

Minimizing Harm:
Three Problems in Moral Theory

by

Alexander W. Friedman

B.A., philosophy (1997)
B.A., mathematics (1997)
Brandeis University

Submitted to the Department of Linguistics and Philosophy
In Partial Fulfillment of the Requirements for the Degree of Doctor of
Philosophy in Philosophy

At the

Massachusetts Institute of Technology

June 2002

© 2002 Alexander W. Friedman
All Rights Reserved

The author hereby grants to MIT
permission to reproduce and to
distribute in print, paper and
electronic form, in its thesis
documental system, or in part

Signature of Author.....

Department of Linguistics and Philosophy
May 23, 2002

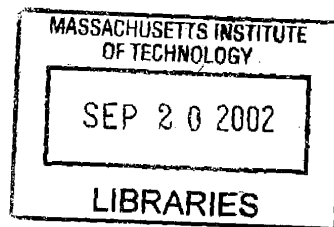
Certified By.....

Judith Jarvis Thomson
Professor of Philosophy
Thesis Supervisor

Accepted by.....

Vann McGee
Professor of Philosophy
Chairman, Committee on Graduate Students

ARCHIVES



MINIMIZING HARM:

THREE PROBLEMS IN MORAL THEORY

by

ALEXANDER W. FRIEDMAN

Submitted to the department of Linguistics and Philosophy
On May 23, 2002 in partial fulfillment of the
Requirements for the Degree of Doctor of Philosophy
in Philosophy

ABSTRACT

Chapter 1 (Distance and Morality): This chapter is a response to Frances Kamm's "Faminine Ethics: the Problem of Distance in Morality and Singer's Ethical Theory". I argue that Kamm fails to produce a pair of cases in which a moral difference is present that is not attributable to factors other than distance. I claim that all of Kamm's pairs of cases that purportedly show that distance *per se* matters in morality, and in which Kamm's moral intuitions seem plausible, can be subdivided into two categories: (1) the cases are not, contrary to first appearances, properly equalized for all non-distance factors that might reasonably be expected to play a moral role; and (2) the cases are properly equalized, but only in virtue of stipulations that are very hard to believe or even fully comprehend; thus our intuitions refuse to reflect those stipulations even if we seem to accept them. I also point out that Kamm's attempts at explaining why distance could possibly matter in morality fall far short, as she herself admits. I conclude that there is no reason for us to believe that distance matters in morality, and offer an explanation of why it might nevertheless appear to us that it does - there are numerous factors that apparently are of moral significance, and that are normally correlated with distance. In the process, I offer a few methodological remarks and suggestions, the importance of which, I believe, transcends this particular discussion.

Chapter 2: (The Problem of Numbers): In this chapter I discuss whether or not numbers matter in morality in the way that most people think they do, i.e. in the way which is best exemplified by the claim that all other things being equal, if two non-overlapping groups of people cannot both be saved, we should save the numerically larger group. My goal is to defend John Taurek's criticism of the standard view of the role of numbers in morality, and to establish Taurek's position, despite its shortcomings, as a viable alternative to the standard view. After pointing out the

under appreciated fact that in "Should the Numbers Count?" Taurek has two independent arguments for his conclusion, I proceed to defend both arguments from some of the more prominent objections that have been raised against them. I also discuss the role that fairness plays in the debate. Ultimately, in light of considerations derived from several of the objections that I address, it becomes clear that Taurek's first argument has to be rejected. In addition, several modifications have to be made in Taurek's overall position in order to make it consistent and plausible. However, the second argument, which focuses on the claim that interpersonal aggregates of losses or suffering carry little, if any, moral significance, and which, I argue, is a much more powerful and illuminating argument to begin with, still remains to challenge the standard view. But Taurek's position is not without serious difficulties either - in the last section I point out and try to address several problems which should be a concern to anyone who is sympathetic to Taurek's views, including the problem of very large numbers. I do not claim to have the answers, and at best just offer a direction in which one may look for a solution.

Chapter 3 (The Return of the Trolley): This chapter is about the Trolley Problem. Very roughly and simplistically stated, the problem is to explain why it is apparently morally permissible to divert a runaway trolley to a track where it will kill one innocent person rather than five, but not morally permissible to kill an innocent healthy person in order to use his organs in saving the lives of five others. I examine in detail several prominent purported solutions to the problem, and find them all severely deficient. In light of the systematic failure of efforts to solve the Trolley Problem, I suggest that perhaps no solution has been found because there is no solution to find - the actions in question in the two cases fall on the same side of the moral permissibility line. I proceed to give a positive argument for the claim that diverting the trolley is not morally permissible, and try to give an explanation of how the intuitions of the majority could have been so deeply mistaken. Since my proposed "solution" relies heavily on the alleged moral role of the making/allowing distinction, I conclude with a discussion of that distinction as well as a defense of its moral significance.

Thesis Supervisor: Judith Jarvis Thomson

Title: Professor of Philosophy

ACKNOWLEDGEMENTS.

First and foremost, I would like to thank my thesis advisor, Judith Thomson, for her help, support and encouragement, without which I would not be at the stage at which I am right now, either academically or in life as a whole. My other two committee members, Alex Byrne and Ralph Wedgwood, also provided invaluable assistance in the form of comments, suggestions, discussion and lots of patience. In addition, I greatly appreciate Vann McGee's putting up with me during the first two years of my graduate studies, when I had no idea what I was going to do or how I was going to do it, and he had the thankless task of being my advisor. I also want to express my gratitude to all the graduate students, faculty and staff for creating a wonderful, friendly and thoroughly enjoyable atmosphere at the department, as well as for discussing some of my ideas with me. I would like to thank my wife, Alisa - even though she never managed to make it all the way through my dissertation, she made a valiant effort nevertheless, while also listening to my constant attempts to sound out everything that was going wrong with the papers that make up this thesis. I am greatly indebted to Professors Yourgrau, Berger, and Samet of Brandeis University for introducing me to philosophy and thus making all of this possible. And, last but not least, I would like to express my undying gratitude and appreciation to the inventors of KOOSH - the true sport of a real philosopher.



Table of Contents

<u>Chapter 1</u>	<u>Distance and Morality.</u>	9
	Section I	An Introduction to the Problem	11
	Section II	Does Salience Matter?.	16
	Section III	Supernatural Abilities	22
	Section IV	On Kamm's New Definition of Distance.	27
	Section V	On Different Kinds of Nearness and Kamm's ranking of them.	33
	Section VI	A few more of Kamm's Arguments for PDM addressed	44
	Section VII	Some Final Thoughts on the Subject.	52
<u>Chapter 2</u>	<u>The Problem of Numbers</u>	61
	Section I	Introduction to the Problem	63
	Section II	Taurek's Arguments.	66
	Subsection 1	A Few Preliminary Thoughts.	66
	Subsection 2	Taurek's First Argument.	68
	Subsection 3	Taurek's Second Argument.	73
	Section III	Derek Parfit and Agent-Relative Permissions. . .	77
	Subsection 1	Parfit vs. Taurek's Second Argument - "Sums of Suffering"	80
	Subsection 2	Parfit vs. Taurek's Second Argument Continued - Maximin	85
	Subsection 3	Parfit vs. Taurek's First Argument - Agent-Relative Permissions	93
	Section IV	The Objection From Transitivity and Other Criticisms by Gregory Kavka	102
	Subsection 1	Taurek on Rights and Obligations	102
	Subsection 2	The Transitivity Objection	109
	Section V	Arbitrariness, Restrictions on Knowledge and Other Objections by Ken Dickey	118
	Subsection 1	Arbitrariness and the Inadvertent Demise of	

	Subsection 2	Taurek's First Argument	119
		The Argument from Probabilities	126
	Section VI	The Problem of Fairness	134
	Section VII	Some Final Thoughts on the Subject.	146
Chapter 3		<u>The Return of the Trolley</u>	157
	Section I	Introduction to the Trolley Problem.	159
	Section II	Some Attempted Solutions.	163
	Subsection 1	Introduction and the Doctrine of Double Effect.	163
	Subsection 2	New Threats versus Old Threats.	171
	Subsection 3	A Few More Attempted Solutions.	179
	Section III	Kamm and her Principle of Permissible Harm	186
	Section IV	Judith Thomson and Hypothetical Prior Consent	197
	Section V	Back to the Killing/Letting Die Distinction.	211
	Section VI	A Few General Remarks on the Killing/ Letting Die Distinction	227
	Subsection 1	Jonathan Bennett and the Quantitative Approach	227
	Subsection 2	The Contrast Strategy	237
	Section VII	When Does the Killing/Letting Die Distinction Make a Moral Difference?	244
<u>Bibliography</u>		255

-Chapter 1-

Distance and Morality

DISTANCE AND MORALITY.

Section I.

An introduction to the problem.

Let's consider the following moral dilemma. Suppose that right outside my house a driver loses control of his car and it smashes into a lamppost. I can see that the driver is alive but very badly injured. It seems that my course of action is clear - I have to call an ambulance and then run out and see if I can be of any assistance. But suppose I also know (say, from a report I've just heard on TV) that if I do not make a phone call pledging a small sum of money to an international relief organization in the next 30 seconds, that organization will not be able to provide the next food shipment to a starving village in Somalia and at least one child will die of starvation. Suppose, furthermore, that by the time I would be done calling an ambulance it will be too late to save the child. What should I do? My moral intuitions, as well as those of more or less everyone else, seem to give a clear answer - I ought to call an ambulance. Examples such as this have convinced many people, and among them many philosophers, that the moral pressure¹ to help is higher with regard to events nearby than with regard to events that are distant from us. At the very least, it seems that in some cases

¹Most authors in the literature use the term 'obligation'. But since that term is possibly a loaded one and may mislead us into thinking that we know what is behind it, I prefer to use the more vague 'moral pressure'.

the distance between the agent and a stranger is relevant to our intuitions regarding how much moral pressure the agent is under to help the stranger. That, however, creates a serious difficulty in moral philosophy - how can we justify these intuitions? This problem - the problem of finding a theoretical justification for our moral intuitions to the effect that distance makes a moral difference - has become known as the *Problem of Distance in Morality* (PDM).

But, as Frances Kamm correctly points out at the beginning of "Famine Ethics: the Problem of Distance in Morality and Singer's Ethical Theory", which constitutes a groundbreaking attempt to analyze PDM, before we can deal with PDM we have to first understand what it is, i.e. the problem needs to be stated correctly. It should be immediately clear that my original example is ridiculously insufficient to generate PDM. There are many differences between the plight of the crashed motorist and the relation in which I stand to that plight on the one hand, and the plight of the starving child and the relation in which I stand to his or her plight on the other. To name just a few: the motorist's life is in danger because of an accident, the child's - probably because of a "basic social injustice"; the motorist is quite likely to be a member of my community, perhaps even a friend of a friend, the child is definitely neither; the motorist is not a child²; assisting the motorist requires no financial expenditures on my part, assisting the child does; if I call an

²The age of a victim may matter to some people's intuitions, so I am including this factor as well.

ambulance, I can be relatively certain that the medical personnel will do everything in their power to save the motorist, whereas I cannot be as sure of what the relief organization is capable of doing, let alone of what it would actually do; if I do not help either the motorist or the child, the motorist will probably die much sooner than the child; and, for all I know, maybe there would still be time for the child to be saved somehow even if I assist the motorist first; in some countries, I may be required by law to assist the motorist³, but nowhere would I be required by law to send money to a starving child in Africa, provided that the child has no legal connection to me; and, last but certainly not least, I am presumably in the best position out of anyone in the world to help the motorist, especially if we assume that there are no other houses nearby and the road is empty, whereas many other people have heard the reports on TV and are in as good a position to help the child as I am - that might make a moral difference even if I am 100% certain that no one else will, in fact, help the child.

The reason why all of these factors make my example insufficient to generate PDM is simple: the fact that our intuitions tell us that we should help the motorist may be, and quite probably is, caused by some or all of the differences, other than distance from me, between the motorist's and the child's situations, and the way I relate to them. And while it might, should we find the need to do it, be very difficult, if not impossible, to find a theoretical justification for the purported

³France is, I believe, an example of such a country.

intuition that distance matters in morality, we should not have much of a problem explaining why some of these other factors matter in morality. To quote Kamm, "if we are trying to find out whether a factor *x* matters *per se* in our intuitions, we must construct a set of comparable cases, one with factor *x* and one without it, and hold all other factors in the two cases constant. I call this "equalizing the cases" ⁴. Kamm quite rightly points out that none of the pairs of examples that can be found in the literature so far, are properly equalized. She devotes a significant portion of the remainder of her paper to trying to produce adequately equalized pairs of examples in order to elicit our intuitions regarding whether or not distance by itself matters in morality. Her conclusion is that all other things being equal, i.e. even after all other differences are filtered out, our intuitions still leave us with the sense that we are under stronger moral pressure to help someone who is near us than to aid someone who is distant.

My example at the beginning of this paper was deliberately chosen to be as poor as it is in demonstrating PDM (it is even more inadequate than the few pre-Kamm examples found in the literature) in order to emphasize exactly how hard it is going to be to make all other things equal. I am going to argue that while Kamm's requirement that the pairs of cases being considered need to be equalized is a necessary one, she only manages to succeed

⁴F.M. Kamm "Feminine Ethics: the Problem of Distance in Morality and Singer's Ethical Theory", Singer and His Critics, ed. by Dale Jamieson, Blackwell's 1999; p.179.

All future quotations from Kamm are from this article.

partway. I will try to show that in all of her examples, whenever there is a difference, as far as our moral intuitions are concerned, between two cases, only one of which is a case of nearness, the difference can be fully explained by factors other than distance - namely, some psychological, epistemological, and, possibly, even sociological factors that Kamm did not, or was not able to, equalize for. Thus I will argue that at least for now we do not have to solve the problem of distance in morality because at this point it has not been established that we are faced with such a problem to begin with. Finally, I will attempt to show just how unlikely it is that distance could possibly matter in morality, but at the same time try to explain why we initially tend to think that it does.

I should point out that I do not claim to be able to show conclusively that distance doesn't matter in morality. In particular, I agree with Kamm that if we can find at least one case or, even better, a group of similar cases in which our moral intuitions clearly and conclusively tell us that it is distance and not some other factors that make the moral difference, then we are faced with a PDM, albeit perhaps one with a very limited scope⁵. So, just generating a number of cases in which distance

⁵The scope would be limited because we would only be faced with a PDM when a particular factor or group of factors is present. Not everyone even agrees that this is possible - it has often been argued that if a factor is morally relevant, it is always morally relevant, even though other morally relevant considerations may sometimes outweigh it. But this claim is controversial - see, for instance, Shelly Kagan's "The Additive Fallacy" in ETHICS Problems and Principles (ed. by Fischer and Ravizza) for an argument against it. Fortunately, the bothersome possibility that distance is morally relevant, but only sometimes, need not concern us until and unless we are presented with at least

does not matter, or in which we mistakenly thought that distance mattered until we discovered the real source of our intuitions, cannot, no matter how large the number of cases is, conclusively establish that distance plays no role in morality and thus that there is no PDM. However, if I am successful in showing that in none of the cases that have been considered so far, do we have any reason to believe that distance is doing any moral work, and if I can give grounds for thinking that any purported example of distance's role in morality with which we might be faced in the future will, upon due examination, reveal the same inescapable flaws, then, at the very least, the burden of proof shifts back to Kamm, and others who think that distance *per se* matters in morality, to show why we should believe that it does.

Section II.

Does salience matter?

After an introductory discussion, Kamm sets out to dispute the suggestion that perhaps it is salience of need of a "victim"⁶ that generates the additional moral pressure to help in cases where the victim is nearby, where salience of need "refers not

one situation where distance does, indeed, make a moral difference.

⁶Following Kamm's terminology, I am going to refer to people requiring aid as victims, regardless of how appropriate such usage would be in everyday discourse.

only to the obviousness and inescapability of noticing need, but also to the continuing imposition of this knowledge on us"⁷. First, Kamm correctly points out that a victim's need can be salient even if the victim is very far away. She suggests that the need may become salient to us through an unusual faculty such as long-distance vision⁸. On the other side of the coin, the need of someone suffering close by may very well not be salient. Therefore, there certainly can't be a perfect correspondence between proximity and salience and so it can't be the case that in all situations where we have a difference in distance we also have a corresponding difference in salience.

However, Kamm still needs to show that in cases where the two differences do coincide, it is distance and not salience that is doing the moral work; otherwise, it could simply turn out that distance never matters in morality, but in some cases the coincidental difference in salience does make a moral difference. One reason why we should be immediately suspicious of this suggestion is that if there is a group of people suffering the exact same plight in exactly the same way, but one of them finds a way to make his need more salient to us or just simply to stand out in the crowd (Kamm gives an example of one of a group of

⁷Kamm, p.181.

⁸However, for the point that Kamm is trying to make here, it doesn't even seem to be necessary to introduce supernatural abilities - a distant victim's plight may become salient to us in any number of ways, e.g. a report on television or in a newspaper or a verbal account vividly imparted by a reliable and trustworthy third party. But more on supernatural abilities later.

starving people dressing up in a clown suit and "much more dramatically exhibiting his need"⁹), it seems inappropriate to conclude just on the basis of that that we are under greater moral pressure to help this particular person than to help any of the others, even though the psychological pressure might be greater in his favor. But Kamm's main argument against the suggestion that salience of need can generate moral pressure is the following: it is true that salience adds significant psychological pressure to render aid; however, if there is no strong moral pressure on me to help that is independent of the salience, there doesn't seem to be anything wrong with eliminating the salience - for example, by turning off the TV or disengaging my long-distance vision or even taking a sleeping pill - even if it means that I will not end up helping the victim. Let's call this **Intuition 1**. Kamm also presupposes a general principle, which I will call the

Non-Elimination Principle 1: All other things being equal, it is not morally permissible to deliberately change your situation so as to eliminate a factor which is generating moral pressure on you to aid someone simply¹⁰ in order to reduce or altogether avoid that moral pressure.

Non-Elimination Principle 1 in conjunction with Intuition 1 yield

⁹Kamm, p.183.

¹⁰As Kamm points out, it is obviously morally permissible to change such a factor if that is required in order to help the victim. For instance, if the moral pressure on you to aid the victim of an accident is partially generated by your being the only person who is aware of the accident and so in a position to help, it is clearly morally permissible for you to call an ambulance, even though once you do you will no longer be the only person to know about the accident. On the other hand, taking a fast-working sleeping pill to make sure that you are no longer in a position to help does not seem to be a morally permissible option.

the conclusion that in cases where nothing other than salience generates any moral pressure to aid, salience can't be generating any moral pressure to aid either.

But that still leaves open the slim possibility that in cases where something other than salience is already generating some moral pressure on us to help, salience of need may bring additional moral pressure to bear on us. After all, as Kamm states, if there is, antecedently to any salience-related considerations, strong moral pressure on me to help someone whose need also happens to be salient, if eliminating the salience will make it less likely that I will help even though I am still under strong moral pressure to do so, it seems that it would be morally wrong of me to eliminate the salience. Let's call this **Intuition 2**. But there is a better explanation available for Intuition 2 than the claim that salience can sometimes generate additional moral pressure. That explanation is provided by another relatively uncontroversial general principle which Kamm appears to endorse - a principle which I will call the

Non-Elimination Principle 2: If there is genuine moral pressure on you to help someone, it is not morally permissible to alter any factor (even one that is purely psychological and is not generating any of the moral pressure), the alteration of which will make it less likely that you will help¹¹.

Thus, especially with the clown-suit example in mind, there

¹¹An exception would be a case where there is an even stronger moral pressure on you to do something else that is incompatible with the original act of aiding being considered. In that case, eliminating the salience, and, perhaps, other psychological factors as well, might be the only way for you to enable yourself to do what you are under greater moral pressure to do.

doesn't seem to be any reason to believe that salience can ever generate moral pressure, and every reason to believe otherwise.

Kamm makes two further comments at this point, which she believes support her contention that distance does matter in morality. (i) She says that the fact that if we are near someone who requires aid and whom we feel great moral pressure to assist, it is intuitively morally wrong to move further away if it is done simply in order to reduce the moral pressure to help, is an additional reason for thinking that it is the nearness itself that sometimes generates the increased moral pressure. (ii) She also says that "another sign that I believe I have obligations to those who are near me because they are near is that if a free device is available that will help me detect need that is near to me, I think I have an obligation to get it, whereas I believe I have less obligation to get a free long-distance device in order to know what need exists at a distance"¹².

However, if stated as general principles, I do not share either of those intuitions, even though in certain particular cases my intuitions might conform to those general principles. As far as (i) is concerned, my intuitions will only coincide with particular instances of it when whichever feature(s) of the case is(are) generating the moral pressure on me to aid, will be coincidentally removed by my moving further away from the victim. That makes it question-begging to rely on the general principle as a reason for believing that distance matters in morality. About

¹²Kamm, p.183.

(ii), at least as stated, I just don't seem to have any moral intuitions either way except, perhaps, that I am not under much moral pressure to obtain either device. I would probably have intuitions regarding some particular instances of (ii), but only if the situation were described in much more detail, with a number of morally relevant circumstances filled in.

In, fact, I suspect it might be the case that Kamm is working backwards here. The two generalizations might not be antecedently and independently held moral intuitions, confirming her claims about the role of distance in morality. Rather, these intuitions might be the results of applying some principles similar to the Non-Elimination Principles in conjunction with Kamm's belief that distance matters in morality. At any rate, even if these are Kamm's genuine moral intuitions, it is unclear what they are based on and so they can't by themselves provide a substantial reason for someone who doesn't share those intuitions to believe that distance matters in morality. So, even though I believe that Kamm's argument is conclusive in showing that salience cannot make a moral difference, I have to disagree that she has given us a good reason to believe that distance can. On the other hand, and this is extremely important to the entire discussion of the problem of distance in morality, a change in salience can definitely give us the illusion of a change in the level of moral pressure on us to act by changing the level of psychological pressure that is affecting us. Therefore, in order to be able to judge whether or not distance ever matters in morality based on examples, we have to make sure that, among other things, the

examples are equalized for salience. Otherwise, our intuitions may simply mislead us into believing that it is distance that is making a difference in the level of moral pressure, when in fact there is no difference there to begin with. After all, we are human and are all too often subject to all kinds of psychological pressures to which we might not, in our calm and reflective moments, choose to be subjected when faced with decisions of grave importance, such as decisions regarding the kinds of moral questions that are the subject of this paper.

Section III.

Supernatural abilities.

Another issue which I think affects a significant portion of the analysis of whether or not there is a PDM is Kamm's discussion of unusual abilities (which, in most cases, are, as I will explain shortly, more properly classified as supernatural). Kamm claims to have an intuition that she is morally permitted to "turn off" her long distance vision in order to "eliminate my salient knowledge of the far need"¹³. On the other hand, if knowledge of a near need comes via an unusual machine or sense, e.g. a machine or sense that can detect heart beats through a wooden door, Kamm believes that she is not morally permitted to "ignore his plight

¹³Kamm, p.183.

once I have found out about it or to move so I am far".¹⁴ I take it that Kamm offers this as an additional argument for the purported fact that distance matters in morality, because it is not whether or not the method of finding out about the need in question is "usual" or "unusual", but whether or not the victim is near, that makes the moral difference. The argument is aimed against those of us who would say that the reason we have the intuition that it may be morally permissible for me to turn off my long distance vision and go on with my life as usual, but not to turn around (or close my eyes or walk further away) so as not to see the victim with my regular vision in a putatively equalized case, is that in the first case, but not the second, my knowledge of the victim's plight comes via an unusual sense. Such a person would, *prima facie* conclude that there is nothing in this web of intuitions to support Kamm's thesis.

But now let's consider Kamm's response¹⁵. First of all, her choice of the adjective "unusual" both to describe a machine that can detect heartbeats through a wooden door and to describe the sense of long-distance vision is highly misleading. The machine in question is unusual only in the sense that very few, if any, people currently own one. However, there is nothing supernatural (on a par with a time machine or a matter-energy transporter) or

¹⁴Kamm, p.183.

¹⁵So as to not be distracted from the real issue in this section, let us ignore the fact that these cases are not equalized for many extremely relevant factors. For instance, if someone is right outside my door, I seem to be in the best position to help and I am likely to be the only person to know of the victim's situation - not so, most likely, in the long-distance vision case.

magical (on a par with an instantaneous wish-granting machine) about it. In fact, with a little knowledge of engineering, programming and human biology, we would probably be able to construct such a machine quite easily. On the other hand, long-distance vision seems supernatural and, perhaps, even magical. An additional reason for thinking that is given by the question of what justifies us in classifying so-called long-distance vision as a case of vision. If long-distance vision is to allow us to observe events that are taking place hundreds or even thousands of miles away, and that seems to be intended by Kamm, it clearly can't be just like regular vision but with an unusually large range, as regular vision allows us neither to see through obstacles nor to see beyond the horizon. Thus long-distance vision would probably need to fall under the category of extra-sensory perception or clairvoyance, but definitely not under the category of vision¹⁶. Also, whereas it is clear how we would use the heartbeat detection device - it would have an on/off switch, it would perhaps alert us if there were a heartbeat or heartbeats within a certain distance from our door, etc. - we can't make heads or tails of what would be involved in using long-distance vision, how reliable and safe it would be, and so on.

This is important for at least two reasons that I can think of. Neither of these reasons can make a difference to the actual level of moral pressure we are under, but they can, and, I believe, almost without fail do, affect our moral intuitions.

¹⁶I am indebted to Judith Thomson for directing my attention to this point.

Thus they can, so to speak, make our moral intuitions lie to us, just as salience can. The first consideration is that of how other people would react once they found out that I did not help the victim. In most cases, it seems obvious that, all other things being equal, if I both found out about the situation and could have helped using common, garden-variety, very well understood means, people would tend to judge me much more harshly than if supernatural, magical, or even just highly unusual means were involved. While this difference appears to be purely psychological (although maybe in some cases sociological as well), as all factors, such as the reliability of the means and the probability of success in using them, have supposedly been equalized for, and so cannot affect the moral pressure on me to act, it can certainly unduly affect my moral intuitions.

And that brings us to the second, more important issue. I am supposed to be aware of the fact that all the factors other than the "unusualness" of the means of information and/or the means of aiding have been equalized for. In particular, I am supposed to be aware of the fact that in both cases the probability of success, should I decide to help, is the same. However, it may not be enough, especially if the cases become very complicated and/or divorced from our current physical reality, to simply stipulate that a number of factors have been equalized for. A further step may be needed - we might have to convince our intuition-generating faculties that the factors have really been equalized for. Otherwise, the stipulations end up being useless, as our intuitions do not really track the actual moral situation

delimited by the stipulations.

For instance, if I am contemplating whether or not to act on information received via highly unusual, perhaps even supernatural or magical, means, or whether or not to employ equally unusual means for the rescue, it may not be enough to just stipulate that the means of receiving the information are just as reliable as the usual ones and that the means of rescue are just as safe and likely to succeed as some usual ones¹⁷. As a general rule, life has taught us to be very mistrustful of the highly unusual, especially if potentially life-threatening situations are involved. That cannot be overcome by mere stipulation once the situation gets complex or science-fictional enough. This is a very significant problem, I believe, for a large part of Kamm's analysis, and so I will come back to it and to how it appears to be distorting some of her conclusions later.

But, in any case, the two considerations that I have just offered provide a plausible a plausible account of the source of Kamm's intuitions, which are the topic of this section, without relying on the supposition that distance plays a moral role. In fact, it is very likely that Kamm's intuitions would be greatly affected by a shift from the use of an unusual device, such as the

¹⁷It may be interesting to note that in the case of some unusual, and perhaps even supernatural, machines, our intuitions may have relatively little trouble coping. For instance, we have read and watched so much science-fiction about matter-energy transporters and time machines, that in hypothetical moral examples involving their use, we might actually be taken in by an illusion that we understand them. None of that, I think, can apply to cases of long-distance vision as well as to a few other examples that Kamm uses and that I will discuss later on in this paper.

heartbeat-sensing machine, to the use of a supernatural sense, such as clairvoyance.

Section IV.

On Kamm's new definition of distance.

Whether or not we believe that distance matters in morality, it is important to figure out what we mean by "distance" in the formulation of the problem. Kamm has several suggestions on the topic. First, she claims that when figuring out the distance between the agent and the victim, it is wrong to simply measure the distance between "where the agent stands and where the needy stranger stands". One reason for that is the **REACH** case -

Reach: "Suppose I stand in one part of India, but I have very long arms that reach all the way to the other end of India, allowing me to reach a child drowning in a pond at a distance".¹⁸

Kamm says that her intuitions seem to be treating this case as a case of nearness and thus suggests that the proper measurement of distance, as it is relevant to morality, is from the "extended parts of the agent's body to the extended parts of the victim's body"¹⁹. She adds that the moral pressure on you to render assistance is only slightly decreased if the part of you that can extend near the victim is not efficacious in aiding the victim;

¹⁸Kamm, p.186.

¹⁹Kamm, pp.186-187.

and so in such a case the moral pressure is still significantly greater than in a case where you don't have any parts that can extend to the victim or any of the victim's extended parts. This is a claim that I will come back to at the end of this section.

Kamm's first assertion may seem *prima facie* plausible. Certainly Reach is morally different from

Standard: I stand in one part of India and in another part of India there is a child drowning in a pond. In this case I don't have extendible arms that can reach her.

But Reach and Standard are not equalized for an entire host of factors other than the extendibility of my arms. The most important of these factors is that if extending my arms to the drowning child in Reach is as simple, safe and reliable as using my arms is normally, then it seems that it is very easy, as well as costless and safe, for me to save the child, and I can be virtually certain that the rescue attempt would be successful. No one would argue that such factors are irrelevant to the moral pressure on me to save the child, but this is not the issue - distance is what we are concerned with. Thus, to evaluate what Kamm is saying, Standard has to be greatly modified. At the very least, I should have the means for saving the child which are (i) just as safe, costless, and sure to produce the desired result as in Reach and (ii) such that I actually believe, when evaluating the case, that the means are just as safe, costless, and sure to produce the desired result as in Reach.

Equalizing the cases, even just to meet these two criteria, is

already a very difficult task. As a result, modifications to Standard might get so cumbersome that our moral intuitions will still treat Reach differently because they will, contrary to the newly-added stipulations, still believe the means in Reach to be more simple, and so more reliable, safe, etc.²⁰ But if the cases are properly equalized in such a way that our intuitions would only be reacting to the morally relevant features of the pairs of cases, rather than to the irrelevant psychological factors related to how complicated and science-fictional the story being told in one or both of the cases is, why on Earth would we think that the moral pressure on me to help is greater in Reach than in modified Standard? I can see no such reason and there don't seem to be any offered by Kamm.

