
(Glover, 174-175).

This case is an interesting collective action puzzle because it is true for each villager that if only one bean was stolen from their lunch (perhaps in secret), they would not have noticed any difference. We can say that one bean is below the 'discriminability

threshold' for each villager. This is why each of the bandit's actions is, individually, sub-threshold. This case is particularly puzzling because each bandit can say for each act of lunch-stealing 'I did no harm, because no villager would have noticed if one bean was missing.' But of course, the bandits together do a discriminable amount of harm, and the villagers do notice when their whole lunch is missing. The puzzle that Glover has exposed is that of how hundreds of acts, each of them too insignificant to notice, can add up to such a very large difference. In fact, since the theft of one bean would not be noticed at all, we could even say that the harm done by the theft of one bean is zero. But if this is the case, how can one hundred acts that cause zero harm add up to a large amount of total harm?

In order to resolve this gap, Glover proposes that we should accept the principle of divisibility. What the principle of divisibility says is that, "in cases where harm is a matter of degree, sub-threshold actions are wrong to the extent that they cause harm, and where a hundred acts like mine are necessary to cause a detectable difference I have caused 1/100 of that detectable harm" (Glover, 173). This means that each bandit is doing to each villager 1/100th of a harm. Because they do this to 100 people, each bandit is as morally culpable on the second day as they were on the first day. The principle of divisibility assigns an amount of harm to the bandits based on the harm that is actually felt, not on the amount of harm the bandits think they are causing. So, the bandits are responsible for causing a small but non-zero amount of harm. The idea that we ought to take 1/100th of a harm seriously is related to the fourth and fifth mistakes in moral mathematics. What the principle of divisibility tells us to do is take seriously the small harms done by each of the

bandits, and then aggregate those small harms over a set of actions.

The other kind of case Glover turns his attention to is the original distinction between "stronger" and "weaker" 'make no difference' cases. Remember that in stronger 'make no difference' cases, an action makes literally zero difference to an outcome. For example, imagine that votes were made public in real-time. Not only was the percent of votes for each candidate publicly displayed, but also the percent of people who had yet to vote. Imagine that in some scenario, candidate A will win the election, even if all the people left to vote voted for candidate B. Because these votes are public, all the townspeople who have voted or are left to vote are aware of this fact. This is a stronger 'make no difference' case, because no voter nor group of voters can make a difference to the final outcome: candidate A will win no matter what happens. In these cases, Glover bites a sort of bullet. In these cases, we cannot appeal directly to the difference our act makes to the final outcome, because it can not make a difference to the final outcome. In these cases, we are best to appeal to what Glover calls "side effects" and "spirals." For example, we could appeal to the side effect that "my vote will help keep up the morale of my party, or else it will help to support the system of democratic elections" (Glover, 180). Or, we could appeal to the fact that people would notice the margin by which one candidate won, and perhaps the voters should vote in order to reduce this margin and make a sort of political point. Glover thinks that in many cases where we make zero difference to an outcome, spirals and side effects such as these can almost always still provide reason for action (Glover, 190). So even if our action makes no difference to an outcome, we might still have auxiliary reasons that tell us to do that action anyway.

We have to admit that Glover's principle of divisibility helps to get a more acceptable answer in the bandits case than what we get without it. But I am not convinced that the principle of divisibility provides a strong reason against participating in collective harms. First, it is not clear what justification Glover has for the principle of divisibility itself, other than that it gets the right answer; that is, it makes the Bandit's acts seem wrong. But secondly, we can achieve the same result by appealing to Parfit's second mistake in moral mathematics. Parfit writes that it is wrong to think "if some act is right or wrong because of its effects, the only relevant effects are the effects of this particular act" (Parfit, 70). Glover makes the second mistake in moral mathematics because he does not see the bandits as a cohort, nor does he see their actions as a 'set of actions' (Parfit, 70). From the perspective of an individual bandit, the first and second methods of lunch stealing are different- on one day, the individual bandit steals one lunch from one person, and on the second, 1/100th of a lunch from 100 people. But if we consider the bandits as a cohort, whose acts are counted as a set of acts, there is no difference in the amount of harm done between the first and second type of theft: on day one, (a cohort of) 100 bandits stole (as a 'set' of individual acts of stealing) 100 lunches, from 100 villagers. The exact same is true on the second day. By avoiding the second mistake, we see that there is no difference in moral wrongness between the first and second days. If the result that the principle of divisibility gets could be obtained through different means, it is not clear why we should prefer the principle of divisibility over some other moral mathematical methodology.

I argue that Parfit, Kagan, and Glover all fail to explain why it is wrong to

participate in collective harms. The attempts to explain this so far can all be labeled 'accounting' solutions⁴. What Parfit, Kagan, and Glover are arguing is that if we avoid making certain procedural mistakes in moral mathematical calculations, we can get to the result that it is wrong to participate in collective harms. However, in Parfit's case, the solutions he proposes raise more questions than they solve, requiring a large amount of additional work before we have a theory that can really be applied to cases. Against Kagan, we should push back on the idea that every collective action problem is some sort of triggering case. We should also push back on his assertion that the expected utility in collective harm cases will be negative. It is unclear what work Glover's principle of divisibility is supposed to do, and how it fits in to a larger scheme of moral mathematics.

J. J. C. Smart provides a novel approach to solving this puzzle. Smart's solution is based on a straightforward act utilitarianism calculation that involves elements of game theory. I think this solution differs enough from the ones presented above to warrant a separate section. Smart provides a different way of explaining why we have reason not to contribute to collectively-caused harms that is more in line with the type of rational-egoist thinking we often see in the language of economics. The primary puzzle that Smart is concerned about is a collective action problem that R.B. Brandt uses to argue against utilitarianism. Smart, however, thinks he can resolve this problem. The case in question can be called the "heater case:"

4 This terminology is owed to David. T. Schwartz from his book "Consuming Choices: Ethics in a Global Consumer Age," 2010, p54: but I believe I am using this terminology in a slightly different way than he originally intended. He means that 'accounting solutions' are those which "assert[s] that individual responsibility for collective wrongdoing is straightforwardly determinable" through division, such as in Parfit's "share of the total" view. I am using it to describe solutions which suggest that individual responsibility is determinable, through whatever mathematical mechanism these philosophers might propose.

"[Consider] the case of a utilitarian in wartime England, and it is supposed that there is a governmental request that a maximum temperature of 50 degrees F should be maintained in homes, so as to conserve gas and electricity. A utilitarian Frenchman who is resident in England might conceivably reason as follows: 'It is very unlikely that the vast majority of Englishmen will not comply with the request. But it will do no harm at all if a few people, such as myself, live in a temperature of 70 degrees F. And it will do these few people a lot of good for their comfort. Therefore the general happiness will be increased by my using enough electricity and gas to make myself comfortable.' The Frenchman thus decides to use the electricity and gas" (Smart, 42).

If we accept Kagan and Glover's reasoning, we know that the Frenchman does do some incremental amount of harm, but because this harm is dispersed among all the citizens of Britain, we know that it does not do any one person a discriminable amount of harm, and will likely go unnoticed by others. Where Smart's reasoning diverges from Kagan and Glover's is that he bites the bullet, and is willing to say that the private good to the Frenchman, if it outweighs the harms done to the public, makes using the extra fuel what the Frenchman ought to do. This is a stronger claim than that the Frenchman's actions are merely permissible: Smart argues that if we are committed act utilitarians, we

must admit that using the fuel is what the Frenchman ought to do. But, of course, this is just one reason for action, and Smart thinks there are other potentially overriding reasons that the Frenchman ought not to use the heat. The first is the probability of getting caught, and the punishment that would result from this. "The act utilitarian will have to agree that if his behavior could be kept secret then he ought in this case to use the electricity and gas. But he should also agree that he should be condemned and punished if he should be found out" (Smart, 42). This is because, Smart argues, we must keep the utility of an action separate from the utility of praise or blame of an action (Smart, 42).

Imagine that we change the example, so that everyone involved is a convinced and dedicated act utilitarian. Now, presumably, every citizen knows that every other citizen reasoned in the same way that the Frenchman did, and also saw this loophole by which evading the government decree could increase personal benefit. Of course, if everyone were to reason in this way and everyone used the heat, the results would be "disastrous" (Smart, 42). But, we are holding constant in this situation that the nobody has access to "premisses about what other people will do, and each of them will not know how to plan his actions unless they know what the rest of the people (including the Frenchman) will do" (Smart, 43). Because of these constraints, Smart thinks our best bet will be in applying game theory. The utilitarian, always concerned about maximizing good and minimizing harm, will see that at some point N , where N is the number of people using the fuel, that the aggregate harm to the community will outweigh the benefit to the N individuals who are using the fuel. Smart concludes (rightly, I think) that this number would have to be very small. Using game theory, the Frenchman sees three options:

comply with the government request, use the heat, or "decide to give himself a certain probability of disobeying the government's request." The number of heat-users should be small, and this, combined with the probability of each person being one of the lucky few heat-users will lead to a probability that "will be so near zero that the act utilitarian would not bother to calculate it, but would just obey the government's request" (Smart, 44). Smart's solution is similar to Kagan's, in that the motivation not to undertake the action is the risk of being a person whose act lowers total utility. However, Smart's theory has the benefit of not being dependent on Kagan's distinction between triggering cases and incremental difference cases.

I think the strongest objection to Smart is that we ought to take the original intention of the heater case seriously. R.B. Brandt believed that the conclusion that the Frenchman had a moral duty to use the heat was an objectionable conclusion. I think this case helps shed light on an important underlying structure of many collective action problems. The only reason that the Frenchman seems to have a moral duty to use the heat is because of the scale on which we are evaluating the actions and its effects. The underlying structure is this. There is an individual who would derive a private benefit from using some resource: this is an individual perspective. The harms that result from using the resource are spread among the community at large: a community perspective. Then we continue to take a community perspective when we say that the private benefit to the Frenchman gets counted as a community benefit, because he is a member of the community. The shifting of evaluative perspectives creates this problem. If we take a community perspective from the beginning, we see that there is no reason for the

Frenchman to use the heat. If the benefit is proportional to the harms (and in this case, it seems to be) then the net balance to utility is zero.

There is another related worry that applies to Glover and Smart's appeal to auxiliary concerns, such as being caught by the authorities. If we once again assume that the private benefit to the Frenchman is worth the cost to the public, then he ought to use the heat. Then, we might factor in the negative utility of getting caught. This could motivate the Frenchman not to use the heat. But it could also motivate the Frenchman to find ways of not getting caught. If the Frenchman is, indeed, a committed act utilitarian, then in order for him to choose to devise a way to not get caught, the opportunity cost of doing so must be lower than the pleasure he gets from using the heat. This could easily be the case. What I am arguing is that an appeal to side effects such as 'getting caught' or 'being publicly exposed as someone who does not support the war effort' cannot always override reasons for acting, and in fact, could motivate the Frenchman to find ways of disguising his behavior. Smart's approach also fails to provide strong reasons not to participate in collective harms. In fact, Smart admits that depending on the features of the situation, participating in a collective harm might be what the committed act-utilitarian ought to do; in which case we must rely on an appeal to side effects or the expected utility of an action in order to make it seem wrong again.

Of course, one way of resolving the heater case would be to appeal to the fact that if everyone used extra heat, the result would be disastrous. But in act utilitarianism, where the moral scale of evaluation typically focuses on individual agents, we cannot appeal to these kinds of concerns. Rule utilitarianism, however, can appeal to

universalizability tests. Rule utilitarianism (generally speaking) is an approach in which utilitarian calculations are used to create a set of rules, and if everyone followed these rules, utility would be maximized. I will be focusing on rule utilitarianism as it is explained by R.M. Hare in "Ethical Theory and Utilitarianism" (1982). One of the reasons that Hare, as a supporter of the principle of utility in general, develops a system of rule utilitarianism is that he notices the difficulty in making act-utilitarian calculations. For example, how are we to weigh the present versus future interests of people? What are their true interests in a deeper sense? How do alternative courses of action weigh in, when we consider that the agent in question may not be aware that these alternate courses of action exist? (Hare, 27). This would also resolve many of the difficulties I have brought up so far, including difficulties about the scales on which we evaluate actions, and what constitutes a relevant set of acts.

Hare sets forth a few ideals about how the utilitarian calculus should be carried out. First, he argues that the principle of utility is necessarily just. "If we ask how we are to be just between given the competing interests of different people, it seems hard to give any other answer than that it is by giving equal weight, impartially, to the equal interests of everybody" (Hare, 26). This should solve some difficulty in cases where the benefit is private and the harm is public. We must take a third-person view and assess the interests of all involved equally. Second, when making moral judgments, we must universalize prudence; so that the judging we do is done as rationally as possible (Hare, 25). Lastly, when we consider moral decisions, we must think of the "desires and likings of those whom I take in to account of affected parties" of any course of action (Hare, 29). This

turns our attention to anyone who would be affected in any way by our actions, including ourselves, but only insofar as we are one of the affected party. Hare calls this kind of thinking "level-2" moral thinking. This kind of thinking is 'leisurely,' and done with complete knowledge of relevant facts.

One reason Hare turns to rule utilitarianism is because, in everyday situations, we do not have the time nor information to employ level-2 moral thought. On a day-to-day basis, we need level-1 moral thought, which consists of weighing an action against consistent and clear rules and principles. These rules and principles are arrived at after careful deliberation in level-2 thought. Because we make decisions as they accord with moral rules, Hare thinks problems like the heater case and issues about voting are simply not problematic for rule utilitarians (Hare, 36). This is for two main reasons. First, if we take a level-2 viewpoint, when the Frenchman uses the heat, his action fails the universal test and the prescriptivity test- he could not prescribe that everyone universally act as he does. Second, his act fails on utilitarian grounds, because he is considering his own interests as more worthy than the interests of the people he is stealing from, were they perfectly prudent and thus aware of the theft. Lastly, if we resume level-1 thinking, it would have disastrous consequences if society adopted a general rule that said that actions like the Frenchman's were permissible.

But how exactly is Hare able to completely evade worries about collective action problems? I suggest that it is because he is unafraid to abandon act-utilitarianism and moral mathematics in general. The only moral mathematics that we do are in level-2 thought, but in those cases, we are completely aware of all the relevant facts and interests

at hand. In level-1 thought, we have no reason to employ moral mathematics. This abandonment of act-utilitarian calculation of course upsets act utilitarians, but I do not object to the abandonment of act utilitarianism in this way. There are other more serious objections to rule utilitarianism; within the philosophical community it seems to be a notoriously indefensible position. I do not have the space to list these objections to rule utilitarianism, but we can at least recognize that it stands on shaky ground. However, despite all this, we should interpret the ideas presented in rule utilitarianism as signposts pointing us in the right direction. If Hare is successful because he avoids act utilitarianism, creates a system of ethics based on rules and principles, and has a system of ethics that is universalizable; we should look for other ethical systems that also have these features.