In fact, I believe that if you take at face value Kamm's assertion that her moral intuitions react to Reach in the same way, and for the same reasons, as to a case where the victim is near the agent, that just shows that perhaps her intuitions were never tracking moral differences created by varying distance, but were tracking other factors all along. There are many factors, changes in which normally accompany changes in how close you are to the victim. Some of them are irrelevant to the actual moral pressure on you to aid the victim, e.g. salience and the "what will my neighbors say" kind of reasoning. Some, e.g. how easy it is to help, how likely it is that you will succeed should you try to help, and what the chances are of somebody else's being in an

²⁰I will say more about this and some related phenomena in section V.

equally good position to help, are, it appears, morally relevant. It seems from the preceding discussion that Kamm's intuitions in Reach are influenced by factors in the latter grouping, rather than by considerations of distance. That strongly suggests that in the other cases where she believes that it is distance *per se* that is influencing her intuitions, it is actually either some of the factors in the latter group doing the moral work, or some of the psychological factors in the former group distracting her faculty of moral intuition from the morally relevant aspects of the situation and thus rendering the resulting intuitions untrustworthy.

An additional reason for thinking that Kamm is reacting to factors other than distance in Reach can be found in the paragraph that she devotes to cases of "non-continuous reach". She states in passing that it is possible that "someone may reach to a and all the way back of a to his centered location, but his arm may fail to extend and stop within area a to b, so he has no reach there. It extends and stops again within b to c. Intuitively, there can be a point somewhere between a and b from which he is distant, even while he is near a physically further point between b and c"²¹. Notice the locution "physically further". That is the notion with which it seems we started out at the beginning of this endeavor - we were trying to figure out whether physical distance mattered in determining the level of the moral pressure on us to aid someone. And it may be worthwhile to note that before we got

²¹Kamm, p.187.

to these messy complications, there was never a need to add "physical" to "distance".

Granted, we might need to specify how exactly the distance is to be determined for our purposes. But if we end up with a notion of distance that is such that possibly A is far from B, C is further from A than B, but C is not far from A, I think we are not talking about distance any more. It seems that we are then trying to define a more or less completely new term, say "distance'", in such a way that the corresponding notion will automatically contain within itself certain factors which generally affect the level of moral pressure on us to aid needy strangers: in the Reach case, that would be the ease and certainty with which the assistance could be rendered. Introducing this new notion does nothing to answer the question of whether or not we have PDM, and it begs the question of whether or not we have PD'M - the problem of distance' in morality, where distance' is a factor which varies not only with distance, but also with other factors such as, for instance, the ease with which we could render assistance.

So, it seems wrong for our purposes to define distance in the way that Kamm suggests. But, to stick with this topic for one moment longer, I also want to point out just how problematic it would be for Kamm to substitute distance' for distance. After bringing up Reach, which she most definitely considers to be a case of nearness, she goes on to say that she believes that in a case where an extended part of her can reach the victim, but that part is not efficacious in helping the victim, we are still dealing with a case of nearness and thus with a case where there

is an increased moral pressure to aid (even if the resources that would be needed in order to assist the victim are far). It seems to make sense that there may be great moral pressure on me to aid the victim in Kamm's modification of Reach, call it

Reach': This is just like Reach except my extendible arms cannot actually do anything to help the victim, but there is a machine in Canada that I can activate via remote control that will somehow make the rescue happen.

However, similarly to what I've argued above about Reach, distance (or distance') doesn't appear to have anything to do with the level of moral pressure in question - there is simply no reason to believe that the moral pressure on me in Reach' is greater than in a modification of Standard in which I have that same machine in Canada available to me, provided, of course, that the cases are properly and believably equalized.

But that's not the end of the difficulties. Let's make Reach' slightly less respectable. Consider the **Freak** case -

Freak: I have a medical condition which, from time to time, causes my gall bladder to temporarily extend by 1000 miles; with time, I learned to control the effect and can actually make one of those freak occurrences happen at will and in the direction of my choosing.

Suppose that there is a drowning child on the other side of India, 1000 miles away from me; and suppose, quite reasonably, I think, that gall bladders are non-efficacious in rescuing drowning children (if they are efficacious, the example can be adjusted to compensate). It seems ludicrous to believe that the level of moral pressure on me to save the child, assuming that I have some

means of doing so such as the magical machine in Canada, is in any way affected by my ability to extend my gall bladder. In fact, even if I had had one of those freak gall bladder occurrences while the child was already drowning, and my gall bladder was temporarily within 5 feet of her, I fail to see how I could have, simply in virtue of that fact, picked up any additional moral pressure to save her. It seems a lot more difficult than Kamm thinks, and than it might initially appear, to get her modified definition of distance to make sense, especially in conjunction with the supposition that distance matters in morality.

Section V.

On different kinds of nearness and Kamm's ranking of them.

But now let's move on to a different part of Kamm's analysis of what determines the relevant kinds of nearness. It has to be conceded to Kamm that there are certain cases in which, at least initially, it seems intuitively plausible that nearness adds to the moral pressure to aid. So far we have only discussed the nearness of the agent to the victim. But, as Kamm correctly points out, our intuitive *prima facie* reactions to these cases are extremely similar, if not outright identical, to our reactions to several other types of cases. These cases involve considering not only the relative locations of the agent and the victim, but also,

where applicable, the relative locations of the threat to the victim, e.g. a runaway train, a hired killer, or a vial containing a deadly virus, of the victim's means, i.e. means that belong to the victim and might be efficacious in aiding the victim²², and the agent's means, i.e. means that belong to the agent and might be efficacious in either combating the threat while it is still far, or in aiding the victim directly, or in retrieving or activating some of the victim's efficacious means. This causes Kamm to revise PDM. "My tentative conclusion is that the PDM should be understood as whether we can justify our intuition that we have a greater responsibility to take care of what is going on in the area near us or near our efficacious means, whether this involves needy victims, threats, or means belonging to the victim".²³

Of course, it is my contention that up to this point Kamm has not given us sufficient reason to believe that PDM exists, either in the original or the revised form. And, furthermore, the need for Kamm to revise the formulation of PDM does, I think, throw an additional shadow of doubt on her contention that PDM exists. That's because the need to include kinds of distance other than the distance from the agent to the victim seems to further indicate that perhaps what's doing the moral work in all purported instances of distance's mattering in morality is factors such as how easy it is to assist and how likely the help is to actually

²²The means may be efficacious either by removing the threat while it is still far or by aiding the victim once the threat is already directly acting on the victim.

²³Kamm, p.195.

rescue the victim. And considerations like these provide a very plausible explanation (the most plausible one, in my opinion) of what it is that unifies all of the different kinds of alleged nearness into a single grouping that might potentially be relevant to moral theory. Also, with the modified definition of distance (the one dealing with extended parts) as well as the new relevant kinds of distance, it becomes utterly unclear how a pair of cases can ever be equalized for everything but distance, especially for factors such as the costliness of the attempt to help and the certainty of its success.

But, be that as it may, this revision of PDM is interesting both for its own sake and because I believe that looking at Kamm's subsequent discussion in slightly more detail might help us figure out why there is probably no PDM even though it might initially seem otherwise. On p.195 of her article, Kamm tries to rank the various distance factors by their contribution to the moral pressure on the agent to render assistance. She admits that the difference may be slight, but, in at least some of the cases, she is certain that there is one. In particular, she is certain that, all things not related to distance being equal (where by 'distance' I mean any of the kinds of distance mentioned above), the importance of the distance between the agent and the victim outweighs the importance of the distance from the agent to the threat to the victim; as well as that the importance of the distance from the agent to a threat outweighs the importance of the distance from the agent to a victim's efficacious means. Kamm also maintains that the importance of the distance from the agent

to a victim's efficacious means ranks at about the same level as the importance of the distance from the agent's efficacious means to the victim. The latter claim seems to be too complicated, and the intuitions regarding it to be too murky, for it to be of much interest to us. But I think it may be worth our while to consider more carefully the claim that, all things not related to distance being equal, the importance of the distance between the agent and the victim outweighs the importance of the distance from the agent to the threat to the victim.

This is what Kamm has to say about it: "In one near pond, there is a threat to a faraway victim and, in another near pond, there is a different (unseen)²⁴ victim. To which pond should the agent go if one life is at stake in either choice? Perhaps to the pond with the victim"²⁵. But what is that intuition based on? To try to answer that question, let me change the example a little bit. The reason for doing this is that it's hard to come up with a realistic and not too complicated story about a threat in a nearby pond to a victim far away. And, as I've remarked before, the more science-fictional and complicated the example, the harder it is to make any sense of our intuitions about it. So, let's consider

Threat vs. Victim: Suppose that there is an unseen and unheard child drowning in a pond near me. Also, at the same distance from me as the drowning child, there are railroad tracks and a switch. If I throw the switch, one mile further

²⁴This is inserted by Kamm in order to equalize the two sides of the example for salience.

²⁵Kamm, p.195.

down the track the train will be diverted to a different track. Suppose that as the train passes by, I simultaneously find out about the drowning child and about another child of the same age, gender, religion, nationality, etc., that is lying incapacitated on the track. That incapacitated child is located one mile beyond the fork and on the same track that the train, which I find out at the same time to be a runaway train, will be on if I don't throw the switch.

We also need a large number of additional equalizing conditions. The children are both strangers to me and I am under no contractual obligations of any kind to them or their families. Neither of them is responsible for his or her plight (or, alternatively, the two are equally responsible). Let's also assume, to make our lives a little easier, that in this case the death by drowning and the death by runaway train are equally quick and equally painful. Also, let's assume that no one else can get either to the train or to either of the children in time to save them and that I cannot prevent both children from dying; but let's suppose that whether I choose to save the drowning child or to throw the switch that would divert the train, I have a 100% chance of succeeding in saving the child I've chosen to save. And whichever child I do not help is 100% certain to die. Furthermore, let us suppose that I am fully aware of all of these circumstances.

So, what are we to make of this example? Let's suppose, for the moment, that we are tempted to agree with Kamm that our intuitions point to its being the case that there is greater moral pressure on me to save the drowning child? What could be making us think that? I can suggest a number of factors, but none of them can actually affect the moral pressure on me to save the

drowning child, even though they can affect the psychological pressure on me to do so. All of these reasons have to do with the problem that I have brought up earlier, in section III - that many of the factors that have been equalized for, were equalized by stipulation only. The circumstances of the case suggest that these stipulations might be false and, in fact, are highly likely to be false, because based on our experience they are usually false when such circumstances obtain. For instance, a factor of major importance is that of how probable it is that my help will actually succeed in saving the victim in question. It is stipulated that the probability is 100% in either case. But perhaps I feel that if I run to the pond and physically pull the child out with my own two hands, that is by far more reliable, especially if I know how to swim, then pulling the switch and hoping for the best²⁶. Of course, some people may feel the opposite way, but that, I should think, reduces the probability of their being inclined to agree with Kamm's conclusion. We might

²⁶This phenomenon is much more obvious, perhaps even blatantly so, when we start talking about putting some money in a mailbox somewhere and being guaranteed by stipulation that the moment we do so, some organization like Oxfam will find out about it and save a particular starving child (Kamm flirts with similar examples on occasion). Our knowledge of past experiences and of some unpleasant facts about life, makes us feel highly uncertain that we can trust any aid organization to be either that honest or that efficient (where by efficiency I not only mean the ability to find out about my "deposit" the moment I make the money available and the ability to procure, deliver and distribute the necessary amounts of food and medicine in time, but also, for instance, the ability to protect the shipments from raiders such as the Somali warlords or local bureaucrats). Therefore, if we have a more commonsensically plausible and trustworthy opportunity to save some child's life, a mere stipulation that the chances of success are equal is not likely to be enough to truly equalize the cases for our faculties of moral intuition.

want to equalize the two options in our example even further by making the activation of a machine that is certain to rescue the drowning child the only way that the child can be saved. But here we would have to be careful. This might actually tip the balance of our moral intuitions in favor of saving the child on the railroad tracks, as based on our experience railroad switches are very likely to be much more reliable than science-fictional devices such as the child-scooping machines that figure in several of Kamm's examples, some of which we will discuss in just a few pages.

Another important issue is the certainty of the death of the child on the railroad tracks if I don't throw the switch. It is stipulated that the probability of death in that eventuality is 100%. But do we really believe it when we are evaluating what our intuitions tell us about the case? It is very difficult in general, I think, for human beings to comprehend absolute certainties, especially in complicated situations, and especially when the outcome in question is as mind-bogglingly horrible as a helpless child's being torn to pieces by a runaway train. All stipulations aside, there is always irrational hope that someone else might stumble upon the child in the time that it takes for the train to get there, or, when all else fails, we can always hope for an "act of God" to stop the train. I think in most cases which are equalized for the basic physical factors and in which we are greatly tempted to go for the near victim rather than a near threat, it is these kinds of hope that are the cause. It is, however, I believe, possible to equalize the two options in the

example for that factor as well. We simply build into our example the clause that if I assist neither child they will die at the same time - that should even equalize for such an irrational factor as the hope for an act of God²⁷.

I am sure that there are other psychological issues, related to the stipulations that are needed in order to equalize the options, that create similar difficulties and that I have not considered. But I think the ones I have brought up are already more than sufficient in order to make my point. Once the factual equalization of the options is carried out and a number of psychological problems are discovered and equalized for, we end up with an extremely complicated example that probably borders on, or even crosses over into, the realm of science fiction. At that point, if it still seems that our intuitions tell us to go for the victim and not for the threat, it seems that the only explanation is that the result as a whole is too complicated for our faculties of moral intuition to grasp, because there just doesn't appear to be any reason why the moral pressure could actually be different. I think there are only two options - either the example is still "graspable" in the relevant sense, in which case we can't avoid the conclusion that the moral pressure pulling me in one direction is equal to the one pulling me in the opposite direction, or we have a plausible error theory for why our moral intuitions are, in

²⁷In "Whatever the Consequences" (1966), Jonathan Bennett aptly describes the tendency to think of outcomes that are stipulated to be certain as increasingly uncertain in direct proportion to the amount of time that remains until the outcomes take place as "perhaps encouraged by an eccentric view of God as powerful but *slow*".

a manner of speaking, lying to us.

There appears to be no justification for Kamm's conclusion that in a properly equalized case there is greater moral pressure to save a nearby victim than to eliminate a nearby threat to a faraway victim, even though there might, conceivably, be psychological grounds for saying that it will seem to many of us that the moral pressure to save a nearby victim is greater. The main reason for this is that, as I've argued a bit earlier, even if all the factors other than distance are equalized for, the illusion of those or even, perhaps, other differences can still remain to skew our intuitions and lead them astray. So, the examples really don't tell us what Kamm argues they do.

The same problem plagues Kamm's examples which are supposed to illustrate the very existence of PDM. If proper moral intuitions were the outputs of a computer algorithm, the facts of the example being the inputs, then, of course, no equalization beyond the factual one could possibly be required in order to determine whether or not distance matters in a particular case. But since the examples have to be subjected to the human test, a "conceptual equalization" is also required: the example has to make our faculties of moral intuition believe that the facts are truly equalized. And that is an extremely hard, if not impossible, thing to do considering the level of complexity that is required to factually equalize the cases, in conjunction with the fact that a change in distance is usually accompanied by other changes that definitely do affect the moral pressure on us to aid - which is the reason why it initially seems so plausible to us that distance

matters in morality in the first place. We can see this illustrated yet again by taking a look at one of the examples Kamm offers at the beginning of her article in order to demonstrate that intuitively distance does matter. Let's consider Kamm's Near/Far case -

Near/Far: "I learn that in a distant part of a foreign country that I am visiting, a child is drowning and someone is near him. Either one of us could as successfully help by depositing \$500 in a device that will trigger a machine that will scoop the child out. Who has a stronger obligation to help?"²⁸

I am choosing this example because I think it is a very good test case for determining what our intuitions are, that also lets us look at the problem from a slightly different angle because only one victim is involved.

I may be willing to concede to Kamm that *prima facie* it may seem to some that the moral pressure on the person who is closer to the child (let's call him Mike) to save her is greater than the moral pressure on me to do so. But this example is not properly equalized, so let me modify it a little. Suppose, first of all, that not only I, but Mike as well, cannot personally participate in the rescue or even observe it. Also, let us suppose that there is only one child-scooping machine and it is located at an equal distance from me and Mike. We can both get to it in an equal amount of time and activate it so that it can rescue the child before it is too late. The reason why I think there should be only one child-scooping machine is similar to the one I offered in

²⁸Kamm, p.179.

footnote 26. If there is one machine near me (and thus quite distant from the drowning child) and one near Mike (and thus very near the child), it has to be purely a matter of stipulation that the insertion of the money into one of the machines is equally likely to lead to that machine's rescuing the child in time, regardless of which machine it is. I think it is enough of a strain on our faculties of moral intuition to be dealing with child-scooping devices in the first place. To also endow such machines with the ability to almost instantaneous transport themselves across vast distances, in a way which we are to believe is as certain to be speedy and successful from very far away as from just a few feet away, is, perhaps, too much to ask.

So now, if we also add to our modifications of Near/Far that both I and Mike found out about the plight of the child in the same way - in particular, that neither of us can either see or hear her - and thus we eliminate the possibility of salience making it just seem that there is greater moral pressure on Mike to activate the machine, why on Earth should anyone even be tempted to think that either one of us is under greater moral pressure than the other to aid the child²⁹? I believe that the Near/Far case is rather unique in that even though it seems initially that distance matters in this case, it is apparently possible to come up with an equalized modification of it without making the example so complex that upon due reflection we still

²⁹Of course, some further conditions have to be introduced, such as that the child is a stranger to both me and Mike, that neither of us has a duty (e.g., by contract) to help the child, etc.

can't trust our moral intuitions³⁰. Thus it looks as if in this case we can actually observe happening what I am arguing is going on in every one of Kamm's examples.

Section VI.

A few more of Kamm's arguments for PDM addressed.

Having, I believe, produced a satisfactory discussion of Kamm's examples which are supposed to provide a direct argument for the existence of PDM, I would like to address in this section a second category of examples. They form the basis for what I will call "indirect" arguments for the existence of PDM. In these examples, Kamm argues, we find it difficult, if not impossible, to explain our moral intuitions unless we assume that distance does, indeed, matter in morality. I will try to show that it is not necessary to make that assumption in order to explain what's going on in these cases.

All three of the examples that I am going to discuss in this section are borrowed by Kamm from Peter Unger. First, let's consider

³⁰In this case there are two potential agents and only one victim; but obviously a case in which I can by an immediate phone call save either a nearby child or a far away child (both of which are in an identical predicament), but not both of them, can be analyzed in a completely analogous fashion to reach an analogous conclusion.

Oxfam: "Someone who has already given a lot to Oxfam feels morally free to refuse to respond to another request for life-saving aid. The same person, however, does not feel that simply because he has already given a lot to Oxfam, he may refuse services to a person he meets on the road who needs life-saving help"³¹.

Kamm goes on to add that "since Unger denies the significance of distance, he concludes that one cannot refuse the additional Oxfam request for aid to distant lands any more than one can refuse the person on the road"³². But that conclusion appears to be blatantly counterintuitive, prompting Kamm to give her own explanation of what's going on in this case. Her evaluation of the example rests on the assumption that as compared to the moral pressure to aid someone who you find on the side of the road, there is in the first place a much lesser, if any, moral pressure on one to give money to Oxfam, or any other organization that claims to aid starving people at a great distance. Whatever moral pressure there may have been to begin with, has been lifted from the agent in the example by virtue of his having given a lot to Oxfam in the past. The moral pressure on the agent to aid someone near, e.g. someone he notices on the side of the road as he is driving by, is much stronger and of a different nature. Thus, the agent's prior history with Oxfam cannot have any bearing on the moral pressure on him to aid the near individual.

Kamm's account seems perfectly reasonable provided that distance does, indeed, matter in morality. However, that account is completely unnecessary in this case in order to explain what is

³¹Kamm, p.185.

³²Kamm, p.185.

going on and show that Unger's conclusion is utterly unjustified. The reason is the same oversight for which Kamm criticizes Unger, Singer and others at the beginning of her article - the two alternative situations in the example are atrociously unequalized. They vary with regard to a large number of factors, with many of those, even considered individually, being sufficient to show that Unger's conclusion is either unwarranted or just plain wrong. Among those factors are the availability of other sources of aid, the certainty of the help's reaching the victim, whether or not the victim belongs to the same community as the agent, and whether or not the victim's need is the result of some much wider social or political problems, as well as probably many others. There is also the issue of salience of need, which, even though it does not affect the moral pressure to aid, may have made it seem to us that the moral pressure to help a person you see on the side of the road is greater than the moral pressure to give to Oxfam even if that were not, in fact, the case. But that is irrelevant to showing that Unger is wrong in saying that the moral pressure on the agent in the two situations is actually equal. At any rate, it is, I think, obvious by now that it is not necessary to assume anything about the role of distance in morality in order to avoid Unger's counterintuitive conclusion.

The second example deals with so-called *futility thinking*, which in many cases seems to keep people from aiding those in need. Futility thinking is "the tendency to not help anyone if one cannot help a significant portion of those who need help...Unger says that the case in which I can definitely save the

only person near me who is drowning is not a case in which I can take care of "the whole problem", any more than I do when I save a few of those doomed by starvation in distant Africa"³³. Kamm suggests that the reason futility thinking doesn't seem to affect us in the case of a single person drowning nearby, is because we are under much stronger moral pressure to help a person who is near than to help the people who might be drowning very far away from us at the same time, and so "when I take care of the only person near me who needs aid, I think I have completely dealt with a problem"³⁴.

Unger, on the other hand, believes that the solution lies in an appeal to salience, i.e. that in some cases, like the case of one drowning man nearby, salience helps overcome futility thinking. However, I even fail to see this as a relevant problem. Futility thinking, just as salience, is a purely psychological phenomenon. Unger, at least according to Kamm's account of his views, believes that salience can actually affect the moral pressure on us to aid and not just the psychological pressure to do so. Therefore, he might similarly believe that futility thinking can affect the moral pressure to aid as well, which is probably what causes him to raise this puzzle in the first place. Kamm, however, has successfully argued that salience by itself cannot have any effect on the moral pressure to aid. And, in light of that argument, it is unclear how futility thinking can be any different in that respect. So, for Kamm this puzzle has to be

³³Kamm, p.185.

³⁴Kamm, p.186.

a non-problem, as it is for anyone who agrees with her argument regarding salience.

As far as Unger is concerned, however, I believe he is partially right. I think that while futility thinking does not actually affect the moral pressure on us to aid, it can make it seem that the moral pressure is smaller than it actually is. But in cases where the victim's need is very salient, salience can play the opposite role by counterbalancing futility thinking to restore the level of moral pressure felt to about the level at which the moral pressure actually is. In any case, there is no need for the alleged moral relevance of distance to play any role in the discussion of this Unger puzzle.

The third example involves a pair of cases in which an agent comes to the rescue of a victim or victims using a third party's means:

Yacht: "An agent takes a boat that belongs to another to help someone near him. The rescue will result in a million dollars of damage to the boat for which he cannot compensate."³⁵

and

Account: "A delivery boy to the office of a rich person can do a computer transfer of funds from his account to Unicef's"³⁶.

To make things a little clearer and somewhat more equalized, let's add that in Yacht, as well as in Account, lives of strangers (both to the agents and to the people whose resources are used) are at

³⁵Kamm, p.191.

³⁶Kamm, p.192.

stake, and, in fact, that the number of lives in the balance is equal. Also, let us suppose that the amount of the transfer in Account is one million dollars. Unger concludes that the two cases are morally equivalent, i.e. that the actions of the agent in Yacht are morally no better or worse than the actions of the agent in Account. However, as Kamm points out, that just doesn't seem to be the case. Our intuitions appear to be clearly telling us that the actions of the agent in Yacht are morally permissible, while those of the agent in Account are not, or, at the very least, if both actions are somewhat wrong, that the one in Yacht is much less so.

Kamm's explanation quite predictably relies on her views about distance. She points out that in Yacht, even though the owner of the yacht is not near the victims, the owner's means that are efficacious in aiding the victims, i.e. the yacht, are near the victims. In Account, on the other hand, neither the owner of the money nor any efficacious means (including the money)³⁷ belonging to him are near the victims; and, for that matter, the agent in Account is not near the victims either. Therefore, Kamm concludes that "in Yacht, when I take the yacht to help someone, I am carrying out someone else's duty for him, based on either his nearness or the nearness of his property. In Account, I cannot

³⁷For simplicity's sake, let's forego the sticky discussion of whether or not someone's money is near if an ATM from which the money could be withdrawn is near, which could make it the case that a very large proportion of the money in the world is near all kinds of victims, threats and efficacious means; creating, on Kamm's view of the role of distance in morality, an unbearable burden of additional moral pressure to aid numerous physically distant strangers.

say I am enforcing an owner's obligation to a victim if nearness is a ground for obligation."³⁸ Thus, she believes, the actions of the agent in Yacht, even if not completely Kosher, are on a much more solid moral footing than those of the agent in Account.

Let's assume for the duration of this discussion that it makes perfect sense to talk about "carrying out someone else's duty for him". That makes Kamm's suggestion a candidate for a good explanation of our intuitions regarding Yacht and Account. So the question is: can we explain what's going on in these two cases without relying on the assumption that distance matters in morality? If we can, we deprive Kamm of yet another indirect argument for the existence of PDM.

First, I want to suggest a reason why it might seem to us that Yacht and Account are not morally equivalent even if, in fact, they are. Whenever an agent "volunteers" someone else's resources to help a third party, we are usually more inclined to be understanding or forgiving if the agent, so to speak, brings something to the rescue as well. In Yacht, the agent not only borrows the boat, but presumably conducts the actual rescue which, for all we know, might be very difficult or even dangerous. The agent in Account, on the other hand, contributes nothing besides someone else's money³⁹ - the actual assistance is rendered by Unicef. I don't see how this could affect the actual moral permissibility of the two actions in question, but it is

³⁸Kamm, p.192.

³⁹Unless we are willing to take into consideration the dubious contribution of risking being arrested for his actions.

definitely a psychological factor that can, and I believe does, skew our moral intuitions somewhat in favor of the agent in Yacht and against the agent in Account.

But let's assume that not only does it seem that the actions of the agent in Yacht are less wrong than the actions of the agent in Account, but that they actually are. In that case the explanation comes from Kamm herself. She points out that Account and Yacht are not equalized for a wide variety of relevant factors. Mostly these include all the usual suspects that have come up many times before, e.g. the availability of other sources of aid, the probability of success, etc., etc., and so I will not discuss them in detail here. I do want to add that in this case even salience might indirectly play a moral and not just a psychological role. That's because we are missing a story on what exactly is relevant in determining the moral permissibility of "carrying out someone else's duty for him". For all we know, not only might the level of the moral pressure on the owner of the resource to make that resource available be a relevant factor, but perhaps whether or not a reasonable person in the owner's place would want to make the resource available is morally relevant as well - after all, the resource does belong to him. And for determining the latter, salience may, indeed, be of some importance.

Kamm believes that even if the cases are properly equalized for everything but distance, the actions of the agent in Account will only become as morally permissible as those of the agent in Yacht, if, in addition, the victims in Account are close to either

the agent or the money or, preferably, both. But if we didn't believe Kamm's argument for a similar conclusion in Near/Far and other cases discussed in this paper so far, why should we accept it here? I believe, for the same reasons as in the earlier cases, that there are no grounds whatsoever to suspect that if Yacht and Account are properly equalized for everything but distance, any moral difference between the two cases will remain. Thus if one doesn't find Kamm's direct arguments for the role of distance in morality convincing, one cannot find her indirect arguments, i.e. the arguments from the alleged indispensability of the role of distance in morality for explaining certain moral facts, even the least bit compelling.

Section VII.

Some final thoughts on the subject.

Previous sections seem to conclusively show that Kamm has failed to provide us with any examples which directly or indirectly establish that distance matters in morality. This, of course, as I stated in section I, does not prove that distance does not matter in morality; for if even one uncontroversial example can be found where distance does matter, the thesis that distance *per se* never matters in morality would be wrong, which would perhaps imply that distance only matters when one or several other factors are present. It is clearly not an option to simply

go through all the possible examples one by one and show that distance matters in none of them. However, the burden of proof is squarely back on Kamm and anyone else who supports her view. And, in light of the discussion above, it is virtually unfathomable that there could be an example purporting to establish that distance matters in morality, which is not instantly defeasible, let alone one that is uncontroversial and conclusive.

There is also an additional reason for believing that distance does not matter in morality. It is inevitable that there be some truths which require no explanation and, in fact, are incapable of having one. After all, explanations do have to end somewhere, and every discipline needs a starting point. Arguably, one such axiomatic fact is that pain and suffering, as well as death, of sentient beings are morally significant. But it defies belief that it could be a brute inexplicable fact about morality that distance *per se*, of any of the kinds discussed in sections IV and V of this paper, affects the level of moral pressure on a person to help a stranger in need. Thus it seems that if there is a PDM, there better be a plausible solution to it - some sort of an explanation, derived from more basic, perhaps even brute, moral facts, of how and why distance in itself can possibly make a moral difference.

Even if there were more or less conclusive examples that appeared to show that distance matters in morality, one might have found the prospect of acknowledging this as a brute fact so unpalatable as to consider rejecting the underlying moral intuitions should the search for an explanation turn out to be

fruitless. But in the absence of conclusive examples, it seems to go without saying that anyone who believes that distance plays a role in morality better have one hell of a strong theoretical argument to back up that claim. That makes it doubly important, not only for the sake of solving the PDM, if it exists, but also to establish the very possibility of its existence, to investigate the question of whether or not a role for distance in morality can be justified.

Since Kamm believes that we are faced with PDM, she makes an effort at the end of her paper to give an explanation for why distance might matter in morality. But she herself finds that her best attempts leave a lot to be desired. First, she says that perhaps, in order to keep our moral duties manageable, there exists "the need to set limits to our duty to aid strangers; being responsible in accord with proximity is a way to set these limits"⁴⁰. But she immediately admits that this way of limiting responsibility is too arbitrary to be justifiable, as other ways of imposing such limits are certainly available - for one thing, it is also the case that "if Asian Indians were [only] responsible for helping those in East Hampton, New York, they would have less strenuous obligations"⁴¹. Also, even setting aside the arbitrariness objection for the moment, this explanation of the role of distance in morality would imply that "proximity matters as a heuristic device that correlates with morally significant

⁴⁰Kamm, p.198.

⁴¹Kamm, p.198.

factors, though it itself is not morally significant"⁴². Thus there would not be a direct role for distance in morality after all.

Kamm's second attempt at an explanation is only slightly more plausible. The suggestion is that "it is commonly thought that one has a moral prerogative to give greater weight to one's own interests rather than giving equal weight to oneself and to others...but possibly, if one takes advantage of the option of giving greater weight to oneself and what one cares about, there is an *associated duty, the flip side of the prerogative*, generated from the perspective on life which leads to the prerogative, to take care of what is associated with oneself: for example, the area near one"⁴³. But, as Kamm herself points out, giving greater weight to one's own interests is at most "a permissible option, not a duty"⁴⁴. So, accepting this explanation of the role of distance in morality would lead us to the wildly implausible conclusion that whether or not distance matters in morality is agent-relative, i.e. it depends on whether or not the agent chooses to give greater weight to his own interests. So, in effect, each agent would get to choose whether or not distance would matter in morality *for him*, which seems like an extremely unattractive feature for an ethical theory to have.

As a result, I think it is clear that we do not at this point have even a hint at a satisfactory explanation of the

⁴²Kamm, p.198.

⁴³Kamm, p.200.

⁴⁴Kamm, p.200.

purported role of distance in morality. Thus, we currently have no reason of any kind to believe that distance actually matters in morality or that we are, in fact, faced with PDM. On the contrary, we have compelling reasons to think that neither is the case.

So, why would someone be initially tempted to believe that distance matters in morality in the first place? It seems obvious, based on the cases which we have looked at in the course of this paper, that distance is very strongly correlated with a large number of factors that do make a moral difference, e.g. the cost of aiding, the probability of success in aiding, and also, quite often, the availability of alternative sources of aid. It is very important to note, however, that the correlation is imperfect - for any one of these factors, usually correlated with distance, it is possible to greatly vary distance without varying that factor. Thus it is not the case that distance can somehow be morally analyzed into a combination of some of these factors, which would, after all, have meant that distance matters in morality. Distance is also somewhat correlated with certain psychological factors, e.g. salience, which, even though they do not make a moral difference, do tend to influence our moral intuitions, not to mention our actions. So, perhaps Kamm hit the nail squarely on the head while discrediting one of her own suggestions - to quote the statement once again, "proximity matters as a heuristic device that correlates with morally significant factors, though it itself is not morally

significant"⁴⁵.

In fact, as we have seen during our attempts to equalize the cases we've examined for everything but distance, such equalizing is an extremely difficult task and the outcome usually ends up being complicated and at least slightly science-fictional. In real life, where all factors other than distance are extremely rarely, probably even never, equalized, it is indeed a very useful rule of thumb that the moral pressure on an agent to help a stranger usually increases with an increase in proximity. But it is nothing more than a rule of thumb, made useful only by the fact that a change in distance is usually, but not always and certainly not necessarily, accompanied by changes in morally relevant factors. It looks like we now have a very plausible explanation both of why distance does not matter in morality and of why it is quite reasonable that we would, *prima facie*, assume the contrary⁴⁶.

⁴⁵Kamm, p.198.

⁴⁶In the example on the first page of this paper, for instance, in which it seems clear that I ought to call an ambulance to aid the injured motorist rather than phone an international relief organization, I think it is by now clear, in light of my discussion in the preceding few sections, that should we be able to equalize the options for everything but distance without treading too far in the realm of science fiction, we should expect the moral pressure on me to help the motorist to be equal to the moral pressure on me to save the starving child. I do not, at this point, wish to hazard a guess or an argument on the very interesting, though treacherous, topic of which of the factors listed in section I, that are not equalized for in the original case, make a more significant moral difference, especially in light of objections brought by Unger, Singer, and some others against the moral relevance of some of these factors. This topic I will leave for a different discussion. In the meantime, the fact that we have every reason to expect equal moral pressure in an equalized case, should, in conjunction with there being no theoretical explanation in the offing for how distance could possibly matter in morality, be more than sufficient to conclude that there is no PDM.

Finally, there is one last issue that I would like to address. There seems to be a purely theoretical reason why Kamm wants it to turn out that distance matters in morality. In several places over the course of her article she expresses concern about a particular type of cases - an example of which is a case where you have a fixed amount of resources which you can use either to save a large number of people who are very far away, e.g. 100 starving strangers in Africa, or to save one stranger nearby, e.g. a homeless person "living" on your street. Kamm is concerned that if distance doesn't matter in morality, it will turn out to be the case that you are under much greater moral pressure, or, even worse, perhaps morally required, to aid the 100 people far away. Kamm quite rightly finds that conclusion troubling, both theoretically and practically, as well as somewhat counterintuitive. The conclusion is especially disconcerting because unlike most, if not all, of the other examples that have been discussed, this is a dilemma that does occur quite often in real life.

I have two suggestions in response to this difficulty. First, if the alternatives are properly equalized for all factors other than distance and the number of needy strangers, the conclusion might begin to seem somewhat less counterintuitive and disturbing.⁴⁷ Second, even if the problem still remains in an

⁴⁷Of course, we also need, as usual, to keep in mind that the alternatives should be equalized not only factually, but conceptually as well, i.e. in such a way as to make the putative equalization palatable to our moral-intuition-generating faculties. Otherwise, the moral conclusion under discussion might very well be correct, but still seem counterintuitive and disturbing.

equalized example, there may be a solution that has nothing to do with distance. It is possible that what we should do in order to rid ourselves of the unpleasant consequences of the purported moral fact that we should help the much greater number of distant strangers, is not to assume that distance does, in fact, matter in morality, but to argue that the number of the strangers involved does not. But that is the topic of the next chapter.

-Chapter 2-

The Problem of Numbers

THE PROBLEM OF NUMBERS.

Section I.

Introduction to the problem.

Ask any reasonable person whether or not it is morally permissible to kill a randomly chosen innocent healthy man in order to use his organs for transplants that will save five others, and you will hear that it most certainly is not. This is the classical **Transplant** case. The example is usually brought up to elicit strong intuitions which indicate that it is not morally permissible to kill in order to bring about what would otherwise be considered a preferable outcome. But there is a very important question concealed here that usually goes completely unnoticed - is the state of affairs in which five people's lives are spared really a better one than, and morally preferable to, the state of affairs in which one person survives a deadly threat, but the five do not?

In Transplant, the number of people who are in urgent need of organs is stipulated to be five partly in order to evoke our intuitions more clearly. After all, if it is morally wrong to kill one innocent person to prevent five others from dying, surely it is wrong to do so just to prevent one person from dying. But, on the other hand, it is usually assumed without argument that the converse is not necessarily true, i.e. that it does not have to be the case that something that is morally impermissible if done to save just one life is morally impermissible if it is done to save

many more lives. This assumption seems to be strongly supported by the fact that virtually everyone who believes that we are morally permitted to divert a deadly threat from five innocent people to one equally innocent person, would not feel the same way if instead he were considering the possibility of diverting a deadly threat from just one innocent person to an equally innocent person. And no one in their right mind would suggest that, all other things being equal, it is morally permissible to divert a deadly threat from one person to five people.

But should we accept these intuitions? There is great room for discomfort if we do. It seems that in allowing the number of people involved to count as a moral factor, and a factor of great force at that, we are apparently ignoring the people in question as individuals, especially the people that form the smaller group. Also, as Taurek suggests in "Should the Numbers Count?", by allowing the numbers to matter in morality we are treating human beings as objects: each one has a certain value to us, and in deciding which group of people should be saved we follow the same procedure we would follow if we had to decide which group of objects to keep - we add up the individual values of each object in a group and keep the group for which the resulting sum is greater¹. In addition, the standard view that numbers should count turns out, upon closer examination, to be much harder to justify than it might appear at first glance. These issues cause some people, John Taurek and myself included, enough discomfort to

¹In cases where people are involved, we can just multiply the number of people in a group by a unit.

seriously entertain the suggestion that perhaps numbers should not play a role in moral considerations or, at the very least, that they should not play the role they are usually taken to play.

In this paper I will investigate the question of whether or not numbers matter in morality. I will begin by summarizing Taurek's arguments for the conclusion that they do not. I will then proceed to address several major objections that have been raised against Taurek's arguments and position. I will also discuss the role that the notion of fairness plays in Taurek's account. Finally, I will raise a few difficult problems for Taurek's conclusion, to which I do not believe that I have an entirely satisfactory response. I am sympathetic to Taurek's view, but, due to the complexity of the problem and the numerous difficulties that his account clearly faces, it will not be the goal of this paper to demonstrate the correctness of Taurek's conclusions. The most I can hope for is to establish his position as a viable alternative to the standard view of the role of numbers in morality - an alternative that addresses some of the concerns created by the standard view. This will not be an easy task, but it is a task that I believe to be worthwhile.

Section II.

Taurek's Arguments.

Subsection 1: A Few Preliminary Thoughts.

In this section I will lay out Taurek's arguments for the conclusion that the relative sizes of two non-overlapping groups of people do not, all other things being equal, play a role in determining which of the two groups ought to be saved from a fatal threat. It is extremely important to realize that in "Should the Numbers Count?" Taurek has two independent lines of argument for his main conclusion. Many philosophers seem to miss this point and thus appear to falsely believe that undermining just one of Taurek's lines of argument is sufficient to uphold the standard view of the role of numbers in morality. In subsection 2, I will state what I subsequently call Taurek's First Argument. In subsection 3 I will present Taurek's Second Argument. But first - several important preliminary points.

Taurek focuses on a group of cases where an agent can prevent some harms from befalling one, but not both, of two non-overlapping groups of people. In particular, Taurek considers several cases where the agent has to decide between sparing one person a certain harm and sparing five different people a similar harm. We are to assume in all of these hypothetical situations that there are no special considerations involved that should influence the moral predicament that the agent is in. For instance, it is not the case that the agent is somehow responsible for the fact that some or all of the people involved are under the

threat of suffering the harm in question. It is also not the case that the agent has any contractual or quasi-contractual obligations towards some of the people involved to protect them, either from harm in general or from the immediate threat they are faced with in particular.

The category of situations involving quasi-contractual, i.e. implied rather than explicitly contracted, obligations should definitely cover cases which feature such phenomena as the doctor-patient relationship and the parent-child relationship, as well as perhaps a few other familial relationships, probably including the one between spouses. Some people might also be inclined to include friendship in this list, but let's assume for the purposes of the discussion that is to follow that friendship does not generate any quasi-contractual obligations.

Taurek also wants his examples not to be influenced by considerations which might make one or more of the people involved more or less "valuable" than the others, e.g. the fact that one (or more) of the people involved is (are) on the verge of discovering the cure for some horrible disease, or the fact that some of them are "driveling old people" or "idiot infants, loved by no one". The list of these special considerations is open-ended, but that shouldn't be a problem for us, as from the partial list above it is already clear what we are after. In order to examine the role of numbers in morality we want to only look at cases which will be untainted by other considerations that might turn out to be morally relevant and thus tip the moral balance in favor of one of the groups for reasons unrelated to the relative

sizes of the groups.

Subsection 2: Taurek's First Argument.

We are now ready to take a look at some specific examples. First let's consider

First Person: The agent (let's call him David) has a supply of some life-saving drug. Unless he takes all of it, he will die. However, if he takes the drug, five other people, all strangers to him, each of which would require just one fifth of the drug in order to survive, will die.

Surely, Taurek says, it is morally permissible for David to keep all of the drug for himself. After all, the drug belongs to him and he is, by assumption, under no contractual or quasi-contractual obligation to relinquish the drug to the other five people. Those five have no right to the drug and so, should they not receive it, cannot complain that their rights have been violated². Also, Taurek asks, how could it not be morally permissible for David to keep the drug for himself? Some might be tempted to say that by giving up the drug to the five strangers, and consequently dying, David would prevent a much worse outcome from occurring. But, Taurek responds, that claim is unjustified.

²As I will discuss later on, while this claim itself is unproblematic, it might not have the consequences that Taurek is looking for. Certainly, if David is under no contractual or quasi-contractual obligation to relinquish something that belongs to him, he has the right to keep it for himself and, most likely, he would violate no one's rights in the process. However, that may be compatible with its being the case that it would be morally wrong for David to hold on to his possession. On this view, Taurek lacks a conclusive argument for his judgment regarding the First Person case. However, he can simply resort to basing it on a moral intuition, which, after all, is shared by the majority of humankind; this will be discussed in a lot more detail in section IV.

While it is certainly better for each of the five strangers that they should get the drug and survive, it is better for David if he keeps the drug for himself and survives. In what sense could it be just plain worse, "not necessarily for anyone in particular, or relative to anyone's particular ends, but just a worse thing in itself"³ that the five strangers rather than David die?⁴

Taurek also claims that if outcome A is better than outcome B, but you express a preference for B over A, that appears to have bad implications for your moral character, as "it is a moral shortcoming not to prefer what is admittedly in itself a better thing to what is in itself a worse thing"⁵. But it just doesn't seem right to Taurek to say that David is morally deficient for not preferring his own death to the deaths of the five strangers.

Now let's consider the **Friend** case.

Friend: I have a dose of the drug: it belongs to me and I am under no contractual or quasi-contractual obligation to give it to anyone. I can either give the entire dose to David, someone who I "know and like", which will save his life, or give a fifth of the dose to each of five strangers, thus saving all five of them. I cannot save both David and the five strangers.

Taurek argues that it is morally permissible for me to save

³John Taurek "Should the Numbers Count?", ETHICS Problems and Principles, ed. by John Martin Fischer and Mark Ravizza, Harcourt Brace Jovanovich College Publishers 1992, p.220; All future quotations from Taurek are from this article.

⁴This is an extremely important and underrated question, which also lies at the heart of Taurek's Second Argument.

⁵Taurek, p.221.

David.⁶ First, he points out that, as in the previous case, nobody's rights can be violated by my giving the drug to David. Second, Taurek argues, if it is not morally impermissible for David to keep the drug for himself in First Person, it cannot be the case that it is not permissible for me to give David the drug in Friend. Here Taurek relies on a general principle, which I will dub the

Non-Relativity Principle: "if it would be morally permissible for B to choose to spare himself a certain loss, H, instead of sparing another person, C, a loss, H', in a situation where he cannot spare C and himself as well, then it must be permissible for someone else, not under any relevant obligations to the contrary, to take B's perspective, that is, to choose to secure the outcome most favorable to B instead of the outcome most favorable to C, if he cannot secure what would be best for each".⁷

In Taurek's opinion, if there are grounds for its being the case that I am morally required to give the drug to the five rather than to David in Friend, those grounds would also be sufficient to generate a moral requirement that David give away the drug in First Person. Thus, since Taurek is assuming that it is morally permissible for David to keep the drug for himself in First Person, it appears to follow that it is morally permissible for me to give David the drug in Friend. As Taurek puts it, "unless it is for some reason morally impermissible for one person to take the same interest in another's welfare as he himself takes in it,

⁶Taurek also has the option of relying on our moral intuitions to deliver the verdict that it is morally permissible for me to give the drug to David in Friend, but it is an option that Taurek chooses not to exercise.

⁷Taurek, p.219.

it must be permissible for me, in the absence of special obligations to the contrary, to choose the outcome that is in B's best interest. And, of course, this is what I would do if B's welfare were more important to me than C's."⁸

I will discuss the validity of the Non-Relativity Principle, as well as just how crucial it is in arguing for Taurek's main conclusion, at a later time. For now I just want to finish laying out Taurek's first argument. It is important to note that the two judgments he reaches in the two examples I have already mentioned seem to agree with most people's moral intuitions, even among those who adhere to the standard view of the role of numbers in morality. What Taurek claims about the next example, however, clearly does not conform to the moral intuitions of the majority. I will call this example

Six Strangers: I have to choose between saving one total stranger by giving him all of the drug and saving five total strangers by giving each of them a fifth of the drug. I am, as in the two previous cases, under no contractual or quasi-contractual obligation to any of the people involved.

Once again, Taurek argues that it is morally permissible for me to give the drug to the one rather than to the five.

First, as in the earlier cases, no one's rights will be violated whatever I decide. Second, the Non-Relativity Principle, together with our verdict regarding First Person, already imply that it must be morally permissible for me to give the drug to the one stranger rather than the five. And, last but not least,

⁸Taurek, p.219.

Taurek asks the following question: if in general there are grounds for a moral requirement that five people rather than one should be saved if there is no way to save everyone, why are those grounds insufficient to generate a similar requirement in Friend? Can it really be the case, Taurek asks, that the moral force of those grounds is so weak that it can be overridden by the mere fact that I know and like David? As he puts it, "generally, when the facts are such that any impartial person would recognize a moral obligation to do something as important to people as giving this drug to these five would be to them, then an appeal to the fact that one happens to be partial to the interests of some others would do nothing to override the moral obligation"⁹.

Taurek adds that if the agent were, for instance, under a contractual obligation to deliver the drug to the five, this obligation could not be morally overridden either by the fact that the sixth person who needs the drug is someone the agent knows and likes, or even by the fact that the sixth person is the agent himself. So, in Six Strangers Taurek concludes that both alternatives are morally permissible¹⁰, just as in First Person and Friend. However, the Six Strangers case is unlike the Friend case in that I don't have any reasons for a personal preference that would ground my choice between the two morally permissible

⁹Taurek, p.217.

¹⁰Strictly speaking, for Taurek these appear to be permissible outcomes rather than permissible alternatives, since he seems to believe that giving the drug to either group outright, without, for instance, flipping a coin, is morally wrong. I will return to this in more detail in Section V.

outcomes. Taurek suggests that the fairest and morally correct thing to do in such a case is to flip a coin, or to perform some probabilistic equivalent, thus giving each of the strangers involved an equal 50% chance for survival. Anything else, he claims, would be arbitrary. If I do flip a coin, "it would seem to best express my equal concern and respect for each person. Who among them could complain that I have done wrong? And on what grounds?"¹¹

Subsection 3: Taurek's Second Argument.

When arguing against Taurek, many philosophers concentrate on the way he proceeds from First Person to Friend to Six Strangers in what I call his *First Argument*, claiming that if there are no grounds for a moral requirement to save the greater number in one case, there can't be such a requirement in any of the others. But, as I've indicated before, it is a mistake to believe that defeating this line of reasoning is sufficient to uphold the standard view of numbers in morality.¹² Since in the end, the standard view is usually justified by some sort of an appeal to the purported fact that by saving five people rather than one we avoid the greater of two evils, anyone who wishes to retain the standard view has to contend with what I will from now on refer to

¹¹Taurek, p.220.

¹²For one thing, if there was no more to Taurek's arguments against the standard view than his first argument, one could simply uphold the standard view by denying the intuitions backing up the claim that it is morally permissible for David to keep the drug for himself in First Person. Of course, denying those intuitions seems to be an extremely unattractive option.

as Taurek's Second Argument - the argument for the conclusion that, certain other things being equal¹³, we are not avoiding a worse outcome by saving five people rather than one; and that there is no interpersonal objective point of view from which it is just plain better if five people continue to live than if just one does^{14,15}. For reasons that will emerge in later sections, the second argument is a better, stronger and more interesting one than the first argument. In fact, at the end of the day, it is the second argument, but apparently not the first, that will remain standing to challenge the standard view of the role of numbers in morality. So let's look in detail at Taurek's

¹³The reason the claim has to be qualified with "certain other things being equal", is that I would like to retain the hope, against certain Relativist views, that there is, indeed, some interpersonal perspective from which it is, for instance, better that a serial killer be allowed to die than five, or even just one, of his potential victims.

¹⁴It has been argued by several philosophers, Judith Thomson among them, that there is no such property as goodness *simpliciter* and no such relation as being better than *simpliciter*. By talking about one alternative being better than another, I by no means wish to beg the question against this view. The reason I talk this way is that both Taurek and all of his critics that I am familiar with make free use of such locutions. Thus I choose to carry on this discussion on their common terms and table the very difficult issue of whether or not goodness and betterness exist for the purposes of this paper.

¹⁵Ralph Wedgwood has suggested that perhaps it can be argued that it is an intuitively obvious brute fact about morality that the size of the group facing a deadly threat is relevant. If so, then Taurek would be wrong in demanding an explanation for this purported fact from proponents of the standard view. But I do not find this suggestion at all plausible. There are, indeed, it seems, intuitively obvious brute facts about morality, such as, perhaps, that human suffering and death are morally significant. But the alleged fact about the role of numbers in morality doesn't have the same intuitive feel - it seems that if true, it should be justifiable using more basic facts about morality.

discussion of the claim that by saving more people we would prevent a worse outcome from obtaining.

It is, I think, helpful to keep in mind throughout my presentation of this argument that it is, in essence, a negative argument. Taurek doesn't claim to prove that there cannot, in principle, be a reason why it is morally preferable if five people are saved in a case like Six Strangers. What he tries to show is that the reason that most naturally comes to mind and that is likely to be cited by anyone who feels the need to justify the standard intuitions - that the deaths of five people involve more suffering, or a greater loss, than the death of just one person - is not a defensible reason; it cannot justify the standard view of the role of numbers in morality.

First, Taurek reiterates his refusal to think of human beings in the same way in which one would think of inanimate possessions - "it is not my way to think of them [the six strangers] as each having a certain objective value, determined however it is we determine the objective value of things, and then to make some estimate of the combined value of the five as against the one"¹⁶. What matters to Taurek, and this is quite possibly one of the most important points of his paper, is "the loss to the individual, not the loss of the individual"¹⁷. This is how he explains it: "The loss of an arm of the *Pieta* means something to me not because the *Pieta* will miss it. But the loss of an arm of a creature like me means something to me only because I know he will miss it, just as

¹⁶Taurek, p.222.

¹⁷Taurek, p.220.

I would miss mine. It is the loss to *this person* that I focus on. I lose nothing of value to me should he lose his arm. But if I have a concern for him, I shall wish he might be spared his loss."¹⁸

Next, Taurek argues that we can't add up different people's losses or suffering in a morally significant way. In *Six Strangers*, each of the six strangers involved would lose something very important to him if he does not get the drug - his life. But it seems wrong to argue that the loss that any one of the five strangers, each of whom requires only one-fifth of the drug, would endure, can be increased by the fact that four others will also not get the drug and share his fate. No one of the six strangers is in danger of suffering a loss greater than the loss of one's life. Should I choose to give the drug to the one rather than the five, no one will suffer a greater loss than the one would have suffered had my decision been different, let alone a loss five times greater than that loss.

Taurek illustrates this point further by offering the example in which the agent can spare someone a minor loss by suffering a much larger one. Very special circumstances aside, there is no sound moral reason for the agent to agree to such a tradeoff. Now suppose that the same agent can spare three people that minor loss by suffering a much greater loss himself. It seems clear that there is no better reason in this case than in the previous one for the agent to accept the tradeoff. That is because if the

¹⁸Taurek, p.220.

agent were to agree to the tradeoff, no one would be spared a loss close in intensity to the one he would suffer - all he would achieve is that each of the three other people involved would be spared a minor loss, the same loss the agent would have spared one person had he accepted the tradeoff in the original example.

If the agent refuses the tradeoff, three people will indeed suffer a minor loss A , but no one will suffer the loss $3A$, or, in fact, any loss greater than A , as a result. Thus there is no way to make sense of the claim that the total of suffering in this eventuality would be $3A$ in such a way that it could serve as an impetus for the agent to reconsider agreeing to the tradeoff, assuming, of course, that he would not have agreed to it in the initial one-one situation. Adding two extra people to the equation doesn't seem to make a moral difference.¹⁹ In contrast, had the agent had the opportunity to spare one person three losses equivalent to the minor loss in the previous examples, instead of

¹⁹It has been suggested by Judith Thomson that preventing the suffering of three people rather than the equivalent suffering of just one person, while it doesn't maximize goodness because there is no such property (see footnote 14), maximizes goodness-for in the sense that it is good for three people rather than just one. This seems reasonable enough; and it may be a promising avenue of thought for anti-Taurek theorists to pursue. However, I see no serious reason at this point why morality should be concerned with maximizing goodness-for in this way. Thus, until I am faced with an explicit argument to the contrary, I choose to table this discussion for lack of evidence that is threatening to Taurek's views.

A suggestion offered by Alex Byrne, that is, perhaps, roughly along the same lines, is that in cases where you prevent the suffering of many, the result is better than the sum of its parts (i.e. the prevented suffering of individuals). Again, I see no reason to believe that this is the case and so no *prima facie* reason to discuss the idea in detail until and unless I am faced with an explicit argument that threatens to undermine Taurek's position.

just one, he would perhaps have had a new reason to consider making the tradeoff.

The moral of this part of the story is that if there is no compelling reason for the agent to agree to trade his loss A for the loss B of any one of the individuals involved, there is no compelling reason for the agent to agree to trade his loss A for the loss B of every individual involved. Note that the conclusion appears to stay in effect even if the number of people that are facing the prospect of suffering B is changed from 3 to 10, or 25, or even 1,000,000. Similarly, Taurek believes, since no special considerations are involved, if there is no reason for me to prefer saving any one of the five in Six Strangers individually to saving the sixth stranger, it seems that there is no reason for me to prefer giving the drug to the five rather than the one. Thus, in order to be fair to everyone, I should flip a coin.

This concludes my presentation of Taurek's views. It should not come as a surprise that his controversial arguments and conclusions have come under a lot of fire from every direction ever since his article was first published. It is, on the other hand, rather surprising that philosophers in general, and ethicists in particular, have to a large extent rejected the main points of Taurek's article out of hand, without giving full and fair consideration to the complexity of the problem and the truly compelling parts of his arguments. In discussions of somewhat related problems in moral theory, such as the Trolley Problem or the significance of the killing/letting die distinction, it has become common practice to just drop a footnote acknowledging that

Taurek disagrees with the view that numbers should matter in morality, and then simply refer the reader to one of the more famous critiques of his article, thus believing the case to have been closed. In the sections that follow, I will attempt to reply to some of the more serious objections that have been leveled against Taurek's views by his opponents, in the hope of showing that Taurek's conclusions cannot be dismissed so easily and that, contrary to what most philosophers seem to believe, the case is far from closed.

Section III.

Derek Parfit and Agent-Relative Permissions.

In "Innumerate Ethics", Derek Parfit's response to Taurek's "Should the Numbers Count?", Parfit disputes Taurek's conclusion that numbers do not matter in morality. First, in subsections 1 and 2, I want to discuss Parfit's argument for the claim that all other things being equal, it is a worse thing when more, rather than fewer people die²⁰. Should Parfit's arguments for that proposition fall short and thus fail to derail Taurek's second

²⁰Those who object to talking about worse things or worse outcomes without specifying who or what they are worse for (see also footnote 14), can replace the offending expressions with "outcomes that there are moral reasons to avoid". For the purposes of this paper I will use these expressions interchangeably, as in most cases just saying "worse" as shorthand for the much longer but less controversial expression is very convenient. Parfit says something very similar in footnote 10 of his article.

argument, as I believe they do, I will argue that Parfit will already have failed to reach his objective of disproving Taurek's main conclusion, regardless of anything he can add to counteract other arguments and claims made by Taurek in "Should the Numbers Count?". Then, in subsection 3, I will proceed to discuss the other part of Parfit's paper, the one dealing with Taurek's first argument and agent-relative permissions, anyway, as I believe that some very important issues will inevitably be raised in the process.

Subsection 1: Parfit vs. Taurek's Second Argument - "Sums of Suffering".

To strike down Taurek's conclusion Parfit wants to show that Taurek is wrong in claiming that if everyone cannot be saved, there is no moral reason for saving five strangers rather than just one. Taurek suggests that if there were such a reason, it would have to do with the amount of suffering caused by each of the outcomes. That is, indeed, the reason why most, if not all, proponents of the standard view of the role of numbers in morality defend it: they believe that all other things being equal, greater suffering is caused by the deaths of five people than by the death of just one person. This is not proof that no moral reason that is not derived from the comparison of losses or suffering can exist for preferring to save the five in Six Strangers. But it is hard to imagine what such a reason could possibly consist in, and so the burden of proof shifts from Taurek to his opponents to show

that such a reason can, in fact, exist²¹. Thus the task at hand is to analyze whether or not it is plausible that the amount of suffering that would be caused by the deaths of the five strangers in Six Strangers is greater than the amount of suffering that would be caused by the death of the one.²²

Parfit says that "those who believe that suffering is "additive" do not believe that many lesser pains might be the same thing as one greater pain. What they do believe is that the

²¹Although see footnote 19 for possible glimpses of an idea...

²²Before I proceed, I want to digress for a moment to address an indirect argument for the conclusion that the deaths of the five is not a worse outcome than the death of the one. Taurek uses this argument in his discussion of the Friend case. Parfit summarizes it in the following way:

First Premise: If we prefer the worse of two outcomes we are morally deficient.

Second Premise: David would not be morally deficient if he preferred that we save him rather than the five.

Conclusion: The deaths of the five cannot be a worse outcome than the death of David.

Parfit is quite correct in pointing out that there is no reason why the first premise has to be true. It is far from obvious that preferring a morally worse outcome is always a sign of a moral deficiency. It is also far from obvious that the second premise is true. It is very natural for someone in David's position to want to be the one saved and, in fact, almost anyone would have just that preference. That, however, does not rule out the possibility that David is morally deficient for having such a preference, as is everyone else who would prefer to be saved in his place. As far as both premises are concerned, Taurek is just relying on his intuitions without further justification. And those intuitions are far from completely obvious. So, regardless of whether or not the conclusion is true, both of the premises cannot be assumed to be true without a much more serious discussion. So Parfit is right in discarding the argument. However, as Parfit himself must have realized, this argument is just a comment made by Taurek in passing to make one of the premises for his major arguments sound more plausible; and it is not intended to be one of Taurek's big guns.

lesser pains might together be as bad."²³ For one thing, Parfit remarks, pains do seem to be additive when we are only talking about a single individual. For instance, Taurek is unlikely to dismiss as utter nonsense a person's decision to avoid fifty minor headaches that would be suffered over some relatively short period of time instead of avoiding one major migraine. The comparison between the amount of suffering that would be caused by the migraine and the amount of suffering that would be caused by the fifty headaches is, indeed, very rough. One would probably be quite incredulous if told that the suffering caused by one migraine is precisely equal to the suffering caused by thirty seven minor headaches. However, it certainly seems to make sense to suppose that even though a migraine is much, much worse than a minor headache, if the number of minor headaches that need to be endured in order to avoid a very bad migraine is high enough (be it fifty, a hundred, or even more), it is wiser to refuse the tradeoff. In fact, there is probably a number high enough so that no one in their right mind would accept the tradeoff, provided no special factors are at play, such as that the person is in very weak health and a major migraine could kill him whereas no amount of minor headaches can.