I argue that consequentialist and utilitarian frameworks fail to adequately explain reasons why agents ought not to participate in collectively-caused harms, let alone why they ought not refrain from participating in collectively-caused benefits. The accounting-type solutions of Parfit, Glover, and Kagan fail to explain how, using moral mathematics, we can consistently explain why individuals ought not participate in the causation of collective harms. Like Kagan, Smart suggests that using game theory, the expected utility of an action can be shown to be negative. And like Glover, he suggests that often, the side effects of actions can make them praiseworthy or condemnable for reasons not directly related to the collective action problem at hand. But what I hope to have shown is that these solutions are ultimately not plausible. Hare's response is to abandon moral mathematics almost entirely, and he successfully avoids the puzzle posed by collective

action problems. I suggest that we should interpret Hare's success here as a signifier that perhaps all moral mathematical solutions are problematic. If the problem at hand is a moral mathematical puzzle, perhaps instead of approaching it with different mathematical solutions, we should approach it from non-mathematical ethical frameworks.

Section 3: Deontology and the Group Agency solution

Deontology is an ethical system based on ideas about underlying principles and universalizability tests; and as such, is typically far less mathematical than consequentialist frameworks. More specifically, deontology focuses on how the rightness or wrongness of an action is based on how it discharges or fails to discharge rights, duties, and obligations. Where consequentialists look at the effects of an action, deontologists are generally more concerned with an act's intentions. Initially, deontology looks like it will have no problem handling the kind of problems I have been discussing. First, deontology offers a number of ready-at-hand explanations as to why some of the acts we have been discussing might be wrong. For example, it might be wrong to pollute because doing so violates some right, like the right of people downstream not to be made sick. Or, in the beans and bandits case, we could say that the bandits wronged the villagers by the mere fact of stealing the beans, because stealing violates a moral duty against theft.

But deontology faces its own type of problems in handling collective action cases. One of the central ideas in deontological thought is the idea that 'ought implies can.' 'Ought implies can' is the premise that in order for an agent to be obligated to do something, the obligation must be something they could actually do. This is relevant both

to collectively caused harms and benefits. Imagine some potential collective benefit, like famine relief, that could be achieved if enough people donated to Oxfam. Suppose also that there is some sort of prima facie obligation to end great moral wrongs such as famines. However, the end of famine relief is not a possible action that any individual alone could achieve: achieving this end would require a group of agents. Because ought implies can, individuals cannot be obligated to do what only a group of people could possibly do. This puts us in a peculiar place- our participation in the production of a collective benefit is contingent on the fact that enough others are also participating in this scheme. If others are not participating, we can have no obligations concerning that end. The deontologist has at least some grounds for concluding that “if nobody else is doing X, why should I?”

This, of course, is a straw man portrayal of deontology, and no sophisticated deontological theory of ethics agrees with this conclusion in this way. In the following section, I will explore two common methods for resolving these issues. The first method we could use is to take a Kantian approach. Immanuel Kant was an 18th-century German philosopher whose work is considered a foundation of deontology. But in many ways, Kant’s philosophy is not particularly well-built to handle the kinds of problems we have been discussing. For example, take Kagan’s pollution example. I have the intuition that what the factory owner does wrong in this case is that he causes many people to get sick. If we apply familiar Kantian tests to these kinds of problems, we don’t get intuitive results. For example, Kant’s second formulation of the categorical imperative tells us to always treat humanity as ends in themselves, never merely as a means (Kant, 29). We

could say that the factory owner does wrong because he treats the sick people as mere means to the end of production in his factory. But this account of wrongness seems misguided. The factory owner does not intend for people to get sick, and it is not the case that people must get sick if the work at the factory is to continue. Getting people sick is a side effect of the factory owner's intended actions, and the factory owner may not even be aware of this side effect. It does not seem like an appeal to treating people as ends is going to process this situation properly.

We get similar results if we apply Kant's first formulation of the categorical imperative, in which he says 'act only in accordance with that maxim through which you can at the same time will that it become a universal law' (Kant, 10). Under this formula, the factory owner should ask himself, what if everyone acted in accordance with a maxim such as 'pollution is permissible.' In a world that operated on this maxim, there would be no contradiction in will- that is, the factory owner's attempts to pollute would not be defeated. Second, there is no contradiction in conception- that is, there is no reason we cannot imagine a world in which everybody pollutes: this world would be an unfortunate one, but it is a conceptual possibility. My argument here is not that Kant cannot somehow derive reasons not to participate in collective harms. What I am arguing is that familiar Kantian tools are not easily nor straightforwardly applied to collective action cases.

A more promising account comes from a recent development in deontological theory that argues that groups can be moral agents. If groups can be agents, then they can bear obligations and act in accordance with maxims. The idea is that if groups can be agents, they can be obligated to do things that groups could do, like enter into projects

like producing collective benefits. Also, group agency could work in a negative way, by saying that groups are responsible for collective harms, or for failing to discharge important obligations. In order to do this, deontologists will have to explain a number of things. First, they will have to explain what criteria a group must meet in order to be considered an agent. Next, they will have to explain how group obligations translate to individual obligations. If this can be done, it could yield the result that individuals have strong reasons to participate in groups that can produce collective benefits or avoid collective harms.

Stephanie Collins makes one such attempt at fleshing out an argument for moral groups and group agency. Collins wants her theory to be rooted in standard deontological concepts. For example, it is a common assumption that only agents can bear duties (Collins, 2). So, in order for a group to bear a moral duty, it will have to be an agent. Collin's method of arguing for this conclusion is to show how groups meet basic criteria for moral agency. But Collins wants to resist "relaxing our concept of agency so much that it loses its distinctiveness" (Collins, 3). In order to avoid this, Collins creates two criteria for group agency. First, the group of people must have a decision-making procedure that processes reasons. Second, the decision-making process must output roles. If a group meets these two criteria, it meets what Collins calls the agency condition. These group agents are more accurately called 'collectives' (Collins, 2). In order for an obligation to hold for a collective, it must also meet the 'can condition,' that is, it must be able to bring about the outcome required by the duty.

The agency condition is most important to Collins. The presence of a decision

making procedure seems to do the majority of the work in securing a group's status as an agent. The decision-making procedure is important because it takes in reasons, such as action-demanding duties (Collins, 2). Once these reasons are internalized, the group decision-making process generates three outputs. The first is a set of group aims. The second is a set of individual roles which are "each necessary and jointly sufficient to achieve those aims/s," and third is a distribution of these roles to the members of the group (Collins, 8). These roles are Collins' second criteria for agency. These roles are how group agents act as collectives. It is not enough that the group take in and process reasons; they must try to achieve their ends. The only way that a group can act in a coordinated way is by the assignment of roles. By making this move, Collins is supporting a strict idea about collective action. For Collins, it is not enough that a group merely share an aim or intention. It is also not sufficient if all of the group members do the same action. For Collins, an act is not a collective action unless it is a highly coordinated, role-based effort to fulfill an obligation through the fulfillment of individual roles.

Once a group meets these conditions for collective agency, we must also assess whether or not they meet the 'can condition.' This is an important solution to the problem posed by the 'ought implies can' problem. A collective meets this condition if it 'is sufficiently likely to ϕ , given that it tries' (Collins, 12). Some collectives will fail this condition. Imagine that there are three people present at the scene of a house fire⁵. Let's also suppose that these three people each recognize that there is a prima facie moral obligation to put out this house fire. The group processes this reason using a decision-

5 Credit is owed to Dr. Richard Lee for this thought experiment.

making procedure which distributes roles to these agents. If the group does this, they meet the agency condition. But, the group members might soon see that they fail the can condition: that is, no distribution of roles can change the fact that they lack sufficient numbers to form a bucket brigade that reaches the nearby pond. They cannot put out the fire. We can imagine that if the size of the group was twenty people as opposed to three, they might meet the 'can condition,' because they would have enough people to form an efficient bucket brigade capable of putting out the fire. But, as it stands, the group is not able to put out the fire, and so their obligation to do so dissolves.