Parfit now wants to make the jump from one person suffering fifty headaches to fifty different people suffering one headache

²³Derek Parfit "Innumerate Ethics", ETHICS Problems and Principles, ed. by John Martin Fischer and Mark Ravizza, Harcourt Brace Jovanovich College Publishers 1992, p.231;
All future quotations from Parfit are from this article.

each. He believes that the sum of suffering would be about as great in the two cases, where the phrase "sum of suffering" is to be understood the same way in the case of fifty different headache sufferers as in the case of one person suffering fifty headaches. In both cases, he argues, any comparison between such a "sum of suffering" and a single loss by a single individual (or another "sum of suffering") would be rough and imprecise, but not meaningless, as Taurek might have us believe. After quoting a passage from C.S. Lewis that supports Taurek's contention that all attempts at morally significant interpersonal aggregation of suffering are nonsensical, Parfit asks: "like Taurek, Lewis assumes that any "sum of suffering" must be felt by a single person. Why not add that it must be felt at a single time? That would reduce even further the Problem of Pain..."²⁴ But, he responds, both requirements are unjustified - "suffering at other times is more suffering. So is the suffering of other people."²⁵

Parfit claims not to see any difference between the two cases. But I believe that his analogy is very strained. The reason why we seem to care about suffering in the first place, and why we believe that it is relevant to morality, is that somebody suffers it. When we are talking about someone enduring fifty minor headaches, the reason why it is important that we are talking about fifty headaches and not just one, is that there is a person, a conscious living being, who suffers (and, presumably, remembers suffering) all of them. And it is much worse for any such being

²⁴Parfit, p.232.

²⁵Parfit, p.232.

to suffer fifty headaches than to suffer just one, which is why a person's pain does not necessarily have to be suffered at one instant in order for its aggregate to have moral significance. However, in the case of fifty different headache sufferers there is no being, self-aware or otherwise, that suffers an aggregate pain of more than one headache. If the universe itself were conscious and aware of the pains of the parts that make it up, then and only then would it perhaps make sense in this case to talk of a morally significant aggregate of suffering roughly equivalent to that of fifty headaches.

Parfit offers the following: "I suggest that he [Lewis] confused two different claims. He makes the factual claim that the suffering of more people cannot be more suffering. He may have meant the moral claim that it cannot matter more. He may have thought the suffering of one person to be as great an evil as the suffering of a million"²⁶. He adds, "when he [Taurek] says that "suffering is not additive", he too may not mean that the pains of different people cannot be more pain. He may mean that these pains cannot be *morally* summed - that they cannot together make an outcome worse."²⁷ I agree that this is, indeed, what Taurek and Lewis both claim. But it is unclear to me what kind of confusion Parfit is accusing them of. You can, of course, attach some meaning to the claim that the "sum of suffering" of five people with a headache is greater than the suffering of just one person with the same kind of headache. After all, there are four

²⁶Parfit, p.232.

²⁷Parfit, p.232.

more people suffering the headache in the former case than in the latter. But that was never in dispute. What was and still is in dispute is whether the fact that there are more people suffering a pain or a loss A makes the "sum of suffering" greater than A in any meaningful way - in particular, in a way that is morally significant.

Taurek's point and, I presume, Lewis' as well, is that no moral significance can be attached to such an aggregate of suffering because there is no one who suffers it. Of course, that doesn't mean that people can't agree on a non-moral convention that considers five people's suffering of A a greater amount of suffering than the suffering of A by one person, provided that the convention is not in itself intended to carry any moral significance. In fact, such an agreement seems to be in place in everyday discourse, which is why we have to be extremely careful about not drawing any moral conclusions from the convention without further justification, which it seems that Parfit was unable to provide. Parfit's analogy with one person's suffering of fifty minor headaches fails to counteract the arguments by Taurek and Lewis, as in such a case there is someone, namely that person, who suffers all of those headaches. Thus, so far, Parfit seems to have been quite unsuccessful in his struggle against Taurek's second argument.

Subsection 2: Parfit vs. Taurek's Second Argument Continued - Maximin.

While I believe that Parfit has not made any kind of a dent in

Taurek's claim that individual losses suffered by different people cannot be morally summed, he does make an additional attempt to undermine one of the ways in which Taurek arrives at that claim. Before discussing it, however, I do want to point out that even should Parfit's objection succeed, in light of what's been said in the preceding subsection, he will still have failed to bring down Taurek's main conclusion altogether, having only managed to discard one supporting argument. The best Parfit can hope for here is to make Taurek's claim sound slightly less plausible.

Parfit's criticism is focused on a principle derived from our intuitions regarding rather uncontroversial cases, such as the ones discussed towards the end of section II. Parfit calls this principle

(N): We ought to save one person from harm rather than saving any number from smaller harms.²⁸

The first remark that needs to be made is that this is certainly not Taurek's view. The most he would commit himself to is that there is a moral reason for saving one person from a great harm rather than any number from a smaller harm. It should be clear from his discussion of choosing between saving one person's arm and another person's life, however, that he would never accept the "ought" in the formulation of principle (N). Whether or not he is correct in this opinion is a separate issue, which, fortunately, is besides the point here.

That objection aside, this is how Parfit reconstructs Taurek's

²⁸Parfit, p.232.

argument. Unless we believe that

(M): If one person is harmed, that would be just as bad as if any number are each equally harmed^{29 30}

is true, he has Taurek arguing, we can't explain the truth of (N) or even the truth of the much more obviously intuitively true

(O): We ought to save one person from harm rather than saving any number from *much* smaller harms.³¹

Parfit suggests that instead of (O) we should accept

Maximin: We should give priority to helping those people who would be the worst off if they aren't helped³²

as the principle that best describes what we should have in mind for the kind of tradeoff cases that Taurek considers. Parfit then tries to argue that Maximin, unlike (N) or (O), can be explained without reference to (M).

This is how I understand what Parfit is saying here. He believes that it is Maximin, rather than (N) or (O), that is really our, as well as Taurek's, view. (O), he claims, gives us the intuitively correct answers only when those answers coincide

²⁹Parfit, p.232.

³⁰Of course, we would need to insert some sort of an "all other things being equal" clause here if we want to be precise.

³¹Parfit, p.232.

³²Here I will once again gloss over the fact that Taurek would never accept any principle like this which includes a "should" or an "ought", even if he agrees that there is a moral reason to do what such a principle tells us to do.

with the answers that Maximin would give. But Maximin, unlike (O) or (N), has, at least at first glance, nothing to do with any kind of a notion of a "sum of suffering". Therefore, Parfit believes that Maximin can be justified without relying on a thesis like (M), and so Maximin can explain all the intuitions that Taurek claims we need (N) or (O) to explain, without committing us to any Taurek-like conclusions about cases like Six Strangers.³³ Parfit's hope is that in cases like Six Strangers, where each of the six strangers involved stands to lose as much and end up as badly off as any one of the others, we can then still conclude that it is the five strangers that should be saved rather than the one because more suffering would be caused by the deaths of the five.

I see many problems lurking here. First, Parfit relies in the end on claiming that he has already established that there is, in a morally significant way, more suffering involved in the deaths of five strangers than in the death of just one. And unless the "sum of suffering" if five die would be greater than the suffering involved in the death of the one, it seems that Parfit has no grounds for arguing, as he does, that numbers should be used as a tiebreaker when Maximin fails to deliver a verdict, as, for instance, in Six Strangers. However, as I have claimed earlier,

³³I am very far from convinced that Maximin can possibly be justified without appealing to a principle like (M). It seems to me that even though, *prima facie*, it appears as if Maximin has a lot less (if anything at all) to do with aggregates of losses or suffering across individuals than (O) does, it would ultimately turn out to be the case that this is an illusion and that it would be exactly as hard (or as easy) to justify Maximin without falling back on a principle like (M) as it would be to justify (O). But I have enough bones to pick with Maximin, as stated by Parfit, and with Parfit's entire argument, to not need to place much weight on this objection.

it seems that Parfit has failed in his attempt to argue for morally significant sums of suffering.

Second, even if Parfit were allowed to make the "sum of suffering" assumption, it is unclear what would justify using Maximin in cases of "uneven" suffering, but relying on some sort of a "sum of suffering" argument as a tiebreaker when Maximin can't give you an answer because everyone stands to end up in the same boat (in Charon's boat, if we are talking about Six Strangers). Why not flip a coin when Maximin has nothing to say? That's a conclusion Taurek would be quite happy with. And finally, I am unclear about how Maximin, as stated by Parfit, is to be justified, and about how it can fit in with the intuitions that (O) fits so well. These concerns support my belief that Parfit is wrong in claiming that (O) is not our real view.

Parfit seems to assume that when Taurek talks about the loss that someone would incur or the harm that would come to them, no facts about the prior condition of the person are included in the deliberations. But does that really make sense? It is a much greater loss for a blind person to also become deaf than it is for a person who is not blind. It is also probably a much greater loss for someone who is already paralyzed to also become blind than it is for someone who is not paralyzed. And there is a greater amount of suffering involved in enduring an additional ten minutes of a migraine headache for a person who has had a nonstop migraine for ten hours already, than there is in having a ten minute long migraine headache for someone who has been feeling fine for a long time. Parfit suggests the following example,

which I will call

Agony: "For each of many people, yesterday was agony. For some other single person [call him Z], it was a day of minor pain. Ought we now to save this person from a day of agony rather than saving each of the many from a day of minor pain?"³⁴

First of all, it has to be made absolutely clear that there is no issue of unfairness involved here, or at least there shouldn't be if the example is to be of any use to us in our current enterprise. The way Parfit phrases the example, it might sound as if Z somehow deserved to be in agony yesterday but managed to escape it in some sneaky way. If that were the case, the example would be entirely useless to us. So let's assume that that isn't the case. Parfit believes that there is no moral reason for helping Z which is what he claims Maximin would dictate as well. On the other hand, Parfit claims that according to (O) it is Z that should be helped and spared the agony.

But why does Parfit say that? I can see only two alternatives in analyzing Agony. The first alternative is that having suffered through the agony makes it much more difficult to endure the additional minor suffering. In that case I believe that Parfit is wrong in maintaining that (O) tells us to help Z, as the potential harm to each of the many today would not be much smaller than the potential harm to Z today. The second alternative is that the agony that was suffered yesterday does not make it more difficult to endure the minor suffering today or, perhaps, does make it more difficult but only slightly. In that case, (O) would, indeed,

³⁴Parfit, p.233.

tell us to help Z. But then Parfit would seem to be wrong in maintaining that there is no moral reason to help Z. In fact, I believe that even Maximin would not support that claim of his under the circumstances. In either case, Parfit's verdict regarding (O) seems to be misguided, leaving him with no grounds for stating that it is Maximin rather than (O) that should be adopted.

As for Parfit's example of saving one arthritic from blindness rather than saving any number of the blind from arthritis, the same argument seems to apply as to the Agony case. We should determine, assuming that such a determination is, indeed, possible, whether or not it is a much greater loss for an arthritic to become blind than it is for a blind person to become afflicted with arthritis. If it is, then I think it is obvious that regardless of what Maximin would say about it, we should help the arthritic, as (O) would, indeed, tell us to. If it isn't, then, contrary to what Parfit would apparently say, (O) would not require us to help the arthritic (a possibility which I do not, in all honesty, find very plausible). In that unlikely, bordering on impossible, case, we would not be compelled to help the arthritic even if the tradeoff involved only one blind person, even though we would, perhaps, have a moral reason for helping the arthritic anyway since his potential suffering would still likely be at least somewhat greater.

In general, I believe that Parfit has failed to discredit (O). I also have serious reservations about Maximin. Unless the person somehow deserved to have lost or suffered more in the past than he

has, especially under circumstances that relate to the current dilemma, it is unclear how much we should take into account who would be the worst off altogether, provided that there is a great disparity between the losses that would actually be suffered as a result of our decision. If I were given a choice between saving the arm of a person who had previously lost his little toe and saving the little toe of a person who had previously lost his arm, why on Earth wouldn't I have a valid reason to save the first guy's arm? To me it seems ridiculous to do what Maximin, at least under Parfit's interpretation of it, would dictate, and to be guided by the fact that if I help no one the two would end up equally badly off, rather than by the much more morally relevant great disparity between the potential losses. And what looks even worse for Parfit's Maximin is that it apparently concludes that a "fully intact" person's arm ought not to be saved if doing so would mean that a one-armed person would lose his little toe. Maximin, as formulated and applied by Parfit, just doesn't seem to be a morally valid principle.

I believe I have shown that Parfit has not managed to discredit (O) and thus appears to be unable to explain some of our moral intuitions without referring to a principle like (M). Therefore, he has failed to impeach this particular argument for (M); and from (M) Taurek's main conclusion seems to follow quite easily. In addition, as I've argued in subsection 1, Parfit's direct attempts to argue for the opposite of (M) have apparently fallen short. Thus I have to conclude that Parfit has not succeeded in undermining Taurek's second argument.

Subsection 3: Parfit vs. Taurek's First Argument - Agent Relative Permissions.

In the first part of his paper, Parfit tries, through his use of the notion of agent-relative permissions, to argue that Taurek's first argument, i.e. the line of reasoning that proceeds from the First Person and Friend cases to the Six Strangers case, is flawed. Parfit maintains that in Six Strangers the agent ought to save the five rather than the one. However, in First Person, as well as perhaps even in Friend, the agent has an agent-relative permission to save the one rather than the five. But, in light of the fact that the second part of Parfit's paper has apparently not been successful, Parfit cannot hope to achieve all of his objectives. Without a successful defense against Taurek's second argument for his conclusion in Six Strangers, Parfit has failed to show that I ought not flip a coin in Six Strangers to determine who to save. Thus the most Parfit can conclude, at least without presupposing his own desired conclusion, is that Taurek's first argument is flawed and what is really going on in First Person, as well as perhaps in Friend, is that the agent has an agent-relative permission to forgo the coin toss and simply save the one rather than the five. That conclusion would hardly be very damaging to any of Taurek's main theses or arguments³⁵. However, I believe that it's worth it to consider Parfit's argument regarding agent-relative permissions for its own sake. And the topic of agent-

³⁵At least it seems that way now. See section V for a less optimistic (as far as Taurek is concerned) analysis of this problem from a slightly different angle.

relative permissions will resurface in section V.

In this part of his paper Parfit focuses on a case which Taurek himself considers as a potential counterexample against his first argument - the example involving a tradeoff situation where one person stands to lose his life while the other stands to lose only his arm. Taurek believes that just as it is morally permissible for David in the First Person case to save his own life rather than the lives of five strangers, it would be morally permissible for David to save his own arm even if that would prevent him from saving somebody else's life. I am not entirely convinced that this is correct, but I will not get into this issue here as it is not very crucial for the purposes of the discussion at hand. If Taurek's first argument is valid, a similar line of reasoning would lead to the conclusion that it is morally permissible for me to save one stranger's arm at the cost of not being able to save another stranger's life.³⁶ Taurek does not have a problem with this as he explicitly embraces the seemingly counterintuitive claim in his paper. Parfit, however, finds the conclusion too outrageous to swallow.

This is how Parfit reconstructs Taurek's argument: Taurek assumes

(B): "It would not be true that we ought to save this stranger's life at the cost of our own arm".³⁷

³⁶Of course, we are only talking about situations that don't involve any of the special considerations that were discussed at the beginning of section II.

³⁷Parfit, p.228.

and

(C): "Whether we ought to save this stranger's life cannot depend on whether it would cost us nothing, or our own arm";³⁸

and so concludes that the following is false -

(A): "If we could save either one stranger's life or another stranger's arm, and it would cost us nothing to do either, we ought to save the first stranger's life"³⁹.

For now we have agreed to let (B) pass. So, as Parfit points out, we now have to choose between (A) and (C). He finds (A) to be much more plausible, and thus believes that it would be much more reasonable to reject (C) than to give up (A). However, I believe that the way Parfit phrases (C) has a lot to do with the ease with which he arrives at the choice he favors. Consider the following modification,

(C)': Whether we ought to save this stranger's life cannot depend on whether it would cost us or someone else an arm.

I believe that (C') captures what Taurek had in mind a lot better than (C). Exactly how implausible does (C') seem? I don't think Parfit would be as quick in judging the negation of (C') to be less counterintuitive than the negation of (A), as he was with (C) in place of (C'). In fact, if I had to choose between (C') and (A), I believe I would discard (A). I will return to C' and explain why I believe that it is true, shortly, but first let's consider Parfit's argument.

³⁸Parfit, p.229.

³⁹Parfit, p.228.

Parfit gives an explanation for why he thinks that (C) is false. Since the difference between (C) and (C') is mostly psychological (the latter simply draws attention to a factor that the former glosses over - namely, that someone else stands to lose an arm and that the only difference is that it's not the agent himself who will suffer the loss), Parfit's explanation, if valid, also justifies the rejection of (C'). In any case, it leaves Parfit free to keep (A) and, by analogy, the standard view of the role of numbers in morality⁴⁰.

Parfit says that along with Taurek he believes that one may give priority to one's own welfare. That priority is, of course, limited, e.g. "perhaps Y could save his arm rather than X's life, but he ought to save X rather than his own umbrella"⁴¹. Parfit also believes that we may give priority to the welfare of others, e.g. our children and, perhaps, other family members or loved ones. It is also possible that we may give priority to the welfare of close friends. But that, Parfit argues, is as far as the unequal weighing of people's welfare can go. This is why, Parfit explains, it is morally permissible for David to keep the drug for himself in First Person, and why it might be morally permissible for me to give the drug to David in Friend, although we would be on much safer ground if David were a relative or a

⁴⁰Provided, of course, that we reject Taurek's second argument against the standard view, which, I have argued, Parfit has failed to convince us to do.

⁴¹Parfit, p.229.

loved one, or at least a close friend⁴². But, according to Parfit, it is not morally permissible for me to give the drug to the one stranger rather than the five in Six Strangers, with or without tossing a coin first. The permission in First Person is *agent-relative*.⁴³

Agent-relative permissions might also be useful in responding to Taurek's comment that were we under contract or military duty to give the drug to the five, it would not matter one iota whether or not we or someone we know and like, or even someone we love, needed the drug. Parfit responds that "contracts and military duties give rise to special obligations. Perhaps we ought to carry out these even at a heavy cost to ourselves. But this need not be true of everything we ought to do. It may not be true of saving the lives of strangers"⁴⁴.

⁴²Parfit does not believe that the agent-relative permission to not give the drug to the five applies to the case where the one is someone the agent knows and likes, even though it would apply if the one were a close relative or a loved one. But there is no reason why we have to agree with Parfit on where to draw the line if we are willing to draw it in the first place.

⁴³Aside from the objections to agent-relative permissions that I am about to discuss, there is also the following complaint: it seems that without a much deeper explanation of what is going on and what justifies agent-relative permissions in the first place, Parfit can be accused of simply restating alleged facts by using a new term when he attempts to explain his moral intuitions by claiming that agent-relative permissions are involved. What Parfit is doing can, perhaps, be perceived in the same way as if a hypothetical proponent of the standard view of the role of numbers in morality tried to explain the purported difference between Six Strangers and a case in which I have to choose which one of two people to save, by claiming that the moral permission to save one person in the latter case is numbers-relative, and argued that the issue is thus settled. Such a remark would hardly be illuminating. I thank Judith Thomson for directing my attention to this problem.

But how plausible is Parfit's explanation? First, we may ask with Taurek why exactly it is the case that we may give priority to the welfare of ourselves and certain special others, but not to the welfare of strangers. I don't find what Parfit says sufficient to answer this question. Second, I'd like to offer an example that I think causes problems for the notion of agent-relative permissions. Consider the Five-One-One case -

Five-One-One: I have a quantity of the life-saving drug. I need all of it in order to survive. A stranger, let's call him Jack, also needs all of my drug in order to survive. Each of five strangers, A, B, C, D, and E, needs only one-fifth of the drug in order to survive.

According to Parfit, I have an agent-relative permission to keep the drug for myself even though in general one ought to save five people rather than just one. However, Parfit is apparently committed to saying that if I choose not to act on my agent-relative permission to keep the drug, I ought to give it to the five because, agent-relative permissions aside, one ought to save five people rather than one.

I find that conclusion unacceptable. Selflessness is usually considered a virtue and, at the very least, it is very rarely considered a vice. If I am morally permitted to withhold the drug from A, B, C, D, and E, I fail to see how it can be the case that I am not then morally permitted to perform a selfless act and sacrifice my life in exchange for Jack's. That may, indeed, seem like a very strange thing for me to do, but that's beside the

⁴⁴Parfit, p.230.

point. This is one reason why I believe that the notion of agent-relative permissions may be inherently flawed. Another reason is that I think it is generally wrong to sacrifice something that belongs to someone else, especially if you would not be willing to make the same sacrifice for the same goal in that person's place. Similarly, I believe that it can't be morally impermissible for me to allow a very bad outcome A (e.g. the deaths of five people) to obtain in order to prevent a loss B (e.g. death) from befalling someone, if it would not, under similar circumstances, also be morally impermissible for me to allow A to obtain in order to prevent B from happening to me.

As an example, consider a variant of the standard Trolley case. Suppose that I am the only passenger on a trolley, and the trolley driver has just died of a heart attack. This way issues of job-related duties shouldn't enter into our moral considerations in a way in which they might have had to if the agent were himself the trolley driver. Suppose also that if the trolley continues straight ahead, it will derail and hit a nuclear power plant, causing a nuclear explosion that will kill millions, including me and everyone else involved. My options are to turn right, which would result in five deaths, or to turn left, which would result in only one death. It seems clear that whatever else I decide, I should turn the trolley one way or the other. I just don't see how the true theory of morality can possibly require me to divert the trolley so that it kills one person rather than five, if it wouldn't also require me, under otherwise similar circumstances, to divert the trolley so that it goes over a cliff

and I die in order to avoid the deaths of the same five people as well as the deaths of millions of others.⁴⁵ That is the reason why I prefer (C') to (C) - (C') emphasizes the fact that someone's arm will be lost to save a life, the only question is whose arm it will be. As is by now clear from my thoughts on the subject, I think that regardless of anything else (C') should be accepted.

If we keep (C'), we have to choose between (A) and (B). Either, special circumstances aside, one is both morally required to refrain from saving one's own arm instead of a stranger's life and to refrain from saving another stranger's arm instead of the first stranger's life; or, special circumstances aside, one is never morally required to save B's life rather than C's arm. Taurek is convinced that the latter is the case. I am not so sure, but I am definitely inclined to agree that (A) and (B) cannot be true at the same time⁴⁶. I also agree with Taurek's psychological explanation for why many of us want to have it both ways. He suggests that most people believe that in general a

⁴⁵This connects in interesting ways to the issue of externalizing costs, which is fascinating in its own right. Unfortunately, it falls outside the scope of this paper and so I have to leave a discussion of it for a later time.

⁴⁶In fact, I definitely disagree with the choice Taurek makes as, even laying initial intuitions aside, accepting (B) and the negation of (A) leads to what appear to be quite absurd and objectionable consequences for a stranger's arm vs. another stranger's life tradeoff situation. Should we flip a coin in that case just as in Six Strangers? Taurek is not definitely committed to that as (I will discuss this in section V) he can maintain that it is not arbitrary, although not obligatory either, to prevent the maximum amount of suffering to any one person. But it seems to me somewhat counterintuitive that it is could even be morally permissible to flip a coin in such a case. Would flipping a coin then also be permissible in a finger vs. a life case? A toenail vs. a life? An umbrella vs. a life?

life should be saved rather than an arm, provided that both cannot be saved. But they recognize the weakness of human nature in themselves and admit that they would, because of the great personal cost involved, find it virtually, if not entirely, impossible to give up a chance to save their own arm in order to save a stranger's life. On the other hand, that personal cost isn't there if the choice is between saving B's life and C's arm. That is what makes keeping both (A) and (B) so attractive. It allows people to retain a general view that they believe to be correct, while avoiding having to face very unpleasant consequences of that view, which, if confronted, might brand them morally deficient. However, I think that Taurek does, unfortunately, have it right in maintaining that this easy way out can't be morally justified and we have to choose between (A) and (B) after all.

Since I believe it to be possible and, in fact, quite likely that (B) is true while (A) is false, I also have to admit, by analogy, the possibility that Taurek's judgments regarding the First Person and Six Strangers cases are both false, provided, of course, that I temporarily disregard Taurek's second argument for his conclusion in Six Strangers. I find that to be a highly unlikely possibility. Apparently Parfit does too. Whatever his beliefs about Six Strangers, Parfit definitely thinks that one is morally permitted, special circumstances aside, to save one's own life rather than the lives of five strangers. Thus I believe that at this point I have no reason to consider Taurek's first argument against the standard view of the role of numbers in morality to be

in any peril as a consequence of arguments offered by Parfit⁴⁷. Combined with what I perceive to be a failure of Parfit's in his attempts to undermine Taurek's second argument, that leads me to the conclusion that Parfit has struck out and Taurek's view remains a viable alternative to the standard one, at the very least until the next section.

Section IV.

The Objection From Transitivity and Other Criticisms by Gregory Kavka.

Subsection 1: Taurek on Rights and Obligations.

In section 1 of the first part of his paper "The Numbers Should Count", another critique of Taurek's views and arguments, Gregory S. Kavka also discusses something along the lines of agent-relative permissions, although he doesn't call them by that name. He points out, quite correctly, that if morality permits partiality to one's own very important interests, as well as the very important interests of loved ones and friends, then Taurek's judgments regarding the First Person and Friend cases shed virtually no light on the problem of numbers in morality - at

⁴⁷I will, however, return to this topic briefly in section V, as I believe that the analysis that will be required in order to discuss objections by Gregory Kavka and Ken Dickey shows that there may be a reason for Taurek to be gravely concerned about agent-relative permissions after all.

most, we can find out that there are some considerations that can override the numbers even if numbers do, indeed, matter. Such a view of morality would make it very difficult for Taurek to reach any interesting conclusions regarding Six Strangers by analogy with those two cases, and thus would invalidate Taurek's first argument, although it would, of course, leave Taurek's second argument untouched.

Unlike Parfit, Kavka offers some reasons for accepting something like agent-relative permissions. In particular, he argues that a system of morality is more practical and socially valuable if it incorporates rules allowing, under certain circumstances, partiality to one's own interests and the interests of loved ones. However, in light of the discussion in the previous section, and especially since this line of criticism doesn't target Taurek's second argument at all, I have to conclude that Kavka doesn't give us a sufficient reason to reopen the discussion of agent-relative permissions until, as I have promised, section V.

In the second section of the first part of his paper, Kavka discusses Taurek's views on moral obligations and how those views affect Taurek's first argument⁴⁸. This line of criticism doesn't concern Taurek's second argument either, and so by itself doesn't

⁴⁸The term 'obligation', used more or less indiscriminately in the literature without much clarification of what is meant by it, is very murky. However, since all of the philosophers involved in this discussion use it, I will often have little choice but to do the same in the course of my paper. If you feel great discomfort at the sight of "A is obligated to ...", or "A has an obligation to", feel free to mentally replace them with "A is under moral pressure to ...", or some equivalent expression.

threaten to disprove Taurek's conclusions, but the discussion is very interesting and, I believe, does have crucial repercussions for Taurek's view of morality as a whole. Kavka argues that Taurek subscribes to what he calls "the libertarian account of moral obligation". He describes the libertarian account as the view that recognizes only three kinds of moral obligations - "duties of non-interference, that require one not to assault or kill others (except to repel or punish an attack), and not to interfere with their actions or chosen uses of their property (except insofar as such acts and uses violate the rights of others) ... contractual and quasi-contractual obligations ... [and] duties of restitution that require one to compensate another if he has suffered harm as the result of one's unjust, reckless, or negligent acts"⁴⁹. A key part of such a view is the belief that "there is no right to survive that would impose obligations on others to use their property to keep one alive"⁵⁰.

Kavka does, indeed, have good reasons to ascribe some such view to Taurek. Perhaps the most serious reason is Taurek's continued insistence that special circumstances aside, the owner of the drug violates no one's rights by choosing not to give any of it to people who need it to survive. Taurek can even be interpreted as saying that if there are no special considerations involved, the agent isn't morally required to help anyone. The

⁴⁹Gregory S. Kavka "The Numbers Should Count", ETHICS Problems and Principles, ed. by John Martin Fischer and Mark Ravizza, Harcourt Brace Jovanovich College Publishers 1992, p.239; All future quotations from Kavka are from this article.

⁵⁰Kavka, p.239.

key factor in his explanation of these opinions seems to be simply that the agent owns the drug and so doesn't violate anybody's rights by not giving it to them unless special considerations are involved. And it is important to note that the absence of special considerations, as discussed at the beginning of section II, precisely rules out the second and third kinds of duties in Kavka's account of the libertarian view, while the first kind is inapplicable to cases like Six Strangers to begin with. As a result of this and other remarks which Taurek makes in "Should the Numbers Count?", I will concede without argument that Taurek holds some sort of a libertarian view.

Kavka suggests that, as most people holding some kind of a libertarian view of moral obligation, Taurek subscribes to some sort of a **Libertarian Aid Principle (LAP)**

(LAP): "Person A is obligated to aid person B *only* if either (i)A has contracted to do so (or is bound to do so in virtue of some quasi-contractual relationship) or (ii)A is responsible for B's needing aid or owes B compensation for other past harms"⁵¹.

As Kavka immediately points out, it seems to be the case that (LAP) has "some very counter-intuitive consequences. It implies, for example, that it is permissible for the owner in case 2 [the Friend case] to simply pour the drug down a sewer, and save neither his friend nor the five strangers."⁵² Also, if it is something like (LAP) that Taurek uses to justify his verdicts in First Person and Friend, then his discussion of those cases can't

⁵¹Kavka, p.240.

⁵²Kavka, p.240.

succeed in telling us much about the role of numbers in morality. Kavka also argues that any thesis about the role of numbers in morality follows from (LAP) only if we also accept the following principle

(O): "Moral reasoning and decision making about giving aid is concerned solely with determining one's moral obligations (in the libertarian sense)".⁵³

The conjunction of (LAP) and (O) does, indeed, appear to imply that numbers don't matter in morality. Unfortunately, that conjunction also implies that virtually nothing matters in morality, including many things that all of us, including Taurek, want to turn out to matter, such as whether or not human lives are at stake. For instance, to use an example that Kavka offers, if in Friend or Six Strangers the agent doesn't need the drug for his survival but feels like drinking it because, for instance, he likes the way it tastes, (LAP) and (O) together imply that it is morally permissible for the agent to drink the drug and let all six die.

I doubt that Taurek wants things to come out that way. And I don't know what Taurek's views actually are regarding (LAP) and (O), and whether or not he has a way of making them compatible with common sense and his other views. But my task in this paper is not to defend everything that Taurek says or believes, but only to defend his challenge to the standard view of the role of numbers in morality. And I believe that there is an easy way to

⁵³Kavka, p.240.

make Taurek's main arguments work in spite of this problem. Whether or not Taurek would endorse this kind of a view is a separate question, although I hope that he may have had something roughly like this in mind.

I think that Taurek is wrong from the outset when he tries to justify his judgments regarding Friend and First Person by an appeal to the purported fact that the agent would be violating no one's rights if he gave the drug to the one rather than the five. And the reason I believe this is a mistake is exactly the reason that Kavka hints at. The question of whether or not anyone's rights are violated, especially if rights are construed as narrowly as they are on the libertarian account presented by Kavka, doesn't entirely settle the issue of whether or not an action is morally permissible. For instance, it should, at least in principle, be possible that the agent violates no one's rights by keeping the drug for himself (say, to save his own arm) and, in fact, even has the right to do so, but it is still morally wrong for him to do it.⁵⁴

Especially if rights are construed in a libertarian way, one's moral theory has to leave room for the possibility that the agent would violate no one's rights by doing P and yet would be doing

⁵⁴It also seems possible (depending, of course, on what the correct theory of rights is) that an agent may find himself in a situation where he has to violate someone's rights if he is to do the morally correct thing. This is, perhaps, yet another reason why, unless Taurek is willing to offer a convincing and comprehensive theory of rights to go along with his arguments, he might be better off if he stays away from basing his conclusions on the contentious question of whether or not someone's rights are violated if they don't get the life-saving drug.

something morally impermissible if he did P. This appears to be precisely the case with the action of drinking the drug for its good taste and letting all six people who need the drug in order to survive die. Any theory of rights, even if not fully libertarian, that respects some notion of private property, would probably say that the agent has the right to withhold the drug for himself even if he doesn't really need it. However, I don't see how it can be morally permissible to drink the drug for its good taste and let innocent people die.

The way for Taurek to keep his first argument on this view of rights and moral permissibility is to back his judgments about the First Person and Friend cases by nothing but intuitions about what is and what is not morally permissible, unless, of course, Taurek can find new arguments that are different from the ones criticized by Kavka. After all, few would disagree with its being morally permissible to save someone you know and like rather than five strangers. And almost no one, if anyone at all, would disagree with the intuition that it is morally permissible to take the drug to save one's own life rather than give it away to save the lives of five strangers. Someone who doesn't share that intuition, however, is in a position to pretty much completely disregard all of Taurek's first argument. But such people would have been in a position to do that even with regard to the original version of the argument. That's one reason why on the whole I think that Taurek's second argument is much stronger and much more difficult to refute than his first one - the former isn't founded solely on intuitions and controversial principles that are very easy to just

deny at the outset.

The way I have just described is, perhaps, not the only way to fix Taurek's first argument so that it doesn't need to rely on (LAP) and (O) at the same time (in fact, it's possible, if not likely, that neither one needs to be accepted), but it seems to me to be a good way. My suggestion also seems to be consistent with Taurek's expressed views, at least insofar as that is possible, as Kavka does seem to have latched on to some contradictions in Taurek's position. Be that as it may, any defender of the standard view of the role of numbers in morality still has a modified and seemingly consistent version of Taurek's first argument to contend with, as well as his second argument, which, I think, is where the really important issues lie anyway.

Subsection 2: The Transitivity Objection.

The second part of Kavka's paper focuses on Taurek's second argument, as well as the compatibility of Taurek's main conclusion with certain widely accepted principles. Kavka's analogy between morality and prudence does not, I believe, deserve any additional attention from us. It adds nothing to Parfit's analogy between the moral summing of different people's suffering and the moral summing of the suffering endured by one person at different times. And that is an analogy which, I think has been sufficiently discussed in section III and subsequently rejected.

Kavka's remaining objection, however, deserves much closer attention. The objection is aimed at the heart of Taurek's main conclusion. Kavka argues against the claim that the agent in Six

Strangers should be morally indifferent between saving the one and saving the five by offering a version of Six Strangers that is slightly modified as far as the options that the agent has, call it case Six' -

Six': "Let us name the six strangers, calling the one who needs all the drug Sam, and the others Mel, Tim, Art, Cal, and Len. Let A1 be the act of giving the whole dose to Sam, and A2 be the act of giving one-fifth of it to each of the others ... [also], he [the agent] could, for example, give one-fifth of the drug to each of Mel, Tim, Art, and Cal, and pour the rest on the ground. Call this act A3."⁵⁵

We need to keep in mind that on Taurek's view we should be morally indifferent between A1 and A2. Now Kavka invites us to compare A1 and A3. Kavka believes that according to Taurek we should be morally indifferent between A1 and A3 as well. That is a claim to which I will come back later. But, assuming we grant Kavka's contention, the point is that "it is a general presupposition of rational choice that indifference (i.e. the absence of preference) is a transitive relation. It follows, therefore, that we should also be indifferent between A2 and A3."⁵⁶ But this conclusion is completely unacceptable as it implies that it is a matter of no moral consequence, should we choose to save Mel, Tim, Art, and Cal, whether or not we save Len as well. As Kavka correctly remarks, this conclusion would be especially hard to sell to Len. Thus, Kavka concludes, Taurek's main thesis is in contradiction with an "accepted general principle of rational choice", and thus should be rejected.

⁵⁵Kavka, p.241.

⁵⁶Kavka, p.241.

Kavka is correct in pointing out that Taurek can't escape this difficulty by simply appealing to the fact that he is not interested in cases involving overlapping groups of people. After all, if his theory regarding non-overlapping groups of people has contradictory or absurd consequences for situations involving overlapping groups, Taurek is in trouble whether or not his original interests included overlapping groups. Kavka also maintains that Taurek is not in a position to rule out A3 as a legitimate option on a par with A1 and A2, even though A3, unlike A1 or A2, involves the wasting of a potentially life-saving portion of the drug on a whim. Taurek, however, definitely seems to believe in some sort of a non-arbitrariness requirement, i.e. that, at least in situations that are as important to the people involved as the situation in Six' is, the agent ought not act arbitrarily. For instance, when he discusses the Six Strangers case, Taurek expresses concern that "any preference I might show, therefore, if it is not to be thought arbitrary, would require grounding"⁵⁷, which is why he concludes that he must flip a coin.⁵⁸

⁵⁷Taurek, p.221.

⁵⁸Ralph Wedgwood and a few others have objected that flipping a coin is an action that's as arbitrary as an action can get. In reply, I want to suggest that the disagreement is merely terminological. What I, and presumably Taurek as well, mean by saying that an action was arbitrary is that there were no moral reasons to prefer it to its alternatives, and that it was selected via an illegitimate procedure, in the sense which includes, but is not limited to (I do not have a precise definition in mind), the procedure's being unfair. For instance, selecting who to save based merely on the race, religious affiliation, or height of those involved constitutes, in my idiolect, an illegitimate procedure; any decision derived solely from such a procedure would be arbitrary, as would be the resulting action. Taurek argues that the size of the group that shares the threatening situation with a person should be on the list with race, religion, and

Pouring out one fifth of the drug most definitely would, according to Taurek, constitute an arbitrary, ungrounded action. It is, however, difficult to see how Taurek would incorporate some sort of a non-arbitrariness principle into his, on the whole seemingly libertarian, framework⁵⁹. So, perhaps Kavka is right when he says that Taurek is hard put to reject A3 as a legitimate option. I, on the other hand, as a provisional defender of Taurek's main conclusion who is not, however, bound to the defense of the rest of his views, seem to be in a very good position to reject A3 as an illegitimate alternative, since I have shown that there is a way to retain a slightly slimmed down version of Taurek's first argument, not to mention his second argument, without accepting both (LAP) and (O). There is no reason why some sort of a non-arbitrariness principle cannot be part of the kind of framework that I've suggested a Taurek defender can accept. In fact, I would be surprised if some sort of a principle like that were not part of such a framework. Thus I definitely may be in a position to reject A3 on the spot as a morally impermissible alternative.

However, there are difficulties. Kavka might object that I owe him the details of the alternative framework, as well as an explicit formulation of the non-arbitrariness principle I've been speaking of; and supplying an adequate response to this objection

height. There is, however, it seems to me, nothing arbitrary in that way about using a random and fair procedure, such as a coin toss, to choose between morally equivalent alternatives.

⁵⁹For a discussion of such a principle and where the adoption of it would lead Taurek, see section V.

might prove to be a very difficult and unrewarding task. Also, I can't completely rule out the possibility that a more complex example can be constructed in place of Six', in which the counterpart of A3 will not be so easy to rule out. So, to take the simple way out of this conundrum, I will simply reject Kavka's objection after showing that the principle of transitivity of moral indifference, if applied in the way that Kavka applies it, also leads to counterintuitive and, in fact, seemingly absurd consequences unrelated to the Taurek controversy.

Consider the Migraine case -

Migraine: I am the only person in town who is both in possession of medications and capable of administering them safely. I can save either Mary's or John's life by delivering a dose of a drug to one of them. I only have one dose of the drug and cannot obtain any more in time to save both John and Mary. John and Mary live on opposite sides of town. If I choose to save John, I will be located close enough to Mike to be able to give him a pill that will immediately relieve his migraine headache. If I do not give Mike the pill, his migraine will go away in ten minutes anyway, and there will be no serious ill effects, either in the long or in the short run, of Mike's having had the migraine for an additional ten minutes. There are, of course, no special considerations involved.

Let M1 be the option of saving Mary's life, M2 - the option of saving John's life and relieving Mike's migraine, and M3 - the option of saving John's life and then going home without bothering with Mike's migraine. In the Migraine case, M3 is definitely a legitimate option, even if some sort of a non-arbitrariness principle is endorsed. It may be the case that if it costs me virtually nothing (say, a five minute walk) to save someone's life, then, even though I have the right to decline the

opportunity, I really ought to save the person's life. But it would be outrageous to claim that, very special considerations aside, I am under a similar kind of moral pressure to save someone from an extra ten minutes of a migraine headache.

It is obvious that one should be morally indifferent between M1 and M3. After all, in each case one life of a total stranger is saved, and neglecting to relieve someone's ten-minute migraine is, barring extraordinarily special circumstances, morally permissible. I also maintain that one should be morally indifferent between M1 and M2. The minor suffering inflicted by a ten minute migraine headache is not even close to being sufficient, on anybody's view I think, to tip the moral balance between saving John and saving Mary in John's favor. Even Mike, I believe, would agree with that. So, if we accept the transitivity principle, we should, according to Kavka, then also be morally indifferent between M2 and M3.

But that means that once I save John, there is no moral difference whatsoever between my going home and my assisting Mike. Granted, I am morally permitted to ignore Mike's quite mild and short-lived suffering, and would be performing a supererogatory act by assisting Mike, but can one really say that it is not morally preferable that I help Mike at basically no cost to myself rather than do nothing? Perhaps the case can be made even more convincing by increasing Mike's suffering considerably, while still making sure that it falls far short of death or permanent disability. But even as stated, I believe the Migraine case shows that the principle of transitivity of indifference cannot be

safely applied to moral decision-making in the way suggested by Kavka.

It occurs to me that some might doubt my claim that there is a moral reason to prefer M2 to M3, even though we should be morally indifferent between M1 and M2. After all, in both comparisons we have a death on each side and a migraine on only one of the sides. I believe that my claim is clearly true and would be surprised to see someone argue otherwise. However, for such skeptics I think that I can offer an even more convincing example by, so to speak, internalizing Mike. The discussion of this case will also suggest at least one possible way of analyzing such examples without having to outright abandon the principle of transitivity of moral indifference. Consider the case of

Broken Ribs: I am the only person who can save either John or Mary, but I cannot save both of them. There are, as usual, no special considerations involved. There are, however, two procedures by means of which I can save Mary's life. One of the procedures would result in the breaking of several of her ribs.

Let's call the option of saving John R1, the option of saving Mary without breaking her ribs R2, and the option of saving Mary while breaking her ribs R3.

Clearly, we should be morally indifferent between R1 and R2. But suppose R2 were unavailable for some reason. In that case, there seems no way to deny that we should be morally indifferent between R1 and R3. What would we say to try to convince Mary that we ought to prefer saving John to saving her? "Surely", we might say, "it is better to save a person's life without any damage to

his or her body than to save another person's life while inflicting an injury, even if it's not permanent or life-threatening." But that's not likely to be very convincing. Mary would probably reply that she wants to live as much as John does and has just as much of a reason for it as John, even if the option R2 is no longer available to her. If the only way to survive is to endure the pain of several broken ribs for a few weeks, then so be it! The price is not even close to being comparable to the potential gain.

In this case, unlike Migraine, there isn't a third person to even suggest that the moral considerations on one side can possibly outweigh the ones on the other in comparing R1 and R3. If we apply the principle of transitivity of moral indifference the way Kavka appears to be applying it, it seems that we should be morally indifferent between R2 and R3. But can we be? How can R2 not be morally preferable to R3? If in doubt, just ask Mary whether or not she prefers the option that would allow her not to suffer through several weeks of pain and limited mobility because of multiple broken ribs.

It looks as if the principle of transitivity of moral indifference is in trouble. But it seems grossly counterintuitive to abandon the principle of transitivity of indifference for rational choice. I see at least two alternatives. The first is to argue that transitivity of indifference does hold for rational choice, but not for moral choice. After all, even though both relations in question are called "indifference", no argument has been given either for the claim that the two are, in fact, the

same relation (i.e., to be more precise, that one is a subset of the other), or for the claim that at least they must have similar properties. This alternative, however, is hardly very intuitive either. Also, examples somewhat similar to my Migraine and Broken Ribs cases have been brought up in discussions of rational non-moral choice as well.⁶⁰ Thus it is, after all, likely that the principle of transitivity of indifference for rational choice and the principle of transitivity of moral indifference stand or fall together in the face of the objection stemming from my examples and others like them.

The second option is to define the preferences and absences of preference more finely with respect to the other alternatives that are available. For example, in Broken Ribs one might say that we should only be morally indifferent between R1 and R3 if R2 is unavailable. If, on the other hand, R2 is available, then we should prefer R1 to R3 and thus only choose between R1 and R2. Similarly, perhaps we should only be morally indifferent between M1 and M3 if M2 is unavailable (although that might be a little bit harder to swallow). Otherwise, we should prefer M1 to M3 and thus only choose between M1 and M2. If so, the principle of transitivity of moral indifference is not endangered by my examples, and neither is the principle of transitivity of

⁶⁰Such examples have come up, for instance, in the context of discussions regarding incomparability and incommensurability. See, for example, Ruth Chang's discussion of "arguments from small improvements" in her Introduction to *Incommensurability, Incomparability and Practical Reason* (Harvard University Press, 1997); although those examples were not intended to directly target either transitivity of indifference or transitivity of the relation "better than".

indifference for rational non-moral choice. But then, of course, this very same strategy, applied to the Six' case, invalidates Kavka's objection to Taurek's conclusion, as Taurek is no longer committed to maintaining that one should be morally indifferent between A1 and A3 in the presence of A2.

There may be, and probably are, at least a few other strategies that Kavka can resort to in defending the principle of transitivity of moral indifference from my objection.⁶¹ But it seems clear that any such strategy can subsequently be used to invalidate his own objection against Taurek. Thus Kavka has failed to discredit Taurek's non-standard view of the role of numbers in morality. We now move on to section V in which, while I believe that I will still be able to defend Taurek's main conclusion, his first argument, at least if it is to be an argument for precisely the conclusion he must have had in mind, will meet its untimely but well-deserved demise.

Section V.

Arbitrariness, Restrictions on Knowledge and Other Objections by Ken Dickey.

In his article "Why I May Count the Numbers", Ken Dickey

⁶¹I certainly don't see the Migraine and Broken Ribs cases as counterexamples to the principle of transitivity of moral indifference, but only as problems to be overcome by modifying the principle while retaining the general idea behind it.

offers an objection to Taurek's views which, I will argue, is entirely ineffective and unsuccessful. In the process, however, without I think realizing it at all, Ken Dickey helps strike down Taurek's first argument for the conclusion that all other things being equal, it is morally permissible to save one person rather than five. At the very least, especially in light of our discussions of the first argument in preceding sections, it appears that Taurek has to abandon the first argument unless he is willing to make his main conclusion much weaker and much less plausible than he wants it to be.

Subsection 1: Arbitrariness and the Inadvertent Demise of Taurek's First Argument.

As a consequence of the Non-Relativity Principle which I discuss in section II⁶², as well as Taurek's judgments regarding the First Person and Friend cases, we arrive at the conclusion that Taurek should believe that in Six Strangers it is not only morally permissible for the agent to give the drug to the one (which is understandably the part of the conclusion that Taurek concentrates on since it's the part that directly challenges the standard view

⁶²A reminder:

Non-Relativity Principle: "if it would be morally permissible for B to choose to spare himself a certain loss, H, instead of sparing another person, C, a loss, H', in a situation where he cannot spare C and himself as well, then it must be permissible for someone else, not under any relevant obligations to the contrary, to take B's perspective, that is, to choose to secure the outcome most favorable to B instead of the outcome most favorable to C, if he cannot secure what would be best for each".

of the role of numbers in morality), but also to give the drug to the five. So, Dickey asks, why can't the agent just act on his non-agent-relative permission to save the five simply because he likes the idea of saving the greater number of people? Taurek definitely seems to think that to do so, rather than to flip a coin, or perform some probabilistic equivalent thereof, and thus give everyone involved an equal chance for survival, would not be the right thing to do. Assuming that Taurek means here that it would be morally wrong not to flip a coin in Six Strangers, rather than that it's simply a worse alternative or just that he personally doesn't approve of it, he seems to be contradicting his own Non-Relativity Principle.

Here is what Dickey has to say: "I believe the best way of characterizing Taurek's position to avoid this dilemma is to say that, according to Taurek, while I have a non-agent-relative permission to save either the one or the five in Case 3 [Six Strangers], I should not act on this non-agent-relative permission in an arbitrary manner"⁶³. I concur. Furthermore, I agree with Dickey that Taurek most definitely believes that special considerations aside, deciding who to save based solely on the relative numbers of people involved is arbitrary. Based on what we already know about Taurek's views, Dickey formulates the following

⁶³Ken Dickey "Why I May Count the Numbers", ETHICS Problems and Principles, ed. by John Martin Fischer and Mark Ravizza, Harcourt Brace Jovanovich College Publishers 1992, p.245; All future quotations from Dickey are from this article.

Arbitrariness Principle (AP): "In conflict cases of the type under consideration, I should not act on non-agent-relative permissions in an arbitrary manner, where "arbitrary manner" (i) includes acting on the basis of the relative numbers of people involved, in itself, as a significant factor in making a decision, and (ii) does not include acting in order to save someone I know and like"⁶⁴.

Of course, Dickey should also add that one does not act arbitrarily in saving oneself. And, in addition, clause (ii) is quite controversial. In fact, Parfit doesn't really agree with it and I am not so sure that I do either⁶⁵. But I certainly think that Dickey correctly captures the relevant parts of what Taurek has in mind. I also want to point out that we probably want to endorse some sort of a non-arbitrariness principle regardless of the Taurek controversy, as presumably in all kinds of tradeoff situations we do not want to allow for choices based, for instance, on the race, ethnicity, hair color, eye color, or height of the people involved.⁶⁶

There is another arbitrariness-related issue left to

⁶⁴Dickey, p.246.

⁶⁵See section VII for a little bit more on that topic.

⁶⁶Some may object to adopting any kind of a non-arbitrariness principle and claim that if there is no moral reason to prefer one alternative to another, the agent is morally permitted to pick whichever alternative pleases him more for any reason, even if the preference is based on a complete whim. In particular, then, on this assumption, Taurek's claim which I will be discussing in a few paragraphs is wrong and an agent would be morally permitted to give the drug to the five outright in Six Strangers. But that would still not mean that numbers are a factor relevant to morality; just as even if we refused to adopt some form of a non-arbitrariness principle about race, and thus assumed that a white racist was morally free, all other things being equal, to save a white person over a black one for no reason other than the colors of their skins, that would still in no way imply that race is a morally relevant factor.

consider. Taurek is very concerned about the losses to persons, even though he is not concerned about the losses of persons. Even though he does not believe that he is morally required to save a stranger's life rather than another stranger's arm, he clearly does not think that it would be arbitrary for him to do so, given that the loss of one's life is a much greater loss to a person than the loss of one's arm. In general, Taurek definitely does not claim that it is arbitrary to act so as to prevent the greatest amount of loss to any one person, at the very least when that loss is considerably greater than any loss that we could prevent someone from suffering by acting otherwise. Hence Dickey's

Modified Arbitrariness Principle (MAP): "In conflict cases of the type under consideration, I should not act on non-agent-relative permissions in an arbitrary manner, where "arbitrary manner" (i) includes acting on the basis of the relative numbers of people involved, in itself, as a significant factor in making a decision, (ii) does not include acting in order to save someone I know and like, and (iii) does not include acting so as to prevent the greatest amount of loss to any one person"⁶⁷.

Dickey is now ready to present his argument against Taurek. But before I address that argument, I want to mention a major consequence of this discussion that Dickey apparently neither intended nor noticed.

On several occasions in the course of his discussion Dickey draws our attention to the word "should" in Taurek's claim that one *should* not act arbitrarily on one's non-agent-relative

⁶⁷Dickey, p.247.

permissions. What kind of force is the "should" supposed to have here? I would suggest that Taurek intends it to have the strongest possible force - that of "ought". The reason I believe this to be the case is that the alternatives seem to be too weak for Taurek's purposes. For suppose that all he means is at best⁶⁸ that flipping a coin is morally preferable to just acting on one of the agent's two non-agent-relative permissions without flipping a coin first, but that all three alternatives are morally permissible. That would basically mean that while there is no moral reason to prefer saving five people to saving one person, which would already be a major victory for anyone opposed to the standard view of the role of numbers in morality, whether or not to count the numbers is up to the individual. Perhaps, it is morally preferable to flip a coin. But if one happens to have a preference for saving larger numbers of people, he or she is morally permitted to act on that preference without flipping a coin.

It is important to note that this would then also have to be true about the opposite preference. If one happens to have a taste for saving smaller numbers of people, he is morally permitted to act on that preference as well. Even this conclusion would defeat the standard view that the larger number ought to be saved. But I believe that this kind of a conclusion is by far too weak for what Taurek has in mind here. And, in any case, I do not

⁶⁸As I said before, the "should" could mean even less - it could, conceivably (but just barely so), simply be expressing Taurek's personal preferences.

find it in the least bit plausible that such an important moral decision as whether or not to count the numbers in tradeoff situations, not to mention which way to count them if they are counted, is left up to the individual tastes of agents by the true theory of morality. Thus I am going to assume that the "should" is intended to have the force of "ought", and that in Six Strangers Taurek maintains that the agent ought to flip a coin, or perform some probabilistic equivalent, in order to determine who to save.

But under this interpretation Taurek has a major problem. In First Person, as well as in Friend, the agent is supposedly morally permitted to forgo a coin toss and give the drug to the one rather than the five. Presumably, that is because of the special relationship between the agent and the one person in the two cases. On the other hand, in Six Strangers the agent is not morally permitted to forgo the coin toss and give the drug to the one outright, even though Taurek's judgments in First Person and Friend imply the opposite when conjoined with the Non-Relativity Principle. Presumably, in fact it appears to be the only explanation that Taurek can consistently give, that is because there is no special relationship between the agent and the one person in this case, and so it would be arbitrary for the agent to use his non-agent-relative permission to save the one, without flipping a coin first to determine who to save. But the situation which I have just described happens to be a paradigm of an agent-relative permission, because there is an action A, namely forgoing the coin toss and just giving the drug to the one, that an agent

is morally permitted to engage in only if he has a special relationship to the one. But Taurek doesn't believe in agent-relative permissions. In fact, his entire first argument relies heavily on denying the existence of agent-relative permissions⁶⁹.

If we continue to reject the weak interpretation of "should", as I believe we have no choice but to do, Taurek has only two options: (1) he can admit that agent-relative permissions exist after all, and so give up the Non-Relativity Principle, or (2) he can reject his initial judgments regarding the First Person and Friend cases, which, considering his conclusion in Six Strangers, would seem to indicate that the agent in those two cases ought to flip a coin as well. Regardless of which option Taurek chooses, his first argument is now deceased if he wants it to be an argument for the strong and at least minimally plausible conclusion that I believe he had set out to argue for^{70 71}.

⁶⁹Accepting agent-relative permissions undermines the Non-Relativity Principle, without which Taurek's first argument cannot even get off the ground.

⁷⁰I stand by the grave concerns that I have expressed in section III regarding the notion of agent-relative permissions, at least as presented by Parfit. However, I do find reversing Taurek's judgment in First Person extremely disturbing and counterintuitive. So, if I am presented with a better argument for some version of agent-relative permissions, that hopefully addresses the difficulties discussed in section III, I will, despite some serious reservations, be more than happy to accept agent-relative permissions in some form.

One other possibility, which, perhaps, can be worked on, is that saving yourself outright at the expense of saving five others in First Person is as morally impermissible as saving the one in Six Strangers outright, but saving yourself in First Person is less blameworthy, i.e. there are no agent-relative permissions as such, but there is agent-relative blameworthiness. If this could be made to work, we would, so to speak, be able to have it both ways, at least to some extent. But this option seems to have more plausibility in the case of my arm vs. a stranger's life, than in a situation where my life is

Ironically, it is not, as Parfit would have us believe, Taurek's denial of agent-relative permissions that brings down the argument, but his inadvertent and hitherto unnoticed acceptance of them. But, as far as I am concerned, good riddance! I believed from the outset that the first argument was a lot more trouble than it was worth, even though a few good discussions that are valuable for their own sake were produced by it. And the second argument, which, as I have claimed before, is a much stronger, interesting and illuminating one, still remains to challenge the standard view of the role of numbers in morality.

Subsection 2: The Argument from Probabilities.

We should now consider Dickey's actual objection against Taurek's conclusion. Dickey suggests at the beginning of his paper that "Taurek's position that the numbers should not count is intended (among other things) to serve as a guide for private citizens in making decisions in real-life situations..."⁷². I see no reason to disagree. Now, having formulated (MAP), Dickey asks whether it is really true that, even in light of clause (iii), there is still no non-arbitrary reason to prefer saving five people in Six

at stake.

⁷¹Trying to rebuild Taurek's first argument by starting with the claim that in First Person and Friend it is morally permissible to flip a coin, although not necessarily to give the drug to the one outright, seems to be a non-starter - while this could perhaps be the conclusion of a very complicated argument, this is hardly a sufficiently strong and plausible intuition to serve as the initial premise of an argument.

⁷²Dickey, p.244.

Strangers. However, since he believes that the thesis

(T): The amount of loss to a person who dies is the same for all persons who die the same type of death.⁷³

is at least somewhat plausible, and he does not wish to be sidetracked by arguing about it in his paper, Dickey prefers to discuss an example similar to Six Strangers which, however, does not involve potential deaths. He offers for our consideration the following variation, let's call it

Six Strangers': "I have a supply of some right-arm-saving drug. Six people will certainly lose their right arms if not treated with the drug. One requires all of the drug if he (or she) is to retain his (or her) right arm. Each of the other five requires only one-fifth of the drug. I do not know any of the six involved and have no special obligations to any of them. What ought I to do?"⁷⁴

It is Dickey's declared goal to prove, by means of Six Strangers', that "Taurek's position that the numbers should not count is inconsistent when applied to cases in which harms other than deaths are involved, *if interpreted as an action-guiding principle for private citizens in real-life situations*"⁷⁵. He attempts to show this by arguing that even in accordance with Taurek's rules of engagement, there are non-arbitrary reasons for preferring to help the five strangers instead of the one in Six Strangers'. But, in order for his argument to proceed, he needs to assume that he doesn't have all the relevant information about the six people involved. This is a crucial assumption and,

⁷³Dickey, p.247.

⁷⁴Dickey, p.247.

⁷⁵Dickey, p.244.

because he cannot avoid making it if the argument is to have a chance to work, that argument cannot possibly show that Taurek is wrong in maintaining that we ought to flip a coin in cases where we know that all the strangers involved stand to lose equally if not helped. The most Dickey can hope to achieve with this line of reasoning, is to show that if we do not know exactly how much each person stands to lose, but only know that all of the potential losses fall under a category of harm with the same generic description, such as "the loss of a right arm" (which is likely to be the case in most real-life situations), there are non-arbitrary reasons for not flipping a coin and just helping the larger group of strangers; thus, Dickey hopes to conclude, saving the larger group outright is, at the very least, morally permissible.

Such a claim is interesting in itself and perhaps somewhat unpleasant for Taurek as well, since I agree with Dickey that Taurek probably intends his conclusion to be applicable as a "practical action-guiding principle in real-life situations". However, it is a well known truth about philosophy that demonstrating that it is difficult, or even practically impossible, to apply a philosophical principle in real-life situations, is usually insufficient to prove that principle false. That is the case here as well: nothing that Dickey can say under the assumption of such a serious epistemic limitation can refute Taurek's main claim that special considerations aside, when the potential losses to all the strangers involved are equal and the agent knows all the relevant facts, the agent should flip a coin to decide between the two groups of strangers regardless of the

relative sizes of the groups. If one concedes even this much to Taurek, one denies the validity of the standard view of the role of numbers in morality. Dickey is, of course, aware of this and would likely be the first to admit it, but I think that it needed to be explicitly remarked and kept in mind throughout the discussion that is to follow.