Collins gives a sophisticated account of how groups of people can be considered agents and under what conditions, and how obligations relate to collectives. But there is still a piece missing in that Collins has not yet explained how the obligations recognized by collectives translate to individual obligations. This is an important problem. Collins gives an example featuring six beachgoers. Five are on the beach sunbathing, and one is drowning. The five could coordinate their efforts and save the drowning person (Collins, 4). However, it is true that each beachgoer cannot, as an individual, save the drowning person. Because ought implies can, each individual cannot be obligated to save the drowning person. "To avoid this result, we need collectivisation duties. These are individual duties to be responsive to others with a view to there being a group that meets Agency and Can conditions" (Collins, 14). Collin's argument for this is detailed, but in its simplest form, it runs like this: "A (the individual agent) has a duty to secure the means to discharging his duties. If A has a duty to ϕ (where ϕ requires a collective), and ψ is the best means to ϕ (where ψ is A's individual responsive collectivising action), then A has a

duty to ψ " (Collins, 14). This argument helps explain some common intuitions about collectivisation. For example, if you witness a house on fire, you might rightly conclude that there is nothing you could individually do to put it out. However, since you recognize a collectivisation duty, you might attempt to form a collective with the bystanders around you in order to discharge the obligation to put out the fire. The obligation to collectivise only dissolves once the agent has thoroughly tried to form a collective.

However, justifying such a duty is a difficult task. Some people argue that group duties are not translatable to individual duties in any way. Torbjörn Tännsjö makes this argument. He thinks that "no sound moral principle can take us automatically from our collective duty to our individual duty" (Tännsjö, 223). First, he points out that there is no reason to assume that in all cases of this type that the collective obligation can be equally divided into shares and parts. Second, in almost all collective agency cases, in order to meet the 'can condition,' the other members of the group must also participate. Given these premises, we conclude that "if the other one is not doing his share, I am under no obligation to do mine" (Tännsjö, 223). This can be further illustrated by appealing to an example. Tännsjö supposes that two people undertake the collective action of pouring water into the pool and jumping in to it. In this case, the 'jumping in' action is conditional on the pool being filled. If there is no water in the pool, the obligation to jump in dissolves.

While I think that Tännsjö's objection is ultimately mistaken, it does raise an interesting worry for Collin's account. I argued that in order for a deontologist to conclude that individuals have duties not to participate in collectively-caused harms, they

will need to explain away the excuse that “if everyone else is not participating, why should I?” I do not think Collins can successfully avoid this worry. For Collins, a collectivisation duty is only fully discharged once all attempts have been made to convince others to join the collective project. She also argues that the importance of the duty to collectivise provides incentive to “affect the phenomena [of the situation] so as to make the collective’s existence more likely.” But the fact remains that if others refuse to collectivise, the obligation to collectivise dissolves.

An other attempt to explain the source of duties to collectivise comes from Holly Lawford-Smith. Lawford-Smith disagrees with Collins about how the collective obligations translate to individual obligations. She divides collective action cases into four distinct types, and argues that the collective to individual distribution of obligations is different in each type. The first kind of case is an 'incremental good' case, where the action of each member 'makes things a little better' (Lawford-Smith, 9). In these cases, each member has an obligation to take a "capacity-relative share" in fulfilling the obligation. The second type of case is 'joint necessity,' where all the collective members must act to bring about an end. In these cases, individuals also have an obligation to take a capacity-relative share, "unless she has the reasonable belief that at least one other member of the collective will not take a capacity-relative share in fulfilling the obligation" (Lawford-Smith, 10).

The last two types of cases are similar. In 'threshold good' cases, a number of people are required to bring about an outcome; say, four people lifting a couch in order to move it. In these cases, the individual has an obligation to take a capacity-relative share unless

she has the reasonable belief that enough others will take a share and the obligation will be fulfilled without her assistance. In the last type of case, 'threshold good with harm,' a number of people are required to produce an outcome, but if too many people assist, "that would be a hindrance" (Lawford-Smith, 10). In these cases, individuals have the obligation to do their share, unless enough others are already participating, in which case the individual should refrain, so as not to cause unnecessary harm (Lawford-Smith, 10). Lawford-Smith's argument is unique, because she explains that context can effect how collective obligations translate to individual obligations. It also makes potentially useful distinctions between types of cases.

Most importantly, Lawford-Smith's argument shows that in some cases, it can be important to undertake an action even if others refuse to collectivise. In 'incremental good' cases, every action towards some end helps. This seems to be the case for donations made to organizations like Oxfam. Although individual donations will not be sufficient to end world hunger, individual donations do alleviate the suffering of individuals in the world, regardless of whether or not others collectivise and also donate. In other cases, where an action is impossible unless enough others collectivise, it does make sense to admit that individual obligations to collectivise dissolve if others refuse to collectivise. This seems to be the case in the bucket brigade case. If others refuse to help put out the fire, there is no point in one individual attempting to put out the fire. This would be a waste of time, because one individual cannot make the fire any incrementally less worse.

Lawford-Smith's distinctions suggest that there may be important differences between different structures of cases, and that the obligation to collectivise may take

different forms in these different types of cases. But it is not clear why we should suppose that all cases of these sorts really do divide so nicely into these four general structures.

Second, it is not clear if there is any justification for the different distributions of collectivisation duties in these cases, other than that it explains common intuitions.

The last account of group-agency we will look at is Terry Isaacs'. Isaacs works to distinguish between relevant types of groups. According to Isaacs, there are two types of collectives: Organizations, and Goal-Oriented collectives. Organizations are highly-structured groups with clear roles, decision-making procedures, and consistent aims, principles, and interests (Isaacs, 42). Examples of organizations are corporations and universities. Goal-Oriented collectives are more loosely structured, and arise out of "shared understandings and a sense of purpose" (Isaacs, 42). Goal-oriented collectives achieve their ends because the intentions of individual agents 'propel' the collective intentions further. Goal-oriented collectives may be more or less cohesive than Organizations. One other important category is social groups. Social groups are not collectives. In social groups, agents are categorized because of some feature they share in common; such as "women," or "cashiers" (Isaacs, 42). The social group of "cashiers" does not meet even minimal criteria for agency, so they are not a collective agent. But it is important to keep social groups in mind with "respect to the more forward-looking collective obligation," meaning "groups that may have the potential to become effective collective agents" should the scenario arise (Isaacs, 42). So even though "cashiers" are not a group agent, it could be that in some case, the social group of "cashiers" could recognize a collectivization duty in order to solve some problem that the group of

cashiers would be best at addressing.

Isaacs' account also provides an alternate form of explaining negative collective obligation, or obligations to avoid participating in collective harms. Isaacs distinguishes between collective responsibility and collective obligation. Collective obligation is forward-looking, and prescriptive in nature (Isaacs, 44). It describes what a collective ought to do. Collective responsibility, on the other hand, is retrospective and evaluative. Responsibility in this sense concerns the praiseworthiness or blameworthiness of actions (Isaacs, 43). For example, we could say that a collective was responsible for the creation of some harm, and blame that collective for this action. Because collective obligation is future-oriented, it "can play a key role in mobilizing groups of people" to solve problems that require collectives (Isaacs, 44). For example, Isaacs argues that the social group of 'men' are causally important in creating a culture of violence against women. Because 'men' are a social group, they are not a collective, so are not a collective agent. Because of this, we cannot hold the social group of 'men' responsible for this problem. But, we can say that the social group of 'men' is a putative agent, that is, capable of collectivising and forming a collective agent, which could then fulfill an obligation. This putative collective agent could both be held responsible in a retrospective way and obligated to solve problems in a prescriptive sense (Isaacs, 52).