So now that all the preliminaries are out of the way, let's grant Dickey his epistemic assumption and see his argument through. It presumably falls within the scope of the epistemic constraint which we have accepted that we do not know about any one of the strangers involved whether or not he or she is a right-handed pitcher in Major League Baseball. But, Dickey claims, the loss of a right arm would be a greater loss for a right-handed pitcher in Major League Baseball than it would be for most other people. And, he says, however small the probability, P , of any given stranger's turning out to be a right-handed pitcher in Major League Baseball is, it is surely greater than zero. And so, the probability of it being the case that at least one of a group of five strangers turns out to be a right-handed pitcher in Major League Baseball, $5P$, is greater than P . Of course, part of what Dickey says here is false, as the probability of there being at least one right-handed pitcher in Major League Baseball in a group of five people is, under the assumptions we have made, not $5P$, but $1-(1-P)^5$. This, however, makes no difference to the argument, as $1-(1-P)^5$ is greater than P for P between 0 and 1.

Thus, Dickey concludes, if he helps the group of five strangers, he decreases the probability of an outcome in which at

least one of the people he does not help, and who thereby loses his or her right arm, is a right-handed pitcher in Major League Baseball and thus suffers a greater than average loss associated with losing a right arm. He adds, "if I am, as Taurek seems to be, seriously concerned with preventing the greatest amount of loss to any one individual, it would be rational in Case 3' [Six Strangers'] to act so as to maximize the probability that I am acting so as to prevent the greatest amount of loss to any individual ... since I can see nothing morally objectionable about acting on relative probabilities in this way, I conclude it would not be arbitrary, according to clause (iii) of (MAP), for me to act in Case 3' [Six Strangers'] so as to save the five - without first flipping a coin"⁷⁶. But that implies that clause (iii) of (MAP) contradicts clause (i) of (MAP), leading, Dickey hopes, to his intended conclusion that Taurek's position is inconsistent if it is evaluated as an action-guiding principle for real-life situations. The argument, Dickey claims in his footnote 14, can be generalized to encompass any generic category of harm.

Unfortunately for Dickey, his argument is deeply flawed in at least several ways. For one thing, even if we grant Dickey that the amount of loss involved in losing a right arm is greater for a right-handed pitcher in Major League Baseball than for almost anyone else, what is it that entitles him to claim that that loss is the greatest loss possible under the description "losing one's right arm"⁷⁷? There isn't even a guarantee that there is a maximal

⁷⁶Dickey, p.248.

possible loss to a person associated with the loss of a right arm. Also, what exactly entitles Dickey to conclude that the loss of a right arm is a greater loss for a right handed pitcher in Major League Baseball than for almost anyone else. It seems to me like a very presumptuous conjecture on Dickey's part. This, however, is an issue that it is very difficult and probably quite fruitless to argue about. So I will grant Dickey this assumption as well. But I still maintain that even if his claim is correct, the differential between the losses is very small as compared to the loss itself.

I now want to draw attention to the precise formulation of clause (iii) of (MAP). After all, if the charge against Taurek is inconsistency, we have to make sure that the clause is stated in a way that Taurek can accept. But I think it is clear from reading Taurek that he would not accept clause (iii) unless it specified that the differential between the losses being considered should be rather large (probably quite significant in comparison with the losses themselves) in order for a preference based on it to be non-arbitrary. For instance, Taurek is unlikely to consider the difference between B's losing an arm and, in addition to that, suffering a minor headache, and C's just losing an arm, significant enough to ground a non-arbitrary preference for sparing B his loss rather than sparing C his loss. In addition, a

⁷⁷For instance, if we stipulate that the loss of a right arm is more of a loss for a concert pianist than even for a right-handed pitcher in Major League Baseball, shouldn't Dickey also worry about how his suggested course of action fares as far as minimizing the probability of a scenario in which a concert pianist loses his or her right arm?

similar objection applies, I think, to the probabilities. Considering the exceedingly small (enough to be considered negligible in everyday life) probability, P , of a random stranger's turning out to be a right-handed pitcher in Major League Baseball, the difference between $1-(1-P)^5$ and P is unlikely, as far as Taurek is concerned, to be sufficiently large to ground a non-arbitrary preference for saving the group of five people in Six Strangers'. Thus I would conclude that Dickey has failed to bring out any inconsistency in Taurek's views by examining the Six Strangers' case.

But that's not all. I believe there is yet another way to show that Dickey's argument doesn't work. In order to do that I will run an argument similar to his, but for the opposite conclusion. There is a probability C , slightly greater than zero, that any given stranger selected at random will develop a cancerous tumor in his or her right arm that will eventually prove fatal. In fact, that probability is quite likely to be greater than the probability of that stranger's being a right-handed pitcher in Major League Baseball. But then the probability of it being the case that at least one of a group of five strangers will at some point develop such a cancerous tumor, $1-(1-C)^5$, is greater than C . By reasoning identical to that of Dickey, I have a non-arbitrary reason for saving the one stranger instead of the five, since by doing so I would minimize the probability of the outcome in which at least one of the people involved suffers a much greater than average loss associated with keeping his or her right arm. I see no way for Dickey to reject my argument without

abandoning his. If his argument can be generalized, as he suggests in his footnote 14, so can mine.

Dickey might respond: "That's fine with me. The conclusion of your argument, as counterintuitive and ridiculous as it, and the reasoning behind it, sound, still shows that it is morally permissible to count the numbers in cases such as Six Strangers', and that Taurek is wrong (at least if he wants his conclusion to serve as an action-guiding principle in real-life situations)". I disagree. I think all that the conclusion of my mock argument shows is that Dickey's argument is fatally flawed and cannot be taken seriously. Once Dickey allows probabilities of outcomes, no matter how insignificant, to seriously affect what is or is not morally permitted in a particular case, he opens a Pandora's box and all kinds of insane consequences follow. As can be seen by comparing my argument with his, letting even negligible probabilities play a role allows us, and, in fact, forces us, to see each potential loss as also a possible benefit and vice versa, thus giving us a non-arbitrary reason to avoid a course of action every time we have a non-arbitrary reason to prefer it (and, of course, vice versa). That is hardly a desirable quality in a modification to an action-guiding principle for real-life situations! When this difficulty is combined with my other objections to Dickey's argument, not to mention the fact that denying the applicability of Taurek's conclusion to real-life situations doesn't even begin to discredit it as a true principle, I think it is safe to conclude that Dickey's argument is not a threat to Taurek's views as a whole.

Section VI.

The Problem of Fairness.

Once Taurek reaches the conclusion that it is not in any way better to save the five strangers rather than the one in Six Strangers (or, to put it differently, that there is no non-arbitrary reason for preferring saving the five to saving the one), he adds that the agent should flip a coin, as this is the fairest and most reasonable thing to do. It does, indeed, seem both fair and reasonable to flip a coin in order to decide between two alternatives, neither of which we have a non-arbitrary reason to prefer. But even this step has not been allowed to pass without some controversy. In her book *Morality, Mortality, Volume I*, Frances Kamm argues that there are procedures other than a coin toss, or some probabilistic equivalent of a coin toss, that perhaps qualify for being fair. Employing any one of those procedures would amount to allowing the numbers to count in some way. John Broome, in his article "Kamm on Fairness", disagrees with Kamm and concludes that Taurek is right in that the only fair procedure in a case like Six Strangers is a coin toss or some probabilistic equivalent. However, he adds, "fairness is not everything. Fairness requires tossing a coin ... but the fairness of tossing a coin is outweighed by the expected badness of the result. Tossing a coin will lead you to save three lives on average (the expectation of lives saved is three), whereas you could save five for sure. Two lives are worth some unfairness, I should say. Therefore, I think you should save the five without

more ado."⁷⁸

Kamm disagrees that fairness should be overridden by considerations of goodness. In her reply to Broome, she says that if Broome were right, "why then not kill one to save two, overriding justice?"⁷⁹ I believe that Broome's view is derived from a notion of fairness that is too narrow. He says, "when claims are equal, I believe fairness requires them to be equally satisfied, and that is all it requires. It does not require any of them to be satisfied to any particular degree ... if a person has a chance of having her life saved, that provides a sort of surrogate satisfaction of her claim to have her life saved. Equal chances provide a surrogate equality in satisfaction, and so a degree of fairness."⁸⁰ Broome also maintains that in a case where there is no way to save everyone, the fairest thing to do is to save no one, although we should, of course, save at least someone out of considerations of goodness.

I would like to suggest that perhaps the notion of fairness should be somewhat more robust. Perhaps helping no one is, in fact, unfair to everyone, instead of being unfair to no one. With a more robust notion of fairness it might turn out to be the case that not only, as Kamm thinks, you aren't morally required to

⁷⁸John Broome, "Kamm on Fairness", *Philosophy and Phenomenological Research*, vol. LVIII, No.4, December 1998, p.955; All quotations from Broome are from this article.

⁷⁹Frances Kamm, "Reply to Critics", *Philosophy and Phenomenological Research*, vol. LVIII, No.4, December 1998, p.972.

⁸⁰Broome, p.956.

sometimes allow considerations of goodness to override considerations of fairness, but that, in fact, fairness and goodness cannot ever conflict. And there is every reason to believe that even on a much more demanding interpretation of what fairness requires, it would still be the case that flipping a coin in Six Strangers is fair.⁸¹ But this is no more than a suggestion and a vague hope based on an even vaguer intuition (plus some sympathy with Kamm's refusal to accept the overriding of fairness).

I don't have any specifics in mind, however, and even if I did, a long discussion would result that would take us far off the topic and wouldn't really be worth the trouble. Fortunately, there is no need to engage in that discussion as there is no reason, in light of what's been said in the previous five sections, to agree with Broome (or Kamm, for that matter) that considerations of goodness are on the side of saving the five rather than the one. Neither Broome nor Kamm, at least in the process of discussing fairness, offer any new grounds for rejecting Taurek's second argument, and so I see no need to revive this issue⁸². Thus the only question that remains for us to

⁸¹Another possibility which would allow goodness and fairness not to conflict is, of course, that fairness is less, not more, demanding than Broome thinks. In that case, many options would perhaps be fair, including, proponents of the standard view of the role of numbers in morality should hope, the option of saving the larger group without performing a coin toss or some probabilistic equivalent. This, though, is precisely the view that Kamm seems to advocate and that I side with Broome in rejecting later on in this section.

⁸²The only argument to that effect offered by Kamm in Morality, Mortality, volume I: Death and Whom to Save from It is centered around a fallacy which is somewhat related to the fallacy in Kamm's

consider is whether any of Kamm's suggestions for fair alternatives to flipping a coin can stand up to scrutiny.

First I will discuss the procedure Kamm dubs '*majority rule*'. It permits something Kamm refers to as '*substituting equivalents*'. She claims that there is an explanation of why the needs of the majority outweigh the needs of a minority, that doesn't rely on there being a suprapersonal point of view from which it is just plain better to do what is in the best interests of the majority. This is important to Kamm because such a suprapersonal point of view is something to which Taurek objects quite strongly, articulately and convincingly. This is where the substitution of equivalents comes in. Broome explains Kamm's argument in the following manner: "The claim of the one conflicts with the claims of the five. Taken individually, it must confront the claims of the five taken individually, and it is actually equivalent to the claim of just one among the five. Then there are the claims of the other four as well. So there are four uncountered claims on one side of the dilemma."⁸³ Thus Kamm argues that it is actually fair (and not just better, which is something she believes as

defense of '*substituting equivalents*', which I am about to discuss. In addition to other problems, both arguments appear to suffer from what Shelly Kagan calls "the additive fallacy", a more or less comprehensive discussion of which falls outside the scope of this paper, but can be found in the next chapter of my dissertation (see subsection 2 of section VI). As a result of the fallacious reasoning, Kamm falsely concludes that by virtue of believing that it is no worse if B and C both suffer the loss L than if only A suffers L, Taurek is committed to the claim that it is no worse if B and C both suffer L than if only B does, and then infers by modus tollens that the first belief must be false. For a related discussion of this issue from a slightly different angle, see also subsection 2 of section IV.

⁸³Broome, p.957.

well) to give the drug to the five simply because there are more of them. At the same time, she believes that since no suprapersonal point of view is presupposed and the individual claims are not aggregated, this procedure respects people as individuals, again offering something that should, in principle, appease Taurek.

Before proceeding any further, I will need to clarify something that will be important both now and for the discussion of the next procedure that Kamm suggests. The kind of "claim" being talked about here, even if it is, strictly speaking, a right (for which no argument has been given), is certainly not to be confused with the kinds of rights which are particularly binding in these kinds of cases, such as ownership rights, contractual rights, or rights of restitution. The assumption that the six strangers all have an equal claim to the drug is not to mean something along the lines of it being the case that all six are part-owners of the drug and thus have property rights to it, or along the lines of it being the case that the six all have a right to the drug as compensation for past wrongs. The kind of claim being discussed here is something a lot less tangible - something, perhaps, along the lines of each person's having a right to be considered as a candidate for having his or her life saved by anyone who has the means and the opportunity to save that life and is aware of the victim's predicament⁸⁴.

⁸⁴But people do not, of course, special considerations aside, have the right to actually be saved.

If Kamm just assumes that these kinds of claims are additive, i.e. that adding extra claims to an equivalent single claim can result in a stronger aggregate claim, the way Taurek admits certain kinds of rights may be additive when he discusses the Coast Guard examples⁸⁵, she begs the question. I believe, however, that she considers her argument not to be question begging. After all, she seems to be saying, she doesn't add anything up, she just points out that when the claim of the one cancels out, in a manner of speaking, the claim of one of the five, four other claims remain that have not been canceled out⁸⁶.

But this is just an illusion, from which I believe Kamm suffers because she finds Taurek's conclusion so implausible that she doesn't take it seriously enough. I would like at this time to offer an analogy to better articulate what I mean. Suppose that I am a beginner studying set theory and am entirely baffled by the purported fact that the cardinality of $w+1$ is not greater than that of w . Suppose I offer the following argument to dispute that fact: w is equivalent to the w part of $w+1$; and there is

⁸⁵In "Should the Numbers Count?", Taurek considers a possible objection to his view, that is based on the intuition that a Coast Guard captain ought not flip a coin to determine whether to save the larger or the smaller group of people from a life-threatening natural disaster. Taurek ends up agreeing that in some cases people actually have a right, akin to partial ownership, to a life-saving resource; and in such cases individual rights can, perhaps, be aggregated in a morally significant way, resulting in the captain's having a duty to save the larger group. Another way we can imagine such a duty arising is through an agreement, explicit or implied, as to how a resource is to be used in an emergency. Of course, neither of these is the case in Friend, First Person, or Six Strangers.

⁸⁶Also, Kamm does explicitly state that she is not aggregating the claims, which is central to her ability to maintain that the 'majority rule' procedure respects everyone involved as an individual.

still a 1 left over that is not "canceled out" (or countered, if we wish to put it that way) by anything in w . Therefore, I conclude that $w+1$ has greater cardinality than w . "Surely, I am not begging the question", I might add, "for I am not presupposing that adding 1 to w adds something to the cardinality of w , I am simply pointing out that you have an w on both sides but one of the sides also has a 1 left over in addition to that".

Nobody would be convinced by such an argument. I would probably be told that I must have been so perplexed by the initial counterintuitiveness of what I was arguing against, that I didn't quite understand the claim and had already unwittingly assumed the negation of what I was trying to deny. Kamm, I believe, makes the same mistake. There is, of course, no literal canceling out of the claims, just as there is no literal canceling out of the w 's. Therefore, it is impossible to justify any conclusion by talking about what is "left over". And since Kamm can't assume (for fear of blatantly begging the question) that the adding of four claims to one equivalent claim adds anything to that single claim that is of any significance either for the question of fairness or for the question of goodness, her argument disintegrates. Broome is right not to take it seriously.

The final procedure I will discuss is the procedure of *proportional chances*. The procedure concedes that the fair thing to do in the Six Strangers case (and thus, on Kamm's view, the right thing to do as well) is to give everyone involved a chance for survival. This would already be a major concession to Taurek and a defeat for the standard view of the role of numbers in

morality. However, Kamm argues that it is fair to roll a six-sided die instead of flipping a coin, thus giving each of the strangers a chance for survival proportional to the size of the group he or she is in. If this were, indeed, a fair procedure, the defenders of the standard view would win a partial victory, or, at the very least, a consolation prize, as well. Numbers would, in fact, count in morality, even though not in the way that the standard view advocates.

To support her contention, Kamm offers us the Lottery analogy. Suppose that six people have an equal chance to win a large sum of money. But five of them decide to organize a joint venture. They have a way of investing the money which is guaranteed to almost immediately increase their wealth by a factor of five. So they decide that if one of them wins, regardless of which one it is, they will invest the winnings in that manner and split the proceeds. Thus, in effect, if one of them wins, they all win. Kamm argues that it would be ridiculous and unfair to prohibit the five from doing that or to announce that the chances of winning should now be altered so that there is a 50% chance that the one will win. In this case, it is fair for the five to have a proportional chance of winning, even though each would in the end win as much, if anything at all, as the one would win. The same, Kamm says, applies to Six Strangers.

Before I discuss this analogy and the procedure of proportional chances as a whole, I have a little bone to pick with Kamm. The claims that people in a lottery have to equal chances to win, sound more like *bona fide* rights to me (akin to property

rights) than simple claims, since, after all, the players did purchase a lottery ticket. And such rights, as Taurek himself admits, may, perhaps, quite reasonably be aggregated⁸⁷. Taurek limits the scope of his conclusion to cases where none of the people involved have a right to the resource in question. Thus if Kamm hopes to defeat Taurek's conclusion, she has to stick to claims only and, so as not to repeat the mistake she made with the majority rule procedure, not assume at the outset that claims may be summed in a way that affects what is and what is not fair.

Broome disagrees with this procedure as well. He says "Kamm's view resembles a particular liberal theory about the distribution of wealth in society. Some liberals think that people ought to have an equal start in life, but that what they make of their opportunities is up to them, so there is no reason why wealth should end up equally distributed. Kamm's idea seems to be that people should have an equal baseline chance, but if they can improve their chances by pooling, that is their right; it is no reason to change the baseline distribution of chances."⁸⁸ Broome admits that Kamm's story may be correct for the case of the lottery, as do I. But in *Six Strangers*, Broome argues, there is no way to distinguish between baseline chances and final chances. He says: "It is not that you are initially able to save just one person out of six, and then somehow five of the six pool their

⁸⁷See Taurek's discussion of the Coast Guard captain examples in "Should the Numbers Count?". I briefly describe what that discussion is about in footnote 85.

⁸⁸Broome, p.960.

chances".⁸⁹ Thus, he claims, Kamm's argument for why it is fair for the six strangers to have uneven final chances for survival becomes baseless.

Kamm responds to such objections by offering the Drifting Icebergs case -

Drifting Icebergs: There are six strangers stranded on individual icebergs. I can only get to one of the icebergs in time to save its "inhabitant". But, luckily, five of the icebergs end up drifting close enough together so that I can save all five of the strangers that are stranded on those icebergs.

Kamm claims that this case is equivalent to her lottery example. Each of the six hapless polar explorers should get an equal baseline chance of one sixth for survival. But if nature ends up pooling the chances of five of them, so be it. It would not be fair to deprive them of their equal baseline chances. Broome disagrees. He responds that "people's position should be equalized *after* nature has done its business. People should be compensated for what nature does, since that is no responsibility of theirs. A person is only entitled to a favoured position if it results from what she herself has done."⁹⁰ In Drifting Icebergs, the five icebergs end up close enough together for the people stranded on them to all be rescued not as a result of something they had done to improve their situation. It is nature, or just mere chance, that helped them out. Thus it is unclear why those five should get to benefit more from the lucky change in their

⁸⁹Broome, p.960.

⁹⁰Broome, p.961.

situation than the remaining one. If, as Taurek would suggest, we flip a coin, we find that the chances of each of the six have improved equally - from one sixth to one half - which, after all, seems only fair.

Kamm could, of course, reply with one more modification. I will call it

Drifting Mad Scientists: There are six strangers stranded on individual icebergs. I can only get to one of the icebergs in time to save its "inhabitant". But, luckily, five of the six strangers are mad scientists, who somehow invent a way to make their five icebergs drift close enough together so that I can save all five of them at the same time.

In this case it is not nature but that is responsible for the pooling of the five's chances, but the five themselves. So is Kamm right after all, at least in this special kind of variation on Six Strangers? I think not. I believe there is an argument that defeats Kamm's conclusion for Drifting Mad Scientists and that can also serve as another reason to reject her verdict regarding Drifting Icebergs. For how are we to make sense of the claim that all six strangers involved in these cases have, or deserve to have, an equal baseline chance? It is true that if I find out that there are six strangers afloat on six different icebergs and I can save at most one of them, the fair procedure is for me to roll a six-sided die, giving each of the six an equal baseline chance of one sixth for survival. Kamm would now need to argue that just because five of the icebergs drifted together or, even better, were made to drift together by five of the six strangers, it does not become fair for me to deprive each of the

five of his or her equal baseline chance.

But one should not forget that even if Taurek's conclusion is meant to guide the actions of people in real-life situations, it is primarily a theoretical principle which requires that certain conditions obtain in order for it to be applicable. In particular, one of the conditions is that I cannot save more than one person or group, not simply that I don't, at the time, think that I can. If the condition does not, in fact, obtain, but I end up rolling the die anyway because I believe that it does, I do something wrong that results in an unfair outcome, even though I cannot be blamed for it since I didn't, and presumably couldn't, know any better.

Thus, even if I had already rolled a die, only to find out later that the chances of five of the strangers were pooled, either by nature or by their own efforts; and even if the roll had indicated that I should save one of the five (and so, in virtue of the pooling, all of the five), it would not be unfair for me to now flip a coin anyway because I had not, having acted on false information, performed a fair procedure the first time. In fact, it would be unfair of me not to correct myself, as the five never deserved an individual equal baseline chance of one sixth in the first place.⁹¹ Thus I concur with Broome that for cases such as Six Strangers, there is no justification for using the procedure

⁹¹Broome may, in fact, have in mind something along the lines of the argument I have just offered when he says "it is not that you are initially able to save just one person out of six, and then somehow five of the six pool their chances", (p.960).

of proportional chances either.⁹² I have to agree with Taurek and Broome that the only fair procedure in cases like Six Strangers is the only procedure that intuitively appears fair at first glance - the procedure which gives each group, and thus each member of each group, an equal chance.

Section VII.

Some Final Thoughts on the Subject.

Up to this point, I believe that I have successfully defended both Taurek's second argument and the consistency of Taurek's main conclusion with common sense, which was placed in doubt by Kavka's objection from transitivity as well as by Dickey's argument. Taurek's first argument eventually had to be abandoned, but not because there is necessarily anything wrong with the conclusion. The argument was flawed from the outset, as it relied on quite controversial intuitions and principles and, apparently, was based quite heavily on the ambiguities within Taurek's views on rights and moral obligations⁹³. Just to give another indication of how

⁹²The procedure of proportional chances may, however, be a fair procedure for cases like the original Coast Guard Captain example discussed by Taurek - an example in which each of the people in the two groups has an equal right to the use of the boat and there are no special facts about the population distribution on the island. For such cases, even the majority rule procedure might, perhaps, be argued to be somewhat fair, although I think the proportional chances procedure has a far better chance.

⁹³I suggested one way of clearing up some of these ambiguities in

fraught with difficulties some of the assumptions of the first argument are, I want to draw your attention one last time to the Friend case.

It is tempting to agree with Taurek's judgment that it is morally permissible in Friend to save David, who is not a close friend, but someone I know and like⁹⁴. But what of the opposite scenario? What if David is someone I know and dislike? Note that David is not my archenemy - he hasn't murdered my parents, or seduced my girlfriend, or even received the promotion that I so richly deserved; he is just someone I dislike for one reason or another, or perhaps even without a reason. Is it as intuitive that it is morally permissible to save the five in this case simply based on my dislike of David, as it is that it is morally permissible to save David in Friend based on my liking David⁹⁵? For the sake of consistency, it better be!

I think this should cast some doubt on the too easily accepted premise used in the discussion of the Friend case. Perhaps likes and dislikes should not be taken as seriously in morality as Taurek seems to have implied.⁹⁶ So, as I've suggested above, I

section IV.

⁹⁴The intuition seems correct even on the picture of rights and obligations which I have sketched in section IV, where property rights and special obligations do not by themselves necessarily carry the day.

⁹⁵A defender of the standard view of the role of numbers in morality is going to find even more problematic a similar case where the agent decides to save David rather than the five, even though David is a stranger, because he knows and dislikes just one of the five.

⁹⁶I believe that the agent's motives do not play a role in determining the moral status of an action at all; and so whether I am considering

think the problem with the first argument lies not in the conclusion but in the controversial and sometimes unclear assumptions the argument is based on, not to mention its heavy reliance on intuition. Fortunately for anyone suspicious of the standard view of the role of numbers in morality, Taurek has his second argument, which is better, stronger, not so heavily based on potentially questionable intuitions, and which I believe I have so far been able to successfully defend from criticism.

So, do I believe that I have conclusively shown Taurek to be right and numbers not to matter in morality? Not quite. As I've indicated at the outset, in section I, the best I hope to achieve is to establish Taurek's view as a viable alternative to the standard account. The reason for this is that I believe there are many difficulties, as yet unmentioned, that a defender of Taurek's position would need to address in order to bring more than a handful of people in the world over to his or her point of view. Even more would likely be required to entirely overthrow the, at least initially, much more plausible and intuitive standard view, which, as far as I know, has been dominant since the dawn of mankind. In the remainder of this paper I will name a few of these problems and give a brief indication of the most promising, as far as I can tell, directions in which one can look for answers

giving (withholding) the drug to (from) David because I like (dislike) him, or for other reasons, cannot be a factor in whether or not it is morally permissible for me to give the drug to David. Of course, my motives can, and do, play a very large role in determining my own moral status. I discuss the relationship between motives and the moral status of acts in much more detail in the next chapter of my dissertation.

to them. I have to confess that even I do not find my suggestions quite satisfactory. And that is precisely the reason why I do not claim to have helped Taurek defeat the standard view.

Among the factors which Taurek wants to exclude, just to be on the safe side, from the cases he considers, are factors which might make one of the people involved more "valuable". One of the examples he gives is that of a person who is about to invent a cure for some horrible disease. There is, indeed, a reason to exclude such factors from Taurek's discussion. While I am not at all convinced that I am morally required to choose saving a doctor who is almost certain, or even completely certain, to invent a cure for cancer, over saving a perfect stranger, I think I very well might have non-arbitrary grounds for doing so. At the very least, I think the world would, so to speak, be a better place without cancer than it would be with it. Most, if not all, people would, I believe, agree with me. But how can Taurek explain this? It is true that if no cure for cancer is discovered, millions of people worldwide would die of it every year, and even more people would suffer lesser, although still significant, harms. But many philosophers (Kamm, for instance) believe that those deaths cannot be placed on the scales when choosing between saving the cancer doctor and saving a regular person. And even if those deaths could be factored into the equation, taking Taurek's view into account, it wouldn't make any kind of a difference. Whether or not we are allowed to consider those millions of deaths, we still ought to flip a coin to decide whether or not the doctor will be the one saved.

Is there any recourse for Taurek? He could perhaps say that there is a non-arbitrary reason for preferring to save the doctor because there is a probability, almost reaching a certainty, that sooner or later at least one of the agent's relatives, loved ones, friends, or perhaps even the agent himself, will die of cancer if a cure for it is not found. But that seems like a cheap cop-out. For one thing, it just isn't plausible that the possibility that a person who you care about will get cancer is what's behind the intuition that there is a non-arbitrary reason to prefer saving the doctor - it appears that the intuition remains even if we change the example by using, instead of cancer, a much rarer disease, and including a guarantee that no one you really like will ever get it. Besides, especially in light of the discussion at the end of section V, we should not want to reopen the door to possibly allowing minute probabilities to influence such decisions. Also, taking into consideration the problems which led us to discard Taurek's first argument, it is unclear that Taurek can resort to this defense anyway, unless he goes back on his explicitly stated opinion and recognizes the existence of agent-relative permissions. So what is one to do regarding the case of the cancer doctor? I will hold off on my attempt at a response until I have stated the next problem.

Taurek says that special considerations aside, he would flip a coin to decide which group of people to save regardless of the relative sizes of the groups. But even if we can overcome the initial intuitive disbelief we may experience in response to Taurek's conclusion in Six Strangers, doing so becomes harder and

harder as the numbers grow. It is a lot more counterintuitive, not to say appalling, to flip a coin to choose between ten million people and one person, than it is to flip a coin in Six Strangers. But why is that, if numbers don't matter in morality? Of course, even if every single human who has ever lived had believed something or had a particular intuition, that still doesn't mean that the belief or the intuition is not, in fact, blatantly false. That could, for all we know, be the case with my example of ten million vs. one, as well as with the cancer doctor case. But these intuitions are very strong, and so I think that a defender of Taurek's position needs to have more to say about them than just that those intuitions may, nevertheless, be erroneous. Of course, anybody who defends Taurek already has to maintain that pretty much everyone's moral intuitions on the subject of numbers are mistaken. But it seems that there is a good chance that the intuitions regarding cases involving colossal numbers have something substantial behind them, even if the more general intuitions regarding the role of numbers in morality do not.

There is only one suggestion that I can offer at this time to reconcile the existence of these intuitions with Taurek's view on the problem of numbers, aside from the obvious but somewhat strenuous alternative of denying that those intuitions represent moral truths. Perhaps there is something, the loss of which is even more important to a person than the loss of one's life. I don't know what to call that something, but I guess I can utilize the overused term 'way of life'. By proposing this, I do not mean to endorse what is unfortunately a widely accepted practice of

sacrificing human lives for abstract ideals of what human society should be like. What I intend is simply to suggest that the destruction of a person's entire way of life, e.g. the destruction of his or her entire ethnic, national or religious community, can be an even greater loss to that person than the loss of his or her life. Perhaps that kind of reasoning can, for instance, even if Taurek's main conclusion is right, be used to ground a non-arbitrary preference for letting one person die rather than allowing every other human being in the world to perish.

Perhaps some of this reasoning can even trickle down to cases where all of humanity, or even all of one nation or religion, is not at stake, but an enormously large number of people, e.g. ten, or even a hundred, million, is. It may be that the loss of hundreds of millions of people is such a great blow to ours, or somebody else's, way of life, that there is a non-arbitrary reason for avoiding it even at the cost of allowing someone to die. This is all very problematic and rough, of course. And I am far from convinced, for instance, that it can be the case that I am actually morally *required* to allow even one person to die rather than hundreds of millions.

Even more problematic is the question of whether or not there is a non-arbitrary reason, in the sense of some further modification of (MAP), for choosing the alternative that causes a great benefit to our way of life, such as saving the doctor who will discover the cure for cancer, over saving one "insignificant" person. I don't know what the answers should be. That's why I am not offering this suggestion as a solution to the problems I have

raised. At best, there is something here which can perhaps be worked on and turned into some kind of a plausible response. One final note on this line of reasoning, however: if there is something to it, it could explain, even on a non-standard view of the role of numbers in morality, why there is a non-arbitrary reason for preferring to prevent a genocide from occurring, even in a sparsely populated nation, when the alternative is saving just one or several people; even while, as I strongly believe is the case, there is no non-arbitrary reason to prefer preventing a genocide in one nation that would kill six million people to preventing a genocide in another nation that would kill only five million people⁹⁷.

There is one final potential problem for Taurek's view that I want to mention. It is, perhaps, possible to defend the standard view of the role of numbers in morality by saying that always acting in such a way as to save the larger number of people, provided that no special considerations are involved, antecedently maximizes every single person's chances for survival in tradeoff situations, as one presumably has a greater chance of ending up in the larger group. First of all, I am not at all convinced that a true moral principle can be based on maximizing the probability of everyone's survival antecedently to any facts being known, in the way suggested. But if it can, there is only one thing that I have

⁹⁷But a possible controversial consequence may be that there would then exist a moral reason to prefer preventing all 10,000 members of some tribe with a unique way of life from perishing over preventing 200 million Chinese people from suffering the same fate.

noticed I should object to. As we've seen during the discussion of fairness in section VI, there is a considerable debate as to whether it is fairer to give everyone an equal chance before any relevant facts become known or before the relevant situation even obtains; or to give everyone an equal chance in every actual tradeoff situation once it arises, provided, of course, that the people involved are not in any way responsible for ending up in one group or the other, which would, perhaps, add special considerations⁹⁸.

I lean on Broome's side of the debate. I think fairness requires us to give an equal chance to each person in a case like Six Strangers, not to come up with a strategy which gives everybody an equal antecedent probability of survival (prior to any tradeoff situations' arising or being known about), but in any actual tradeoff situation deprives someone of any chance for survival. Under the latter arrangement, I believe we would be unfair to the few, even though we would be making our decision in advance, without knowing who those few will be, and even though we are assuming that everyone has an equal chance to end up in that group (and so a greater chance to end up in the larger group and be saved). That's what I think, but, once again, I do not believe that my response is even close to being completely decisive.

⁹⁸The kind of case that I have in mind here is the sort described in one of Taurek's boat captain examples. If there are only two population centers on an island, one much larger than the other, and the inhabitants are free to choose in which one they live, this case should be discussed separately from, and, perhaps, settled differently than, Six Strangers, especially if everyone was antecedently aware of the likelihood of a major natural disaster in the near future and of the fact that at most one boat would be available for the rescue.

Many problems remain unresolved, and much work still lies ahead for anyone who wishes to join Taurek in opposing the standard view of the role of numbers in morality. The best I can offer at present are a few suggestions in that direction. But I believe that Taurek's paper addresses genuine and valid concerns regarding the standard view and gives a lot to think about to anyone who holds it.⁹⁹ I also believe that my defense of Taurek's second argument and his main conclusion is on the whole successful, and so Taurek's view of the role of numbers in morality remains a viable alternative to the standard account. There is, of course, a lot that remains to be said by both sides in the debate as they search for the truth. And that truth may very well lie in neither camp.

⁹⁹Especially in connection with the very intuitive and seemingly justified, although not easily satisfied, need to have our moral theory respect each person as an individual and each person's suffering as the suffering of an individual (see especially sections II and III).

-Chapter 3-

The Return of The Trolley

THE RETURN OF THE TROLLEY.

Section I.

Introduction to the Trolley Problem.

What is the Trolley Problem? Consider the following pair of cases:

Transplant: Five people are about to die of organ failure. But, "luckily", a healthy young patient walks in for a checkup, and he is compatible with the five as a potential organ donor¹. I, a doctor, can save the five by hitting the one over the head, removing his organs and transplanting them into the five. Of course, the one will die;

and

Trolley: A runaway trolley is headed for five track workmen. They cannot get off the tracks quickly enough to get out of the way and will all be killed should the trolley reach them. But I, a bystander, can get to a switch which will redirect the trolley to a spare piece of track. The five will be saved, but the lone workman on the second track will be killed.

People's intuitions regarding these two cases seem to be quite uniform - in fact, they appear to be virtually unanimous. An overwhelming majority would say that I may not proceed with saving the five in Transplant, but that, at the very least, it is morally

¹Unless I specify otherwise, it will be assumed in all of the cases we will discuss that no special considerations are in effect with regard to any of the people involved. In particular, they are all strangers to me, I am under no moral obligation, or an obligation of any other kind, towards any of them, none of them are in any way "more valuable" than the others, none of them are serial killers, none of them bear any responsibility for the fact that the situation we are considering obtains in the first place, etc., etc.

permissible for me to redirect the trolley in Trolley. Some may even go so far as to claim that I ought to save the five in Trolley.

If one is not presented with the Trolley case for comparison, the natural tendency is to explain the impermissibility of acting in Transplant, despite what is widely perceived as a possibility of achieving an otherwise preferable state of affairs, by appealing to the distinction between killing and letting die, a particular case of the making/allowing distinction. After all, if I do not save the five in Transplant, I will only be letting them die, but in saving them I would be killing the one². However, as Judith Thomson points out in her "Killing, Letting Die and the Trolley Problem" (1976), Trolley is exactly parallel to Transplant as far as the killing/letting die distinction is concerned. In deciding whether or not to divert the trolley, I am also choosing between letting the five die and saving them by killing the one. Thomson goes on to conclude that something other than the killing/letting die distinction must be at work in making it morally impermissible for me to save the five in Transplant.

Roughly speaking, the Trolley Problem is the problem of explaining why it is morally permissible for me to save the five in Trolley, but not in Transplant. A large amount of literature has appeared on the subject in what is now almost a quarter of a century since Thomson's article was published. Additional

²One way of spelling out what underlies this explanation is to say that the negative duty to refrain from killing is more stringent than the positive duty to aid someone who would die if not helped.

examples, some of which will be mentioned later on, have sprung up as counterexamples to potential solutions, indicating an unavoidable increase in the scope of the problem³. Considering the amount of literature available on the Trolley Problem, a comprehensive survey of proposed solutions would be neither wise nor possible. However, I will discuss two of them in detail, and say a few things about several of the others.

In the second half of this paper I will argue that not only are none of the proposed solutions satisfactory, but, furthermore, that all attempts in similar directions seem to contain inherent flaws which cannot be overcome. However, it is entirely unacceptable to suppose that it is morally permissible to save the five in Trolley, although not in Transplant, despite there being no explanation whatsoever for the moral difference. So, I believe, if there is no explanation, there is no moral difference. But instead of looking for yet another solution to the Trolley Problem, I will take up and defend a radically different position - I will argue, for reasons that I will go into later, that even though it is almost a consensus that I am at the very least morally permitted to save the five in Trolley, that is, in fact, not the case. In the process, since I believe that the killing/letting die distinction has a very important role to play here, I will have to revisit that distinction and, after concluding my discussion of the Trolley Problem itself, defend its

³Of course, these further examples are extremely helpful in clarifying our intuitions and in weeding out attempted solutions which are carefully crafted to fit the already available group of cases, but are not generalizable to fit our intuitions in other relevant situations.

significance for morality against some prominent objections.

As I will conclude towards the end of this paper, there may be some relatively minor, or even not so minor, differences between Trolley and Transplant, each of which renders saving the five in the latter case more morally objectionable, or, at the very least, makes it seem that it is more objectionable. However, I will argue that saving the five is morally impermissible in both cases. Thus, in effect, my solution to the Trolley Problem will consist in claiming that it doesn't exist. This, of course, will force me into my final challenge - the unenviable task of explaining how everyone, philosophers included, can be so ready to jump to a conclusion which I am claiming is wrong, and even sustain their belief in that conclusion upon a considerable amount of reflection.

But before we embark on our journey on the trolley, there is one more issue that I feel I should address. In light of my previous chapter, one might ask: "Why should you bother with the Trolley Problem at all if you believe that numbers don't count in morality and so there is no moral reason to turn the trolley onto the one in the first place"? In fact, several philosophers writing on the Trolley Problem did find it necessary to include footnotes at some point in their articles, signifying that they are aware of Taurek's challenge, and referring the reader to alleged refutations of Taurek's conclusion. But I think it is important to notice that given that the predominant view is that it is morally permissible to divert the trolley in Trolley, a possible position for a convert to Taurekism to take is that it is

morally permissible to flip a coin⁴ in order to determine whether or not to divert the trolley. On the other hand, hardly anyone on either side of the Taurek controversy would argue that it is morally permissible for me to flip a coin in Transplant in order to determine whether or not to kill the one healthy patient, cut him up, and take his organs. Thus, even if Taurek is right, that does not necessarily do anything to defuse the Trolley Problem.

Section II.

Some attempted solutions.

Subsection 1: Introduction and the Doctrine of Double Effect.

Before proceeding to a more detailed discussion of a couple of proposed solutions to the Trolley Problem, I want to briefly address some very obvious and a few not so obvious, attempts to rise to the challenge. As we have already seen, the killing/letting die distinction, at least by itself, cannot be the answer, since Trolley and Transplant are symmetric with regard to that distinction. Can some sort of a Kantian injunction be the answer we seek? After all, the one in Transplant would be used as a means to the survival of the five - if the one's organs are not made available, the five will die; whereas the one in Trolley is not needed to ensure the survival of the five - in fact, if the

⁴Or perform some probabilistic equivalent, physically or mentally.

spare track were empty, that would make me very happy as I would not even need to hesitate before saving five lives at the cost of none. However, things are not that simple. In "The Trolley Problem" (1985), Judith Thomson offers the following case -

Loop: A runaway trolley is about to kill five track workmen. But I can get to a switch which will divert the trolley to a spare track on which only one track workman will be killed. However, the diverging pieces of track loop back. The only reason the five would be spared if I divert the trolley is that the one is fat enough to stop the trolley on impact. The five are, perhaps, thinner, but their combined weight will be sufficient to prevent the trolley from looping back and reaching the one if I choose not to divert the trolley.

Virtually everyone who believes that it is morally permissible to divert the trolley in Trolley believes that it is also permissible to divert the trolley in Loop. Almost everyone will even agree that killing the one to save the five is permissible in

Modified Loop: This is just like Loop except the spare piece of track loops back in such a way that while the death of the one is necessary in order to prevent the trolley from coming back to kill the five anyway, should I divert the trolley; if I do not divert, even if the five do not stop the trolley, it will just keep going in a straight line rather than loop back to kill the one⁵.

But in these two cases, the one is also needed as a means to saving the five. So a Kantian solution doesn't seem to work.

Another attempted solution which is often brought up in articles concerning the Trolley Problem involves the infamous Doctrine of Double Effect (DDE). Much has been written to hopefully discredit the DDE for good. So, it would not be prudent

⁵One known exception is Michael J. Costa.

to go into a detailed discussion of the DDE. However, I think a brief review of the major problems for the DDE is in order, mostly for the sake of introducing several examples and one major distinction which will be very important later on in this paper.

First, a quick reminder of what the Doctrine of Double Effect is.

The Doctrine of Double Effect (DDE): It is sometimes morally permissible to perform an act which has both a good and a bad consequence, but only if the following conditions obtain:

- 1) the good consequence outweighs the bad (in some sense, the world is a better place if both take place than if neither does);
- 2) the agent intends the good consequence, but does not intend the bad one - he or she merely foresees it;
- 3) the bad consequence is not a means to the good consequence, as it is assumed that the agent cannot intend the end without intending the means.

It seems that the DDE can handle explaining the difference between Trolley and Transplant. After all, in Transplant the death of the one is the means to the continued life of the five, whereas in Trolley the means to preventing the deaths of the five is my redirecting the trolley - the death of the one is merely a foreseen tragic consequence, which, on the traditional view of the role of numbers in morality, is outweighed by the good of saving five lives. Now I will quickly run down the list of four major problems with the DDE, which show that in addition to problems of interpretation, the DDE clearly issues the wrong verdict in many cases, and even the ones it gets right, it gets right for the wrong reasons.

The first problem, indicating that the DDE is too permissive, was brought up by Philippa Foot in "The Problem of Abortion and the Doctrine of the Double Effect" (1975), the article that first introduced an early version of the Trolley example. Consider the case of the

Lethal Fumes: "...There are five patients in a hospital whose lives could be saved by the manufacture of a certain gas, but ... this inevitably releases lethal fumes into the room of another patient whom for some reason we are unable to move"⁶.

Foot claims that it is clearly impermissible to proceed with the manufacture of the gas even though, according to the DDE, there is no reason why that should be the case because the death of the one is not intended and is not a means to saving the five. Many philosophers working on the Trolley Problem agree with her, and, as far as I know, no one has disagreed. Thus it definitely appears to be the case that while it fits the intuitions of the majority regarding the Trolley and Transplant cases, the DDE fails to fit other intuitions which are held just as firmly by those very same people.

Of course, the DDE is supposed to only provide us with necessary conditions for moral permissibility, and not with sufficient ones. Thus, even if, in fact, it is morally wrong to produce the gas in Lethal Fumes, that in itself doesn't invalidate the DDE because the DDE never tells us that an action is morally

⁶Philippa Foot, "The Problem of Abortion and the Doctrine of the Double Effect", reprinted in ETHICS Problems and Principles (EPR), ed. by John Martin Fischer and Mark Ravizza, Harcourt Brace Jovanovich College Publishers 1992, p.65.

permissible. But, if we are right about Lethal Fumes, we can already conclude that the DDE by itself cannot serve as a solution to the extended Trolley Problem, which, as I explained in section I, encompasses not only Trolley and Transplant, but a family of related cases as well. And, in the discussion that is to follow, I will, I hope, successfully argue that the DDE cannot be a part of a solution to the Trolley Problem either. However, I find the judgment regarding Lethal Fumes to be very problematic in any case for those who believe that saving the five is morally permissible in Trolley⁷. So, I wouldn't want my argument against the DDE to rest on Lethal Fumes entirely. It is, at any rate, better to slightly postpone the discussion of Lethal Fumes.

The second problem is that the DDE is also too restrictive. It seems that according to the DDE, diverting the trolley is impermissible in Loop, since the death of the one would serve as a means to saving the five. However, as I've mentioned before, almost everyone who believes that saving the five is morally permissible in Trolley, believes that the same is true of Loop, and even of Modified Loop. So, once again, the results that the DDE yields seem to go against the moral intuitions of the majority.

But there are two even more serious problems lurking here.

⁷I even find that for those who believe that it is permissible to divert the trolley in Trolley, the judgment that saving the five is impermissible is problematic in **Lethal Fumes 2**, a case just like Lethal Fumes but with the difference that the five are actually saved by a surgical procedure requiring the manufacture of the gas; which, for reasons that will become clear later on, is considered even more morally troublesome by some who have worked on the Trolley Problem, than the manufacture of the gas in Lethal Fumes.

One is the problem of individuation of events. To use examples that often come up in related discussions, if I crush the fetus' skull while performing a craniotomy in order to save the mother's life, have I thereby killed the baby or merely caused it to die? If I run over a person on the road, have I thereby killed her, or only caused her to, say, bleed to death? In fact, if I cut up the one in Transplant and take his organs - will I have killed him by that act, or will I have just caused him to die?

The importance of these questions lies in the fact that while it is assumed that I cannot intend the end without intending the necessary means, I should be able to intend the end without intending something that is merely caused by the means, as otherwise condition 2 of the DDE will never be fulfilled in cases where the only way to save five people is to do something that results in the death of some other person. So, if I merely cause the baby to die in performing the craniotomy, or merely cause the one to die in Transplant by removing his organs⁸, it seems that I can intend to perform the craniotomy without intending that the baby die, and that I can also intend to save the five in Transplant without intending that the one whose organs I take die. Thus it appears that according to the DDE I may proceed with saving the five in Transplant after all. But even if some uniform way can be found of distinguishing the act itself from the causal consequences of that act, a deeper question arises: why should

⁸Perhaps I can make this more plausible by crossing over into the realm of science-fiction and offering some sort of an instantaneous, or nearly instantaneous, matter-energy transporter of the Star Trek variety as the method used for the transplant.

determinations like this have any bearing on the moral permissibility of acts such as cutting up the one in Transplant? This issue will also be relevant to the discussion of Kamm's solution in the next section⁹.

And the most serious blow to the DDE is that it seems that what underlies it is a serious error regarding the very nature of ethics. As has already been pointed out in the literature, most notably by Judith Thomson, there are two kinds of ethical questions that one can ask with regard to an act¹⁰. The first concerns the moral properties of the act itself, e.g. whether or not it is morally permissible, whether or not it violates someone's property rights, etc. The second question deals with

⁹In "Actions, Intentions, and Consequences: The Doctrine of Double Effect (1989), reprinted in *EPR*, Warren S. Quinn suggests a modification of the DDE which at first glance seems to avoid this objection - we cannot engage in *harmful direct agency* - "agency in which harm comes to some victims ... from the agent's deliberately involving them in something in order to further his purpose precisely by way of their being so involved". Aside from still leaving the DDE open to the three other major objections that I raise, there is another problem - if "involved" is interpreted as not allowing even merely foreseen involvement, the DDE becomes useless as it prohibits action in all redirection cases; but if foreseeing involvement is allowed, then consider the following modification of the famous Terror Bomber case - I drop a bomb on a major population center of an evil enemy country, that contains a device which will create a telepathic field affecting the entire country in the following way: everyone outside the region will believe that the region has been turned into a desolate wasteland by some new weapon and, in addition, their minds will filter out all attempted communication from the people in the region. This will cause the evil country to surrender, but, by the way, the telepathic field will kill everyone in the region within 24 hours as a foreseen side effect. This case is clearly morally equivalent to the standard Terror Bomber case, but it seems that Quinn's modification will have to allow this kind of bombing.

¹⁰Jonathan Bennett also points this out in section 15 of The Act Itself.

how the act reflects on the moral character of the agent performing it. The agent's intentions, as well as perhaps some of the agent's other mental states, are very relevant to the second type of questions, but have nothing to do with the first.

If Jack goes on a killing spree in a local preschool while under the influence of a mind-controlling drug which was planted in his breakfast cereal by an evil scientist, we might decide that this says nothing about Jack's moral character since he cannot be held responsible for his actions under the circumstances, but we should not conclude that going on a killing spree in a local preschool is sometimes morally permissible. Similarly, if I want the one on the spare track in Trolley to die because I hate him or his entire race, or if I just want to experience the feeling of knowing that I killed someone but will not be punished for it, we should say that this has bad implications for my moral character - I am probably not a very good person; but it seems grossly counterintuitive to say that because of this it is not morally permissible for me to divert the trolley, if simply changing my state of mind could have made it permissible to perform that very same act. Also, if I ask someone whether or not it is permissible for me to divert the trolley, it would be a strange response, indeed, if I heard in reply that the answer depends on my intentions - if I merely foresee that the one will die, but don't want him to or don't care, then it's OK, but if I want him to die, then it's no longer OK^{11 12}.

¹¹Judith Thomson makes a similar argument in "Physician-Assisted Suicide: Two Moral Arguments", *Ethics* 109 (April 1999).

I believe we should conclude that intentions are not the proper basis for judging the moral permissibility of actions, and thus all attempts to bring a version of the DDE to bear on the Trolley Problem are inevitably doomed to fail¹³. If, however, defenders of the DDE try to shift from the actual intentions of the agent to the issue of whether or not it is possible in principle for an agent to act in the manner in question without intending the negative consequence, it is easy to see that we are not really talking about intentions at all any more - we are reducing the DDE to a principle concerned solely with means, ends, and causal relations. Some such principles will be discussed later on.

Subsection 2: New Threats versus Old Threats.

Some philosophers, e.g. Michael J. Costa, have tried to defend the DDE from the charge that it delivers the wrong verdicts in some cases by modifying it to include an injunction against creating a new threat instead of simply redirecting an already existing one.

¹²I want to point out that it is not the case that I can positively influence my moral worth by refraining, because my intentions are inappropriate, from doing something that is morally permissible - there is no tradeoff going on here. My intentions speak poorly of me, but my not doing something (because of my intentions) that is morally permissible and, in fact, might even be considered praiseworthy, can only, depending on the precise circumstances, either leave me with the same poor moral standing or make things even worse.

¹³As a side note, I want to point out that it is not surprising to find the DDE making such a fundamental mistake, as it originated as a religious doctrine, and religious doctrines are usually a lot more concerned with the moral worth of agents than with the moral permissibility of acts in themselves.

Of course, in light of what I offered as the third and fourth objections to the DDE, this modification appears to be a case of "too little, too late". However, the distinction between redirecting a pre-existing threat and creating a new one has been brought up for reasons unrelated to the DDE as well, for instance, in Judith Thomson's "Killing, Letting Die, and the Trolley Problem" (1976) - the article that framed the Trolley Problem and also featured Thomson's first crack at it - and her "The Trolley Problem" (1985)¹⁴. So, I think there is sufficient reason to say a few things about this distinction.

Costa talks about an act that introduces a new threat as "an act that is itself *prima facie* evil and one that is more difficult to justify"^{15 16}. And Thomson discusses in her 1985 paper of a "'distributive exemption', which permits arranging that something that will do harm anyway shall be better distributed than it otherwise would be - shall do harm to fewer rather than more"¹⁷. I see several problems with this approach to the Trolley Problem. The first, in response to Costa, is that I fail to see why creating a new threat is necessarily evil - more evil than redirecting an already existing threat¹⁸. Of course, if you

¹⁴Both reprinted in EPR.

¹⁵EPR, p.306.

¹⁶Thomas Nagel seems to be expressing a similar concern in "Ethics", also reprinted in EPR.

¹⁷EPR, p.287.

¹⁸And it would be hard to argue that no explanation is necessary for why it is worse, all other things being equal, to create a new threat than to redirect an old one, as it seems to be highly implausible that

conjure up an image of a mad scientist creating a hitherto unknown deadly virus to wipeout all mankind, what we are imagining is an evil act. But if we are talking about manufacturing a gas needed to save five people, or, say, stopping a runaway trolley by the only means available, which, in turn, causes a threatening avalanche, why exactly is that intrinsically more evil than diverting a trolley to a previously unthreatened person? I will return to this issue a little bit later.

The second difficulty concerns determining what constitutes a new threat as opposed to an old threat. It is not always easy to draw the line. If I manage to stop the trolley, but its stopping causes an avalanche, the avalanche is presumably a new threat. But what if the trolley's stopping causes it to explode, and the explosion threatens a bystander? Is the explosion a new threat? What if I stop the trolley in a way that causes it to be ripped in half, and one of the halves proceeds to fly over the spare track towards the one? Is that half a new threat, or is it still the original trolley threat?¹⁹ More importantly, why should any truth about morality depend on the answers to those questions and others like it? If I am not mistaken, similar considerations may have been part of the reason why Thomson has apparently given up on the

such a controversial claim with, as I will argue shortly, some very counterintuitive consequences, can simply be a brute fact about morality.

¹⁹In Morality, Mortality: Volume II, chapter 7, p. 175, Frances M. Kamm even goes so far as to claim that the trolley's looping back towards the five in Loop, in case the one's body fails to stop it for some reason, would constitute a new threat to the five, as the trolley would approach from a different direction!

view that the new-threat/old-threat distinction in itself has any bearing on the Trolley Problem.

And now I would like to introduce a few more examples into the picture, which, I believe, will illustrate the fact that our moral intuitions do not support the hypothesis that whether we create a new threat, or simply deflect an already existing one, has any moral significance. Let's consider the cases of

Two Buttons: There is a runaway trolley headed for five people who can not get out of the way in time. There is a spare piece of track on which one person is standing. I can get to a panel with two buttons, A and B, in time to save the five. If I push button A, the trolley will be diverted to the spare track and the one will be killed. If I push button B, a huge wall of iron will come out of the ground in front of the trolley and stop it. But the impact of the trolley against the wall will cause an avalanche which will kill the one. There are no other ways in which I can save the five;

Three Buttons: This is just like Two Buttons, except there is also a third button, C. If I push C, an iron wall will also come out of the ground to stop the trolley, but in this case it will be the emergence of the wall itself, and not the impact of the trolley against the wall, that will cause the avalanche. As in Two Buttons, there are no other ways in which I can save the five²⁰;

and

Four Buttons: This is just like Three Buttons, except that in addition to the panel containing buttons A, B, and C, there is also a separate panel containing button D. Pressing D will also cause the iron wall to emerge from the Underworld, but this time it will be my running over to the panel containing D that will trigger the avalanche.

²⁰The importance of this case, as well as of Four Buttons, and of the difference between Two Buttons and Three Buttons, will not become clear until the next section.

If A were the only button available to me in Two Buttons, the case would be identical to Trolley, i.e., according to everyone writing on the Trolley Problem, it would be morally permissible for me to push A.

Now consider

Avalanche: There is a runaway trolley headed for five people who can not get out of the way in time. I can get to a button, my pushing of which will cause a huge wall of iron to come out of the ground in front of the trolley and stop it. But the impact of the trolley against the wall will cause an avalanche which will kill an innocent bystander. There are no other ways in which I can save the five.

If B were the only button available to me in Two Buttons, the dilemma would be the same as in Avalanche, and structurally analogous to Lethal Fumes²¹, i.e. it would allegedly be impermissible for me to push the button. Now, it seems to be a reasonable assumption that the following is true -

Firm Boundaries Principle: If there is a desirable outcome O, and it would be morally permissible for me to adopt a course of action A in order to achieve O, provided that it is the only way to achieve O, but it would not be permissible for me to adopt a course of action B in order to achieve O, provided that it is the only way to achieve O; then should both alternatives be available, if there is no other way I can achieve O, I ought to either do A or forfeit O. In either case, B should still be off limits morally.

²¹The only difference I can think of is that hospitals have a special obligation towards their patients not to harm them, even in order to help other patients. This may make a difference in causing saving the five to be impermissible in Lethal Fumes. But I am assuming that this can't possibly be all that Foot, Thomson, and others mean when say they say that manufacturing the gas is impermissible, otherwise all that the example would show is that it was chosen poorly and needs to be restated so as to eliminate the factor in question.

For example, if, for argument's sake, I am not morally permitted to cut off a person's arm to save someone else's life, but I am morally permitted to steal \$5 for the same purpose; should a relevantly similar situation arise in which the only two ways for me to save a life are to either steal \$5 or to cut off someone else's arm, it seems that I ought to either refrain from saving the life altogether, or to steal \$5, but I ought not cut off anyone's arm.

But I can see no reason why, especially in light of the Firm Boundaries Principle, I, the one, God, or anyone else should care whether I push A or B in Two Buttons, assuming that on some morbid scale being run over by a trolley ranks as being a death roughly equal in unpleasantness to being buried by an avalanche²². In either case the one will be killed so that the five may live. In either case the death of the one is not, strictly speaking, the means to saving the five, but, to use a DDE term, only a foreseen consequence. So where is the morally relevant difference in virtue of which one course of action is permissible and the other is not?²³ In fact, I can even imagine a case where the one has

²²Perhaps I could make my argument even stronger by making the original threat an avalanche as well, instead of a trolley; but I think that is unnecessary - the Two Buttons case should be able to speak for itself.

²³Another example which might be very helpful in illustrating this point is

Pushing Trolley: This is just like Trolley, except that should I choose to divert the trolley, it will actually stop, but in so doing push another trolley, which until then is stationary, causing it to go down the spare track and kill the one.

It seems obvious that if it is morally permissible for us to save the five in Trolley, it should be morally permissible for us to do it in

previously expressed a preference for death by being buried by an avalanche over death by trolley. Such a preference cannot change anything as far as the permissibility of pushing B in cases equivalent to Lethal Fumes, such as Avalanche, but it seems ludicrous to suggest that still I ought to either push A or do nothing at all in Two Buttons because it is not morally permissible to push B²⁴.

To sum up, I can see absolutely no moral difference between pushing A and pushing B in Two Buttons, and, therefore, I do not believe that there can be any moral difference between Trolley and Avalanche, and so between Trolley and Lethal Fumes. This is the reason why I claimed earlier that the prevalent intuition regarding Lethal Fumes is problematic. Similarly, I do not see how there can be any moral difference between pushing A, B, or C in Three Buttons, as the difference in causal structure between what happens if I push B and what happens if I push C, equivalent to the difference between Lethal Fumes and Lethal

Pushing Trolley as well, contrary to what those who defend the diverting-old-threat/creating-new-threat solution to the Trolley Problem seem to be forced to maintain.

²⁴There may be small room for concern that I am being so overwhelmed by the fact that in either case the one dies, that I end up ignoring a real moral difference which is, nevertheless, there to be found; however, I believe that this concern can be easily alleviated by changing the example so that pushing each button kills a different person. I still see no possible grounds for a moral difference. Also, I can see quite a difference, if the one is on a bridge overlooking the tracks and there is a spare piece of track going to the bridge as well, between pushing a button that will catapult him off the bridge into the path of the trolley, his body then halting the trolley's progress, and pushing the button which will redirect the trolley towards him, despite the fact that he dies either way in this case as well.

Fumes 2 (see footnote 7), seems to be entirely morally inert. Thus I believe that if saving the five is permissible in Trolley, it has to be permissible in Lethal Fumes and Lethal Fumes 2 as well. Therefore, the new-threat/old-threat distinction cannot supply us with a solution to the Trolley Problem.

The solution just considered, based on the new-threat/old-threat distinction, fails on three very important counts. First, it introduces a distinction which is very hard to draw in some cases, and, which in fact, might have to be drawn by fiat in many situations. That wouldn't necessarily be an insurmountable problem, if not for the fact that it is unclear why the distinction should matter in those cases in the first place, even if it can be drawn in a reasonable manner²⁵. And, finally, once we move beyond the Trolley and Transplant cases, the criterion based on the diverting-an-old-threat/introducing-a-new-threat distinction no longer seems to deliver verdicts which agree with our intuitions. This is in general something we need to watch out for when considering proposed solutions to the Trolley Problem - any criterion carefully crafted to fit our intuitions regarding several famous cases is very likely not to be generalizable, i.e. it is very likely to be open to counterexamples.