Despite the arguments made by these philosophers, I think that deontological group theory fails to answer some of the major worries about why it might be wrong to participate in a collectively-caused harm, or fail to participate in a collectively-caused benefit. I do not think deontological group agency theories are able to answer the excuse I

presented at the beginning of this section: that is, “if nobody else is doing X, why should I?” or, alternately, “if everyone else is doing X, why should I refrain?” Because of deontology's commitment to the 'ought implies can' thesis, it will always be subject to the problem that if enough others defect from a collective action, agents no longer have an obligation to carry out their collective role. Unless we agree with Lawford-Smith, in which case we could argue that in 'incremental good' cases, agents have a duty to assist even if others defect, since every bit of assistance helps.

Second, Deontology's focus on the highly sophisticated groups with internal decision-making procedures and the ability to process moral duties limits the scope of these solutions to very few real-world cases. In the real world, collectively-caused harms tend to result from the collective unintended effects of many agents acting independently. For example, climate change is caused by many people who use or purchase things that emit or lead to emissions of greenhouse gases into the atmosphere. Climate change is caused by people who drive to work, purchase many kinds of plastics, run coal-fired power plants, and farmers who raise beef cattle. Groups like 'greenhouse gas-emitters' are large, unstructured, and lacking any of the criteria for agency that Collins or others require. Since this is the case, it seems that collective agency theories cannot explain why any of these acts are wrong. Isaacs comes close to answering this puzzle by positing that social groups can be putative agents, or loosely-structured groups capable of collectivising. But in these large social groups, it is unclear where the motivation to collectivise would come from. In order to collectivise, all members of the group would have to recognize the same morally important obligation, and that this obligation could

only be achieved if all the group members cooperated. If deontological group theories cannot handle these kinds of groups, then deontological group theories would only be applicable in a few limited real-world cases.

One other hurdle that group agency theories must overcome is that in many ways, the consequentialist approach is more intuitive. What seems to matter in most collective action cases is the effects of the group action, or the consequences that the group brings about. In Kagan's chicken case, for example, the group of consumers together trigger the raising and killing of additional chickens in order to keep up with demand. As Nefsky points out, the wrongness of the action is intuitively and closely related to the fact that a purchase is causally linked to the raising of more chickens: that is, "it makes a difference to decisions to increase (or decrease) production by the factory farm" (Nefsky 2011, 370-371). Consequentialism focuses on an intuitive part of this problem: the consequences produced by the group. The deontologist has a harder time explaining this intuition. The group of consumers is not a group agent. Perhaps we could say that they ought to recognize a duty to collectivise, or that they are subject to backward-looking collective responsibility for an obligation they failed to discharge. But this kind of thinking seems to abandon the idea that the wrongness of the group action is related to the negative effects it causes. Nefsky argues that because of this fact, we should look for the answer in consequentialism (Nefsky 2012, 29). I do not think that we must necessarily look for the answer in consequentialism, but we should look for an ethical system that maintains some kind of closer link between the wrongness of actions and their effects.

In conclusion, deontological group-agency theory makes an important and useful

attempt to explain why it is wrong to participate in collective harms, or why we ought to participate in the creation of benefits that could only be produced through a collective. However, deontology fails to evade a number of central worries about the conclusions it reaches. I think that many of these worries stem from the deontological framework itself, especially deontology's commitment to the 'ought implies can' idea. Because of this idea, it seems that in almost all cases, a duty to refrain from participating in a collectively caused harm can be defeated by the fact that others refuse to refrain from participating. Second, deontological group agency theory has a difficult time handling cases involving the large, loosely-structured, and diverse groups that cause collective harms or do not create collective benefits in the real world. Last, the deontological framework maintains a deliberate separateness from the moral importance of the effects of actions, which does not line up with a central intuition in collective action problems.

Section 4: Contractarianism, Contractualism, and Social Systems of Ethics

One other ethical theory we should consider is social contract theory. Social contract theory is a family ethical theories which maintain that the rightness and wrongness of actions should be evaluated with regard to an underlying social understanding or agreement. One of the major branches of social contract theory is contractarianism, which argues that people are primarily motivated by rational egoism, but constrain their self-interested desires when it is beneficial to do so. Contractualism, on the other hand, is concerned with the rights of people, and the motivational desire to justify our actions to people we share social arrangements with. Compared to deontology, contract theories have a very different metaethical underpinning. The deontological world

is made up of obligations, which are taken to be weighty, objective things. The contract theorist understands morality in relative terms, and believes that morality is constructed through the interaction of people.

How might contract theories answer the puzzle at hand? Generally, contractarians evaluate the rightness or wrongness of an action based on whether or not it is consistent with some moral agreement. If we are trying to argue that it is morally wrong to participate in schemes that collectively cause harms, we must show that this action is inconsistent with some general moral agreement. Contractualists are concerned with whether or not we can justify our actions to others. For the contractualist, we must ask if participating in collectively-caused harms is an action that we can justify to others; specifically those who are harmed by the collectively-caused harm. I argue that in these cases, the contract theorist can make sense of the wrongness of these types of actions.

First, we will look at the contractarian branch of contract theory. Contractarianism traces back to Hobbes' thoughts on constrained rational egoism. The contractarian story begins by supposing that people are primarily motivated by rational egoism. Rational egoism is the idea that what it is rational for anyone to do is to maximize their own self-interested aims. But of course, in a world with scarce resources and a society where every individual acts only to maximize their self-interest, our efforts to achieve these aims will lead to chaos. The most common example of this is a collective action problem called the tragedy of the commons, first described by Garret Hardin in 1968. The commons is a public grazing area. Each shepherd, hoping to maximize his personal economic return, will turn as many sheep out to graze as possible. If every shepherd who uses this public

grazing area did this, the resource would collapse, and all the shepherds would make no money. If each acts in a purely self-interested way, they will obtain suboptimal results. In this case, individual rational egoists produce a collective harm.

Contractarianism suggests that the foundation of ethics is in agreements between agents to constrain their rational egoist behavior in order to achieve the best possible ends for everyone. An example of this would be if the shepherds set grazing limits. This would constrain their behavior, but in a way that ensures the maximum sustainable profit for each person involved. For example, each shepherd could agree to set out five sheep each. Doing so would provide modest profits for the shepherd, but would also guarantee that no collective harm is created. In fact, contractarian ethics seems to have been what was on Hardin's mind when he concluded that the only solution to the tragedy of the commons was "mutual coercion mutually agreed upon" (Hardin, 1247). But Hardin was not a philosopher, and was not intending to write a philosophically tight argument. In the following section, we will look at philosophical approaches to puzzles of this sort, and assess their plausibility in explaining the wrongness of contributing to collective harms.

Gilbert Harman is one philosopher who has developed an advanced concept of contractarian ethics. Harman argues that "morality arises when a group of people reach an implicit agreement or come to a tacit understanding about their relations with one another" (Harman, 3). One of the building blocks of Harman's theory is the idea that inner moral judgments only make sense when weighed against "motivational attitudes" like "goals, desires [and] intentions" (Harman, 9). Inner moral judgments are inner judgments about how someone ought or ought not to have acted, or whether it was right or wrong of

someone to act in a certain way (Harman, 4). Harman argues that it is not enough for these kinds of judgments to be made based on an agent A, considerations C, and a type of act D; but that we must also include the motivating attitudes that the agent has.

These motivating attitudes take a specific form in that 'they derive from an agreement.' Harman argues that motivating attitudes are "intentions to adhere to a particular agreement on the understanding that others also intend to do so" (Harman, 11-12). These agreements are reached through a bargaining process. "In order to further our interests, we form certain conditional intentions, hoping that others will do the same" (Harman, 13). But of course, not everyone will have the same conditional intentions. Through implicit bargaining, we discover what intentions we can compromise on, and share as a group. When we say that someone ought to do something, say, recycle, what we are saying is that given the act of recycling in a certain circumstance, and the shared moral intention to 'live in a cleaner city,' they ought to do so.

But Harman bites the bullet in a case where the agent in question does not share this moral intention. Imagine an agent who does not have the moral intention to live in a cleaner city, although the vast majority of other citizens do share this intention. Because this person does not share a morally relevant attitude, we cannot make full-fledged inner judgments about what such an agent ought or ought not to do. Since the cleanliness of a city does not motivate this person, we cannot say their act was inconsistent with this agreement. But Harman also admits that in cases like these, we are not without ways of convincing such agents that these intentions are ones that they ought to share. For example, "we can argue by appealing to inconsistencies, incoherencies, and whether or

not an agreement is self-defeating" in order to show someone that their moral intentions are in some sense, bad or wrong (Harman, 16). Also, we can appeal to the 'ought of rationality,' the 'normative ought to be' and the 'moral ought to do' in order to convince someone that they ought to share in our widely-held moral intentions (Harman, 6). So even if we cannot say that the non-recycler does something wrong, we can say that they ought to share the moral intention to keep a city clean, and that if they shared this moral intention, they would see that not-recycling was wrong.

David Gauthier's approach to contractarian ethics is not based on inner moral judgments, but instead on a fuller account of rational egoism. In his article "Why Contractarianism?", he sets out to prove that agreements about the pursuit of rational interests form the basis of morality. Gauthier wants to ground his moral reason in a way that coheres with "psychological states" like "desires and beliefs" (Gauthier, 16). Through a process Gauthier calls deliberative justification, we make choices and commit to actions which "maximize the agent's expected utility, where utility is a measure of considered preference" (Gauthier, 19). Deliberative justification is not moral, because it is not other-regarding. But morality is found within deliberative justification, and has the same function; it is a process by which we make choices and commit to actions. But unlike deliberative justification, moral justification is necessarily other-regarding, and places constraints on choices and actions. Gauthier's overall argument is that morality arises when, considering self-interested action, we begin to regard the actions of others.

What Gauthier has in mind is the fact that unconstrained deliberative justification often leads to intuitively bad results. "In many situations, if each person chooses what,

given the choices of the others, would maximize her expected utility, then the outcome would be mutually disadvantageous" (Gauthier, 22). We have already seen how unconstrained self-interested action led to the tragedy of the commons. But in these cases, the best possible results could be obtained if everyone cooperated. In these situations, "each person can see the benefit, to herself, of participating with her fellows in practices requiring each to refrain from the direct endeavor to maximize her own utility," (Gauthier, 23). These types of agreed upon-constraints form the basis of morality. Morality is the set of "constraints that would obtain agreement among rational persons who were choosing their terms of interaction" (Gauthier, 25).

Considering some of the objections to contractarianism will help further define the theory. The classic challenge to any contractarian theory of morality is how to answer the question of 'why be moral?' Agreements between individuals lead to mutual benefits, but in some cases, the mutual benefits are smaller than the benefits that could be obtained if the agent acted in a purely-self interested way. If we return to the example of the tragedy of the commons, we see that each shepherd would do moderately well if they mutually agree to constrain their behavior. But one shepherd could do best if everyone else constrained their behavior, and that one person maximized their sheep. Rational egoism seems to lead to the temptation to be non-cooperative, or to free-ride on the efforts of others. In "The Logic of Collective Action," Mancur Olson argues that without some sort of pressure that encourages compliance, "rational, self-interested individuals will not act to achieve their common or group interests" (Olson, 2). On this subject, Elinor Ostrom wrote that at the heart of these kinds of problems is the free rider problem: "Whenever

one person cannot be excluded from the benefits that others provide, each person is motivated not to contribute to the joint effort, but to free-ride on the efforts of others. If all participants choose to free-ride, the collective benefit will not be produced. The temptation to free-ride, however, may dominate the decision process, and thus all will end up where no one wanted to be” (Ostrom, 6).

Gauthier and Harman both have ways of responding to this challenge. Harman does not take it to be a serious challenge that we might make but not actually keep agreements. He writes "To agree in the relevant sense is not just to say something; it is to intend to do something, namely, to intend to carry out one's part of the agreement on the condition that others do their parts" (Harman, 15). If an agent agrees to constrain their behavior, but then acts in a way that is contrary to this agreement, Harman would say that this person never really intended to agree in the first place. Gauthier points out that anyone who fails to keep a moral agreement will be "excluded" from future participation in cooperative schemes, so it is really most rational that all agents make and keep agreements (Gauthier, 25). Anyone who merely pretends to keep an agreement in order to achieve some strategic advantage would be excluded and possibly sanctioned for this kind of behavior. Further, the contractualist has a straightforward way of explaining the wrongness of this action: an action is wrong if it runs contrary to the shared agreement. Any act of non-compliance is obviously wrong.

I think the most serious challenge to any contractarian thesis is that it seems that it cannot always explain why unjust societal institutions are wrong. The classic example of this is that of a society like the prewar American south in which slavery is a widely

accepted institution. The objection runs like this. In this society, there is an explicit moral agreement that certain humans are to be regarded as property, and bought and sold. So, acts of buying and selling human beings cannot be condemned, because they do not run against this agreement. In fact, they are perfectly consistent with this agreement.

Harman's answer to this challenge is that moral systems are always driven towards consistency and a lack of arbitrary distinctions. We can judge the agreements that certain societies or groups come to because they do not cohere with facts about the world, or are based on arbitrary distinctions, and are therefore not general (Harman, 20). In the example of slavery, we could say that although the actions of slave owners are consistent with an agreement, the agreement itself is based on an arbitrary distinction between races of people. Gauthier's response is that, as self-interested rational egoists, each agent in a society has reason to ask whether or not there is some other agreement or set of agreements under which they could do better. It is clear that in the case of slavery, the slaves would do better under a set of moral agreements that did not include the agreement that certain human beings were property. Gauthier says this would lead to moral instability, and provides reason for the disadvantaged people to ask for a renegotiation of moral agreements (Gauthier, 27).

I think this challenge to contractarianism is serious, but I think contractarians have a good way of addressing these problems. This is because rational egoism provides constant pressure for tacit agreements to come to light, and for serious questioning about whether or not the agreements we have made are equitable and reasonable. Consider this example. Today, city planners share a tacit agreement that cities are

designed around personal automobile transportation. Harman argues that "someone may habitually act in accordance with the relevant understanding and therefore may be disposed to act in that way without having any more or less conscious intention" (Harman, 10). In this way, the city planners are acting in accordance with a shared tacit understanding that cars deserve primacy over pedestrians and other forms of transportation. When a large group of people acts in accordance with this tacit agreement, problems like smog and air pollution can be created. It seems that tacit agreements can lead to the creation of collective harms. However, rational egoism provides motivations to constantly revise our agreements. If it is the case that a standing agreement leads to collective harms, those harmed by that agreement have reason to ask for a renegotiation of that agreement.

Contractarianism is able to properly ground reasons why it is wrong to participate in collectively caused harms, or refrain from participating in collectively-caused benefits. I argue that contractarianism explains the wrongness of these actions in an intuitive way: it is because doing so runs against important social agreements. In cases where a social agreement leads to the creation of a collective harm, then, rational egoism and the desire to have a stable set of agreements puts pressure on revising these agreements. Once these agreements have been brought to light and restructured, we can condemn any action that fails to avoid causing this collective harm. In the case of providing collective benefits, the answer is simple: we have reason to participate in cooperative efforts to produce collective benefits when doing so can be agreed upon. If we agree that we ought to cooperate in order to solve some problem, such as world hunger, then any act that runs

counter to this agreement is wrong.

Why, then, should we agree to cooperative schemes such as ending world hunger in the first place? Why should we agree to not participate in activities that lead to collective harms? Agreeing to these types of behaviors often entails a significant change from the status quo for the agents involved, and these changes might come at significant costs. Despite these costs, I argue that we cannot refuse to form cooperative agreements of these types because we cannot justify our refusal to do so to those who would be affected by this agreement. For a further explanation of this position, we turn to another kind of contract theory.

T.M. Scanlon is the most well known proponent of the other major branch of contract ethics, which is referred to as 'contractualism' in order to distinguish it from contractarianism. Contractualists begin from a fundamentally different starting point than contractarians. Instead of supposing that everyone is primarily motivated by rational egoism, Scanlon bases his system on the motivating force of "the desire to be able to justify one's actions to others on grounds they could not reasonably reject" (Scanlon 1982, 116). Fundamental to the idea that this is a morally motivating sentiment is the idea that other people are persons, and worthy of respect. This implies something like the idea that we may not use people as mere means.

The central thesis of Scanlon's contractualism is that "an act is wrong if its performance under the circumstances would be disallowed by any system of rules for the general regulation of behaviour which no one could reasonably reject as a basis for informed, unforced general agreement" (Scanlon 1982, 110). Scanlon's use of 'reasonably

reject' as opposed to 'rationally agree to' is a deliberate difference between his theory and contractarianism. For example, if we remember the heater case, it might be rational to ignore the government decree and use the extra heat, because it would provide a self-interested sort of benefit to the Frenchman. However, it would not be reasonable of the Frenchman to ignore the harm he is doing to the national supply of fuel. In some cases, what it is rational to do is not the reasonable thing to do.

Of course, not every rejection of a principle will count as a reasonable one. For example, imagine that we are considering implementing a system for famine relief. Could anyone reasonably reject a principle that required people of moderate and high incomes to donate a small amount of money to a famine relief fund? We could imagine someone of moderate or high income attempting to reject this principle, but this rejection could not be reasonable. We would need to "consider the weightiness of the burdens it involves, for those on whom they fall, and the importance of the benefits it offers, for those who enjoy them" (Scanlon, 1998, 208). Nor could this rejection be justified to the people who would be affected by the principle, namely, those whose lives could be saved by the implementation of this principle. Under Scanlon's theory, "when we consider a principle, our attention is naturally directed first to those who would do worst under it. This is because if anyone has reasonable grounds for rejection, it is likely to be *them*" (Scanlon 1982, 123; italics in original).

One of the other central features of Scanlon's argument is that when considering how a principle affects people, we consider only affects for individual people. He says that a central value of contractualist morality is "its insistence that the justifiability of a

moral principle depends only on various *individuals'* reasons for objecting to that principle and alternatives to it" (Scanlon 1998, 229). To illustrate the importance of taking this perspective, Scanlon creates an interesting case we can call the "world cup" case: "Suppose that Jones has suffered an accident in the transmitter room of a television station. Electrical equipment has fallen on his arm, and we cannot rescue him without turning off the transmitter for fifteen minutes. A world cup match is in progress, watched by many people, and it will not be over for an hour. Jones' injury will not get any worse if we wait, but his hand has been mashed and he is receiving extremely painful electric shocks. Should we rescue him now or wait until the match is over?" (Scanlon 1998, 235). In this case, there is a singular person who is suffering a great deal of harm. However, there is also a group of people enjoying the collectively-shared benefit of watching the world cup game. I have the intuition that we should rescue Jones, no matter how large the group of world-cup viewers is.

This can be explained through Scanlon's individual perspective. Scanlon's individualist restriction forces us to compare the complaints⁶ of individuals in this scenario. On one side, we have one individual football-match-watcher, whose complaint is that rescuing Jones would mean missing a significant portion of a game he is enjoying. On the other, we have Jones, whose complaint is that he would be forced to endure torture only so that the football match would not be interrupted. In this case, it is easy to see that Jones' claim against torture is stronger than the claim that can be made by any individual match-watcher. Jones' claim is able to limit the weight of the collectively-

6 The idea and terms of an 'individualist restriction' and the 'complaint model' of Scanlon's theory comes from Derek Parfit in Chapter 21 of Volume 2 of his book "On What Matters"

shared benefit of watching the football match. Scanlon's individual perspective explains why, in some cases, we cannot be obligated to participate in the creation of collective benefits: we cannot be obliged to participate in the creation of collective benefits when doing so would require putting some great burden on someone else, such as Jones. In the same way, we cannot have a duty to refrain from participating in some collectively-caused harm if doing so could be reasonably rejected.

On Scanlon's view, individual agents can make reasonable rejections. This puts a limit on our participation, and refines moral agreements so that they do no serious harm. Under contractarianism, rational egoism supplies constant pressure to refine and revise our moral agreements. Contractualism defines another way these agreements can come under pressure: if any single person has a reasonable rejection to it. This pressure can be explained by applying Scanlon's theory to the case of air pollution in a commuter city. The current tacit agreement supports the fact that cars are the primary method of travel in the city. This agreement leads to air pollution and smog in the city. Scanlon's theory turns our attention to individuals affected by this agreement, perhaps those who have health issues do to the air quality. Can a car commuter justify their action to a person who develops respiratory issues? I do not think so. We should seek an agreement which nobody could reasonably reject.

In general, contract theories have the most intuitive framework for understanding collective action problems. Under both main forms of contract theory, we can ground straightforward reasons why it is (usually) wrong to participate in a scheme that results in a collectively-produced harm. We can also ground straightforward reasons why it is

(usually) impermissible to refrain from participating in cooperative activities through which we could produce collective benefits. Scanlon's theory helps set limits on reasons for participating in these kinds of schemes by introducing the idea of reasonable rejection. I think there are a number of reasons these theories are successful. First, they are able to reach the above conclusion. But more importantly than that, they do so in the right way. They reach the above conclusions by appealing to intuitive ideas about why actions are wrong. Contract theories are a socially-based ethical theory, so it makes sense that they can be successfully applied to socially-based problems.

Compare contract theories to deontological ones. I argued that deontologists could not avoid the challenge posed by the excuse that "if everyone else is doing X, I have no duty to refrain from doing X." I also argued that our obligations to collectivise and form group agents with others were defeated in cases where others refused to recognize the importance of an obligation or refused to collectivise in general. Contract theory provides a better framework for understanding this excuse. Let us return to Collin's drowning case. Imagine that you are one of the five people on the beach. If the others refuse to collectivise, we have a straightforward reason to condemn this action: refusing to participate in the saving of a drowning person runs against common social agreements like 'assist those in need, especially when this could be done at little cost.' Second, the acts of the non-collectivisers cannot be justified to the person who is drowning. In both contract theory and deontology, we reach the same conclusion; that the beachgoers who fail to collectivise do something wrong. But contract theories ground the wrongness of this failure in a way that is relative to the person who is drowning; not to an obligation.

We can also compare contract theories to consequentialist ones. Consequentialists focus entirely on the consequences of an action, which is an intuitive way to ground the wrongness of participating in some collective scheme that produces bad consequences, such as pollution. I argued that deontology is not an intuitive solution, because it generally avoids a discussion of the consequences. Contract theories are not similarly oblivious to the consequences of actions. Instead, contract theories ground the importance of consequences in a meaningful way: consequences matter when they affect people we have moral agreements and understandings with. Then, consequences matter insofar as they violate or support these moral agreements. I argue that contract theory is able to maintain the intuitive connection between the wrongness of participating in certain collective actions and the consequences of those collective actions. For contract theorists, it is not just that consequences matter, it is that they matter to people.

Section 5: Conclusions

In this essay I have argued that in looking for ways to explain what is wrong with participating in a collective harm, or refusing to participate in a collective benefit, we should use a contractualist framework. I have also argued that consequentialist and deontological theories fail to provide consistent and plausible explanations. Consequentialist solutions of the moral-mathematics type proved to be a rabbit hole of moral mathematical solutions. Depending on the evaluative scale, depending on the principles in play, depending on the consequentialist tools used, the results could change. Further, the principles invoked to achieve these results were rarely justified in any way other than that they achieved the answer they were designed to find. Mostly, the focusing

on the consequences relevant to collective action problems does not seem to show the whole picture. Contract theories add to this picture by supposing that the consequences that affect people in collective-action situations can be evaluated not just for their content, but how they fit in to a larger picture that includes the relationships between the agents in the situation.

Deontological theories that include ideas about group agency do give a fairly plausible account of why it is wrong to participate in collective harms. The idea that there might be individual duties to be responsive to others with the idea that a collective could form in order to address some problem is a good way to explain a basic intuition in collective action cases. However, it is hard to overcome problems that stem from underlying principles in deontology, like 'ought implies can.' Second, Contractualism can also explain why we ought to be responsive to others in order to create collectives.

I am not arguing that it is impossible for a consequentialist theory or deontological theory to ground strong reasons of the sort I am looking for. Some accounts are more plausible than others. Even some of the accounts which seem initially implausible could, with enough revision, be made to get the right answer. However, some of these revisions are ad-hoc, question begging, or otherwise unjustifiable. What I was seeking was the ethical framework which most straightforwardly and intuitively justified reasons to act in collective benefits, and reasons not to participate in collective harms. I believe I have argued that the contract framework is that framework. This is not to say that the contract framework is without problems. Initially, there are worries about the so-called 'freerider' problem, and there are worries about justice and fairness in making

agreements. However, I think that contract theories are able to respond to these challenges in a way that supports the conclusion that we can ground reasons of the kind I am looking for.

I conclude that the contract framework is the best ethical model to use in explaining what is wrong about collective action problems. I have been arguing in favor of the intuition that there is something wrong about participating in a scheme which collectively creates a harm, even if the individual acts do not seem to be moral wrongs. I have also argued that there is something wrong about refraining from participating in schemes where, if enough people participated, some important benefit could be collectively-produced. We should move away from explaining these intuitions in terms of moral mathematics or objective obligations. Instead, we should favor theories that capture these intuitions in terms of social arrangements and what we can justify to others.

Works Cited

- Brandt, Richard B. "Moral Obligation and General Welfare." *Ethical Theory; the Problems of Normative and Critical Ethics*. Englewood Cliffs, NJ: Prentice-Hall, 1959. 380-406. Print.
- Collins, Stephanie. "Collectives' Duties and Collectivisation Duties." *Australasian Journal of Philosophy* 91.2 (2013): 231-48. *Philpapers*. Web. 2 Feb. 2016.
- Gauthier, David. "Why Contractarianism?" *Contractarianism and Rational Choice: Essays on David Gauthier's Morals by Agreement*. By Peter Vallentyne. New York: Cambridge UP, 1991. 15-30. Print.
- Glover, Jonathan. "It Makes No Difference Whether or Not I Do It." *Proceedings of the Aristotelian Society Supplementary Volumes*. Vol. 49 (1975): 171-209. *JSTOR*. Web. 1 Sept. 2015
- Hardin, Garrett. "The Tragedy of the Commons." *Science* 162.3859 (1968): 1243-248. *Science*. Web.
- Hare, R. M. "Ethical Theory and Utilitarianism." *Utilitarianism and beyond*. By Amartya Sen and Bernard Williams. Cambridge: Cambridge UP, 1982. 23-38. Print.
- Harman, Gilbert. "Moral Relativism Defended." *The Philosophical Review* 84.1 (1975): 3-22. *JSTOR*. Web. 22 Oct. 2015.
- Isaacs, Tracy. "Collective Responsibility and Collective Obligation." *Midwest Studies In Philosophy* 38.1 (2014): 40-57. Web. 15 Mar. 2016.
- Kagan, Shelly. "Do I Make a Difference?" *Philosophy & Public Affairs* 39.2 (2011): 105-41. *JSTOR*. Web. 25 Aug. 2015.

Kant, Immanuel. *Groundwork for the Metaphysics of Morals*. Ed. Johnathan Bennet.

N.p.: n.p., n.d. *Early Modern Texts*. Web.

<<http://www.earlymoderntexts.com/assets/pdfs/kant1785.pdf>>.

Lawford-Smith, Holly. "The Feasibility of Collectives' Actions." *Australasian Journal of Philosophy* 90.3 (2012): 453-67. *Philpapers*. Web. 2 Feb. 2016.

Nefsky, Julia. "Consequentialism and the Problem of Collective Harm: A Reply to Kagan". *Philosophy & Public Affairs* 39.4 (2011): 364–395. Web.

Nefsky, Julia. "The Morality of Collective Harm." (2012). UC Berkeley: Philosophy. Retrieved from: <http://escholarship.org/uc/item/9s49q22t>

Olson, Mancur. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard UP, 1965. Print.

Ostrom, Elinor. *Governing the Commons: The Evolution of Institutions for Collective Action*. N.p.: Cambridge UP, 1990. Print.

Parfit, Derek. "Chapter 21 How the Numbers Count." *On What Matters*. Vol. 2. Oxford: Oxford UP, 2011. N. pag. Print.

Parfit, Derek. "Five Mistakes in Moral Mathematics." *Reasons and Persons*. Oxford: Clarendon, 1984. 67-86. Print.

Scanlon, T. M. "Contractualism and Utilitarianism." *Utilitarianism and beyond*. By Amartya Sen and Bernard Williams. Cambridge: Cambridge UP, 1982. 103-28. Print

Scanlon, Thomas. *What We Owe to Each Other*. Cambridge, MA: Belknap of Harvard UP, 1998. Print.

Schwartz, David T. *Consuming Choices: Ethics in a Global Consumer Age*. Lanham,

MD: Rowman & Littlefield, 2010. Print.

Shafer-Landau, Russ. "Vegetarianism, Causation and Ethical Theory." *Public Affairs*

Quarterly 8.1 (1994): 85-100. *JSTOR*. Web. 16 Feb. 2016.

Smart, J. J. C. "An Outline of a System of Utilitarian Ethics." *Utilitarianism; for and*

against. By Bernard Williams. Cambridge: U, 1973. 3-67. Print.

Tännsjö, Torbjörn. "The Morality of Collective Actions." *The Philosophical Quarterly*

(1950-) 39.155 (1989): 221-28. *JSTOR*. Web. 19 Oct. 2015.