²⁵In fact, I think it would be quite irrational for me to disavow my moral intuitions about a case involving, say, five people threatened by a trolley and a person who would be killed by a flying half of that trolley should I intervene to save the five, simply because someone convinces me that my original intuitions regarding whether or not a flying half of a trolley is the same threat as the whole trolley were mistaken.

Subsection 3: A Few More Attempted Solutions.

The next ill-fated attempt to solve the Trolley Problem that I want to mention belongs to Robert Hanna. He bases his response on the distinction between being a participant and being a bystander in a given situation. As he phrases his own criterion, "in cases in which there is the option of either (a)harming one bystander by forcing him into the threat-situation, or (b)harming one or five participants in the same threat-situation only in order to bring about the best possible distribution of the threat, it is impermissible to do (a) but sometimes permissible to do (b)"²⁶. This criterion, unlike the last one that we were looking at, has the advantage that it seems intuitively clear, assuming we can come up with a reasonable way to differentiate bystanders from participants, why the distinction it is based on could make a moral difference - as Hanna puts it, we can have a valid complaint against a person for involving us in a dangerous situation, but we can't have a complaint against the world for doing so. But this criterion fares even worse than its predecessor as far as being able to fit our intuitions. Also, the criterion is spelled out in a manner which appears to be more or less arbitrarily contrived to deal with the few original cases, and thus proves to be impossible to generalize.

Hanna claims that the healthy individual in Transplant is clearly a bystander and not a participant. But, he goes on to say, "the one in The Trolley Driver [his version of the case I am

²⁶Robert Hanna, "Morality De Re: Reflections on the Trolley Problem", printed in EPR, pp.330-331.

calling 'Trolley'] is clearly a participant, since one of the pathways of the ongoing trolley-event will make him a victim. The mere existence of the possible pathway, as physically represented by the switching mechanism and spur, is sufficient to implicate the one"²⁷ [in what??]. This seems very wrong. The one feature that Trolley and Transplant definitely appear to have in common is that the one in both cases is not really involved in any way unless and until I do something to threaten him. Hanna seems to be saying that the presence of an extra piece of track and a switch makes the one in Trolley a participant, even though the presence of a knife and an operating table does not make the one a participant in Transplant, but that appears to be a completely arbitrary judgment²⁸. For one thing, if the threat in question is a missile, and I have the means to divert it to virtually any inhabited place on the globe, but not to an uninhabited one, is almost every person in the world to be considered a participant in this crisis?²⁹ That does not seem reasonable. Second, suppose that in Trolley there is yet another piece of track to which the Trolley can be diverted, which is empty. Does that mean that the one is no longer a participant, despite the presence of the piece

²⁷EPR, p.331.

²⁸It is also suspect as far as possibly relying on the distinction between redirecting a threat and creating a new one, which we have already discussed and rejected as being irrelevant to solving the Trolley Problem.

²⁹This question is inspired by a similar example from F.M. Kamm's Morality, Mortality vol.2, where it is part of an objection to Montmarquet, who appears to be making unsuccessful attempts to solve the Trolley Problem that are reminiscent of Hanna's.

of track leading up to him and a switch? If not, how can Hanna explain that? If yes, why can't we turn the trolley onto him, as surely we ought not in this case?

And, last but not least, we should consider yet another case which appears to be morally equivalent to Transplant -

Fat Man: A runaway trolley is headed towards five people. I, a bystander, am located on a bridge overlooking the tracks. The only way to save the five is to throw a fat man, who is standing next to me on the bridge, onto the tracks in front of the trolley. His body would then prevent the trolley from proceeding to kill the five.

Clearly, it is not morally permissible to either throw the fat man off the bridge, or even just wobble the handrail slightly so that he will fall on his own. Shockingly, Hanna claims that the fat man, unlike the one on the tracks, is a mere bystander! In this case, the presence of the bridge over the tracks, coupled with my ability to throw him off or wobble the handrail, which I can even supplement by making a button available to me which will catapult the fat man off the bridge, is somehow insufficient to make the fat man a participant. Is it because there are no tracks leading up to where he is? Would the verdict be different if the tracks diverged and one piece of track led up to the bridge where the fat man is standing? If so, why?³⁰ This attempted solution seems to

³⁰Some may try to respond that the one in Trolley assumes a certain amount of risk by walking out onto the tracks, which is not the case with the fat man in Fat Man. But there are many ways to rule out this asymmetry, even if it is significant, without affecting the heart of the problem that Hanna is facing. We can, for instance, stipulate that the spare piece of track in Trolley has not been used in 20 years, and for that reason there aren't even any warning signs posted in the vicinity. Or we can also stipulate that there is some risk that the fat man assumes by going on the bridge, e.g. that there is

be a complete non-starter because any pre-theoretical intuitive notion of "bystander" fails to yield the results that Hanna needs, and without such an intuitive notion Hanna's purported solution is circular - it appears to use moral intuitions to derive some technical notion of being a bystander, and then turns around and uses that notion to justify those very same intuitions.

The final suggestion that I want to address in this section can be found in Thomson's "The Trolley Problem". She argues that in the relevant kinds of cases it is permissible for me to kill one in order to avoid letting five die, provided that (1) I will be doing so by diverting an already existing threat, and (2) the action by which I will be achieving this outcome will not in itself infringe on a stringent right of the one³¹. We have discussed the requirement in (1) already, and this in itself makes the proposed solution inadequate as it stands, but (2) still makes it worthy of discussion, as Thomson is suggesting a new approach here - looking at features of the cases which are intrinsically related to ethics.

It is clear that in cutting up the healthy patient in

also a piece of track on the bridge itself, but which has also not been in use for 20 years. It just doesn't seem that there is any room for a would-be defender of Hanna's view to maneuver here, especially since we can, in principle, quite easily switch from runaway trolley cases to crashing airplane and nuclear missile cases, in which there would be no tracks and so absolutely no reason to say that the people to whom the threat would be diverted assumed any risk by being where they are.

³¹The part about stringency is needed in order to differentiate between violations of such rights as the right not to be pushed off a bridge or the right not to have one's organs forcibly removed, and rights such as, say, the right not to have your property trespassed on in order for me to get to a switch which is located on your land.

Transplant I surely violate a stringent right of his, especially since, in our society at least, organ donation is purely voluntary, and so it may even be considered a violation of a stringent right of his to use his organs once he is dead without his prior consent. It is equally clear that I violate a stringent right of the fat man in Fat Man if I push him off the bridge. But, as Thomson herself admits, while diverting the trolley is in itself an inoffensive act, if there is a person on the track to which you are diverting the trolley, it seems to violate a right of his as well, as becomes immediately clear if we consider a modified version of Trolley in which there are no five people threatened by the trolley.³² However, Thomson argues, Trolley is different because the means which I use to save the five, i.e. diverting the trolley to a different track, "is not *itself* an infringement of a right of anybody's. The agent would do the one no wrong at all if he turned the trolley onto the right-hand track and by some miracle the trolley did not hit him [the one]."³³

First, I have to disagree with the last quoted sentence - if someone shoots a bullet at my head but by some miracle it falls on the ground 1 mm short of penetrating my forehead, I still believe that the shooter violated a very stringent right of mine by doing something that would normally result in my acquiring an extra orifice in my skull. It's just that both of us would get lucky -

³²One can respond that causing someone harm is a violation of his or her rights only if he or she is harmed unjustly. But that, of course, would render any application of a rights-based approach to the Trolley Problem hopelessly circular.

³³EPR, p.288.

I would get to continue living, and criminal charges against the shooter would be downgraded from murder to attempted murder.³⁴

Second, I am not sure that it is possible to square this claim made by Thomson with what she has previously admitted about the violation of rights resulting from my diverting the trolley to an "inhabited" piece of track³⁵. Thomson herself accepts this in her footnote 2 to chapter 7 of The Realm of Rights, where she repudiates her previous attempts to solve the Trolley Problem and offers an entirely different approach which I will discuss in section IV.

And, last but not least, there is the problem of the Handrail case.

Handrail: This is a variant of Fat Man, in which all I need to do to save the five is slightly wobble the handrail of the bridge. That will cause the fat man to lose his balance and fall off the bridge onto the tracks, thus halting the progress of the trolley.

The problem is that the means to saving the five in Handrail is simply the wobbling of the handrail, which is an action hardly

³⁴Thomson disagrees - she believes that in cases where no actual harm, physical, psychological, or of any other kind, is done, an attempt to harm doesn't constitute an infringement on anyone's rights, as evidenced by the putative fact that no compensation is owed to the victim. I am not convinced that this is true, but I want to sidestep this issue, as nothing of any importance rests on it in this paper.

³⁵At least it might be problematic without lapsing into the kind of debacle that the DDE faced, i.e. that I would also not be violating the rights of the one by cutting him up in Transplant because I will be saving the five by means of making cutting motions with my right hand, which would not be a violation of the one's rights if his body didn't happen to be in the space where I am making the motions, or if, by some miracle, my cutting motions didn't hurt him even though his body was "in the way".

more offensive in itself than diverting the trolley onto a different track. However, it is clearly morally impermissible for me to wobble the handrail. Thomson suggests that we should opt for a wider notion of "means" - if "you get a trolley to threaten the one instead of five by wobbling the handrail, ..., the means you take to get the trolley to threaten the one instead of the five include wobbling the handrail, and all those further things that you would have to succeed in doing by wobbling the handrail if the trolley is to threaten the one instead of the five"³⁶. This makes wobbling the handrail impermissible because the means to saving the five would then include having the one fall off the bridge as a result of my actions, which presumably violates his rights quite severely. But, surprisingly, Thomson does not seem to notice that this apparently exposes her answer to the same counterexample as she herself proposed for the Kantian solution - the Loop case. In Loop, by the same criterion, the means would have to include the one's being hit by the trolley in such a way that his body stops the trolley - a violation of a very stringent right of his, I believe. Thomson, however, repeats just a few pages earlier that she believes it is permissible to save the five in Loop.

In conclusion, I want to mention one last point from Thomson's article. She argues that it is important, as far as the issue of rights is concerned, whether something is done to the one or to the threat in order to save the five. I think that in this

³⁶EPR, p.289.

she may be right, and, perhaps, there is a moral difference between the two kinds of cases³⁷ ³⁸. But, as Thomson herself admits in The Realm of Rights, that difference in itself, even if it exists, seems to be insufficient to account for the enormous difference between our intuitions in Trolley, and Transplant or Fat Man, as well as for our intuitions regarding the entire array of Trolley-Problem-related examples, especially in light of the apparent overall collapse of the "rights approach" to the problem. Therefore, this difference cannot be the basis for a satisfactory solution to the Trolley Problem³⁹.

Section III.

Kamm and her Principle of Permissible Harm.

Frances M. Kamm has offered two very closely related purported solutions to the Trolley Problem, which will be the topic of this

³⁷Although I still have a concern here regarding the issue of how to morally distinguish Loop from Handrail to a sufficient degree without exposing oneself to easy counterexamples.

³⁸In addition, this intuition is not uncontroversial - Kamm, for instance, strongly disagrees with it when she presents her Lazy Susan cases in the chapters on the Trolley Problem in Morality, Mortality vol.2. I, however, strongly disagree with Kamm's intuitions ...

³⁹If it were the solution, it would also, I think, be a problem for Thomson's and others' view that saving the five is impermissible in Lethal Fumes. But, of course, as I argued earlier, I don't think that this view is defensible in any case for those who share the standard intuition regarding Trolley.

section. The first attempt⁴⁰ I will dub the

Causal Chain Length Principle (CCLP): "It is permissible to cause harm to some in the course of achieving the greater good of saving a greater number of others from comparable harm, if events that produce the greater good are not more intimately causally related to the production of harm or to other intrusion onto the person or what is his which leads to such harm, than they are to the production of the greater good (or its structurally equivalent component)".⁴¹

Kamm's second, revised and much more complicated, attempt is the **Principle of Permissible Harm (PPH)**, for which she offers two different and, by her own admission, not necessarily equivalent formulations, **PPH(1)** and **PPH(2)** -

PPH(1): "1. It is permissible that greater good (or its structurally equivalent component) have lesser harm as one of its aspects or as its direct or indirect effect, even when this harm sustains that greater good.

2. It is permissible that a mere means to a greater good have lesser harm as an effect, if and only if the greater good (or its structurally equivalent component) is an aspect of the means or noncausally its flipside, even when this harm sustains the greater good."⁴²;

and

PPH(2): "1. It is permissible that a greater number of people being saved from all death threats already facing them have lesser harm as one of its aspects or as its direct or indirect effect if and only if their being saved from all death threats

⁴⁰It originally appeared in "Harming Some to Save Others", *Philosophical Studies* 57(3) (November, 1989): 227-260.

⁴¹Frances M. Kamm, Morality, Mortality: Volume II - Rights, Duties and Status, New York - Oxford, Oxford University Press 1996, p. 201 (footnote 1); Unless specified otherwise, all future quotations from Kamm are from that volume.

⁴²Kamm, p.183.

already facing them is aimed to accomplish the greater good of saving those very same people⁴³.

2. It is permissible that a means to saving a greater number of people from all death threats already facing them have lesser harm as an effect if and only if (a) saving those people from the threats already facing them is an aspect of the means or its noncausal flipside, and (b) this saving is aimed to accomplish the greater good of saving those very same people."⁴⁴

The PPH, in both of its formulations, is supposed to represent Kamm's improvements on the CCLP, the main difference being that the PPH, unlike the CCLP, requires not only that the action in question be more closely causally related to the good it creates than to the evil, but also that the action, if it is merely a means to the good, be noncausally related to it. However, the PPH and the CCLP are based on the same underlying principles and intuitions, and so most of my objections to Kamm will apply equally to the CCLP and to either formulation of the PPH. And many of the problems facing Kamm's account are very similar in kind, and in some cases in substance as well, to the problems faced by the DDE and the attempted solution based on the redirecting-old-threat/creating-new-threat distinction.

The first problem is the problem of determining the length of relevant causal chains. Is it always clear when something that serves as a means has the end as its noncausal flipside, and when it causes the end? Of the original Trolley case, Kamm says that "the turning of the trolley is conceived of as a good event in itself, or one very intimately noncausally related to the greater

⁴³This is intended to deal with cases such as Loop.

⁴⁴Kamm, p.183.

good, which causes the lesser evil of the one being killed"⁴⁵. But Kamm seems to be cheating a bit here. It appears to me that the action of diverting the trolley is as closely related to prolonging the life of the five, as it is to the death of the one. If, indeed, the turning of the trolley is noncausally related to the good in Trolley, then that good must be the five's becoming free of the deadly threat. But then it seems only fair to count as the evil the creation of a deadly threat to the one, thus making it the case that the evil too is noncausally related to the turning of the trolley. If, on the other hand, the evil is merely caused by the turning of the trolley, and so the evil must be something along the lines of the trolley's hitting and killing the one at some time t , I think it is only fair that the good in question be the trolley's not hitting and killing the five at t' , where t' is the time at which the trolley would have hit the five, had I not diverted it. But then the good turns out to be causally related to my turning the trolley, as it is my turning the trolley that causes it to be elsewhere at t' . This is not crucial to the application of the CCLP to Trolley, (even though it reemerges as a concern in more complicated cases), since the CCLP only requires that the action be no more intimately connected to the evil than to the good, but it becomes very important in applying the PPH, as the PPH is more restrictive and only allows an action that causes harm, or is itself harmful, if it is noncausally related to the good.

⁴⁵Kamm, p.172.

The problem addressed in the previous paragraph raises an objection of the same kind as the one that came up in the course of a similar discussion about what counts as the same threat - why should answers to questions regarding the true length of some causal chains have any bearing on morality? Kamm says that morally permissible actions should be at least as closely related to the good as to the evil, but why should that be, especially in light of the fact that Kamm admits that is not, all other things being equal, less morally objectionable to create evil through a longer causal chain than through a shorter one⁴⁶?

The more restrictive PPH actually demands that the lesser evil not be brought about by a mere causal means - it should be produced by the good itself, or by an aspect of the good, or by what is noncausally related to the good. But even putting aside the very problematic issue of telling when something is noncausally related to the end and when it is not, I fail to understand the moral significance of this distinction. Surely, it is bad if the means to an end are themselves evil or produce evil, hence the saying that the end does not justify the means. But isn't it at least as bad if the end itself, or something more intimately related to it than means operating causally, is a

⁴⁶Assuming, of course, that the causal role my action plays in bringing about the causally remote outcome is as significant as the causal role of my alternative action with the causally proximate outcome would be - if my going to the store tomorrow inadvertently causes the Earth to be destroyed by aliens in the year 2576, my going to the store tomorrow is still not likely to be a morally impermissible action; but the explanatorily relevant factor will be not the length of the causal chain in itself, but rather the various other causes, e.g. the aliens, that will play a much, much larger causal role in the destruction of our home planet than my going to the store.

source of evil?

The two concerns that I have just raised seem to be backed up by our moral intuitions about cases. Going back to Three Buttons and Four Buttons, which I introduced in the previous section⁴⁷, there just doesn't seem to be any moral difference between the option in which the means causes the lesser harm, and the option in which the end itself causes the lesser harm, or between the former and the option where a causally-twice-removed means causes the lesser harm⁴⁸. But this is not the only way in which the CCLP and the PPH fail to capture our intuitions properly. Pushing C in Three Buttons or Four Buttons is definitely impermissible according to the PPH, and pushing D in Four Buttons is certainly impermissible according to both the CCLP and the PPH, even though, as I've argued earlier, I believe that anyone who claims that it is permissible to turn the trolley in Trolley has to disagree with those conclusions. But now I want to offer some cases where the CCLP, and even the more restrictive PPH, judge an action to be morally permissible when it is clearly not.

Consider the case of the

Transporter: This is just like Transplant, but the manner in which the five would get the one's organs is different - I would be using a Star Trek-like matter-energy transporter. What I would need to do is to enter the appropriate instructions into

⁴⁷Or, alternatively, the variants of those cases in which the harmful options involve different single potential victims, discussed in footnote 24.

⁴⁸Plus, once again, we have the problem of whether the trolley's impact with the iron wall is causally or noncausally related to the good - a problem, the resolution of which, on Kamm's account, implausibly bears enormous moral weight.

the machine, and then, if I push the red button, the following complicated event will happen - instantaneously, the organs of the one will end up where the afflicted organs of the five are right now, whereas the afflicted organs themselves will materialize in the waste basket.

I think it is clear that this case is morally on a par with Transplant, especially if we assume that the death of the one in both cases is equally painful or painless. I believe we have one event here which has a good aspect or noncausal effect⁴⁹, i.e. the saving of the five, and a bad aspect or noncausal effect which is its flipside, i.e. the death of the one. It appears that according to the PPH, and the CCLP as well, it should be morally permissible for me to push the red button, which is, indeed, a bad feature for a purported solution to the Trolley Problem to have.

Kamm could try to argue that pushing the button is noncausally related to the one's death, but only causes the five to be saved, and so both the CCLP and the PPH prohibit it, but, as I've argued with regard to the Trolley case, such an approach seems to involve cheating. Whatever relationship, causal or noncausal, exists between the removal of someone's vital organs and their death, is precisely the relationship that exists between the endowment of someone with replacement vital organs and their being saved. If, however, Kamm responds that both are cases of causing, and so pushing the button in Transporter is permitted by the CCLP, but not the PPH, whichever account of causation she offers to support that contention can easily be used to argue that diverting the

⁴⁹I am not quite sure how Kamm wants to draw the line between the two, but fortunately it is not very important as far as the PPH is concerned.

trolley only causes the five to be saved in Trolley, thus rendering turning the trolley in that case impermissible according to the PPH⁵⁰.

I can think of only one more objection to my Transporter case - it might be argued that my example is too science-fictional and that, in fact, I am just trying to conceal, by using the word "instantaneously", the inescapable truth that the organs have to be removed from the one's body first in order to be placed into the bodies of the five. That would, of course, make saving the five in Transporter impermissible both according to the PPH and according to the CCLP. I am not sure that this is a valid objection, as philosophers frequently tend to cross the boundaries of physical possibility in describing their thought experiments, but that is usually not considered to have much of an effect on the validity of the resulting intuitions, so perhaps I should be allowed philosophic license to reinvent physical possibility in this case as well. However, instead of arguing about this, I might be better off just offering a different example -

Cloning Machine: This is once again very similar to

⁵⁰I can also see someone pointing out that my pushing the button may be considered to be merely the cause of the complex event which I describe in the Transporter case. That may be true, but in any case, as far as the PPH is concerned, Kamm has to be talking only about the last event in the causal picture of the case before the causal chains leading to the good and the bad effects part company; otherwise, whether or not it is morally permissible for me to turn the trolley in Trolley would depend on how causally complicated the mechanism is by which I can divert the trolley, which is certainly a morally irrelevant factor, unless, perhaps, it is something in that causal chain, and not the redirection of the trolley itself, that would cause the death of the one.

Transplant, except this time I have yet another science-fictional device - a cloning machine. Let's assume that the five's afflicted organs were already unusable and had to be removed, but that I possess a futuristic life-support device, which can keep them alive for a few more minutes. The way I can save them is by pushing the blue button on the cloning machine, which will immediately place a functioning cloned version of the one's organs into the five. Unfortunately, it is a fact about the cloning procedure that the original organs become unusable as soon as the cloned copies are in use, therefore the one will die⁵¹.

I think it is clear that this magical case is also morally on a par with Transplant, even though the actual organs of the one are never removed. It is equally clear, especially in light of our little discussion of possible objections to the Transporter case, that pushing the blue button is permissible according to both the CCLP and the PPH. Thus we have to conclude that the PPH is not only too restrictive, but also too permissive.⁵²

There are other problems with the PPH as well. The talk of

⁵¹It may occur to the reader that this case is very similar in structure to Lethal Fumes. That is true, although I think there may be some morally relevant differences. The main reason why I want to introduce the Cloning Machine case in addition to Lethal Fumes, is that while on a pre-theoretical level it may be quite easy for someone to maintain that saving the five is permissible in Lethal Fumes, although not in Transplant, it should be next to impossible to claim that saving the five is permissible in Cloning Machine, but not in Transplant.

⁵²It might also be interesting to note that while the PPH and the CCLP rule out saving the five in Lethal Fumes 2, which I describe in footnote 7, and which is similar to Operation - a case Kamm introduces and discusses on page 151, it is unclear what happens with regard to Lethal Fumes, as the case is underdescribed in the sense that we don't know whether it is the production of the gas itself that saves the five, or whether the production of the gas only causes the five to be saved. Supposing that the relationship between the production of the gas and the saving of the five is a noncausal one, the PPH renders it permissible to produce the gas in Lethal Fumes, contradicting the generally accepted view which apparently doesn't depend on which interpretation we pick.

"structurally equivalent components" is introduced by Kamm in order to justify diverting the trolley in cases like Loop, even though the death of the one is clearly a means to saving the five. The rough idea behind this justification is that we would turn the trolley in Loop or Modified Loop in order to achieve a state of affairs which is structurally equivalent to the good which we are seeking, i.e. a state where the five are free from all of the lethal threats that they are currently facing; and that is permissible, even though we need the one to be killed, and we need his body to stop the trolley in order to maintain the good state of affairs, and even though if the one were not there to be killed we would not turn the trolley, especially in Modified Loop, where turning would mean killing six instead of letting five die should the one's body fail to halt the trolley's progress. It is, however, unclear why we should be convinced by this justification, especially since Kamm's defense of it makes use of two distinctions which we have already found objectionable in the course of our discussion in section II.

The first is the distinction she introduces between turning the trolley because the one will be killed, and turning the trolley in order that the one be killed. That distinction, just like the DDE, relies on facts about the intentions of the agent, and that, as we have seen earlier, is extremely problematic, if not outright disastrous, as part of a principle meant to judge the moral permissibility of actions rather than the moral worth of agents. The second problematic distinction that Kamm relies on, especially for applications of the PPH(2) to multiple or strangely-behaving

threat situations⁵³, is the old-threat/new-threat distinction, which she draws so finely that even the same trolley's coming back from a different direction is to count as a new threat. On the whole, Kamm's explanation of this part of the PPH is less than convincing, and without it she can't account for the commonly accepted intuitions regarding Loop, Modified Loop, and other similar cases.

I have to conclude that the PPH and the CCLP fail on all three of the major counts which discredited both the DDE and the redirection/new-threat distinction as potential solutions to the Trolley Problem - Kamm's principles rely on a distinction, the lines for which are very controversial even in the simplest of cases which we have to consider; it is unclear why the distinction could have any moral significance even if drawn cleanly; and, last but not least, the CCLP and the PPH don't fit our moral intuitions regarding a large group of situations, both in the direction of being too permissive and in the direction of being too restrictive, which is, of course, to be expected considering the difficulties with the distinctions on which the principles are based⁵⁴.

⁵³I have numerous complaints against Kamm's moral intuitions regarding many of these cases to begin with, but I don't think it would be useful to discuss them in this paper.

⁵⁴Kamm makes further modifications to the PPH in her, as yet unpublished, manuscript "Toward the Essence of Nonconsequentialism". I find these modifications highly problematic as well, but since the heart of the PPH remains the same, and the objections against the PPH, as well as against the CCLP, are numerous and, apparently, insurmountable, I think it would not be wise at this time to go into those modifications in detail in this paper. It would also not be prudent to discuss the problems I have with many of Kamm's moral

Section IV.

Judith Thomson and Hypothetical Prior Consent.

In chapter 7 of her book The Realm of Rights, Judith Thomson introduces a very different approach to the Trolley Problem. First, she points out some features of the Transplant case that are not present in Trolley. For one thing, people dying of organ failure presumably have had some time to deal with their situation and put their affairs in order, but the one who would be killed would not get that chance. But also, even more importantly, in our world there is normally a connection between the way we lead our lives and the medical condition of our major organs. It seems utterly unfair, not to mention morally wrong, to penalize someone who has lived a healthy life full of exercise and poorly tasting, but medically inoffensive foods, by killing him and using his organs to save five people who have ignored health concerns in order to avoid denying themselves any of life's pleasures.

It is, however, part of my footnote 1 stipulation for the Trolley case that none of the six people involved are in any way responsible for the fact that the situation in question has arisen. That automatically excludes the kinds of examples that Thomson wants to rule out at the beginning of her chapter, such as cases where the five are not supposed to be on the tracks in the first place, e.g. cases in which the five are thrill seekers who

intuitions - most notably, those regarding the various Lazy Susan cases and the examples involving killing someone to whom the agent is providing supererogatory life-saving aid.

are deliberately ignoring warning signs. It is, perhaps, helpful, although not necessary, to think of the six as track workmen⁵⁵, who are voluntarily doing their job and are aware of the risks that it may entail, but who, nevertheless, have every right to be on the tracks, and who should be saved if there is any morally permissible way to save them.

Thomson argues that there is a very important feature of the Trolley case that has hitherto gone unnoticed. Suppose that the runaway trolley incident in Trolley occurs at 3 P.M. Suppose, furthermore, that there is no specialization required among the group of workmen we are concerned with, and that at 8 A.M. the six draw lots to determine which five of them will work on the main piece of track, and which one will work on the spare piece of track. Each one of the six has an equal $1/6$ chance to be the lone person on the spare track. Under the circumstances, Thomson says, it would be reasonable for all six to agree, prior to 8 A.M., that should a runaway trolley situation arise, the trolley will be diverted to the one, as that would increase everyone's chances for survival from $1/6$, if the situation arises but the trolley is not diverted, to $5/6$.⁵⁶ Thomson argues that this hypothetical consent, which would be given despite the fact that at 3 P.M. there clearly could be one track workman to whose advantage it

⁵⁵Which is why I was calling them track workmen in my original formulation of the Trolley case.

⁵⁶If we look at the trolley threat more generally, without knowing which track the runaway trolley will be on, the chances for survival, if the trolley is not diverted, would be $1/2$ for each one of the six workmen. $5/6$ is still a significant improvement.

would be if the trolley was not diverted, is the reason why it is morally permissible to turn the trolley at 3 P.M.

By contrast, it is unlikely that the six in Transplant, or any other randomly chosen six people, for that matter, would agree in advance to have one of them cut up in order to save the other five from organ failure. For one thing, Thomson says, that is because there is never a time when we are completely ignorant about whether or not our organs are likely to fail in the near future, as normally whether or not our organs fail depends on our past medical history, habits, etc. There is no equivalent of a random drawing to determine who will be among the five and who will be the one, and so there is no equivalent of the time before the drawing at which it would have been reasonable for all involved to agree to a transplant should the appropriate need arise. That is why it is not morally permissible for me to proceed with the operation in Transplant.

There are a few minor details that need to be cleared up before we can proceed to evaluate Thomson's approach. First, is it always morally permissible to turn the trolley in Trolley, provided that the six are all track workmen and that their location for the day was determined by a fair random procedure? Thomson is not committed to that. She says that if some of the six have religious or moral beliefs that require them not to benefit from actions such as the turning of the trolley in Trolley, that may, indeed, affect the permissibility of diverting

the trolley⁵⁷. After all, we would no longer be entitled to conclude that an agreement to divert the trolley, should the need arise, would have been to the advantage of all six.

Of course, in considering cases such as Trolley, we usually assume that the agent is a stranger to all six of the people involved; therefore, presumably, he will know nothing about their religious and moral beliefs. I think we should keep it in the back of our minds as a potential concern, however, that there is a real possibility that in diverting the trolley I might be doing something so objectionable even to one of the five, that it is not to his overall benefit that I turn the trolley - in fact, he might not have consented to it, had I had the chance to ask him. I believe there is a realistic chance that this moral/religious minority may outnumber the few people who would object to having their legs amputated even if that were the only way to save their lives, which is an example that Thomson uses several times in her chapter on the Trolley Problem. However, I think it may be best to table this concern for the rest of this discussion, and just assume that none of the people involved hold such beliefs.

The second clarification concerns the question of whether or not raising the probability of one's survival already counts as a real gain, even for the one that actually ends up dying. Thomson argues that it does, using the example of a man, Edward, whose

⁵⁷This may, of course, raise some unpleasant issues. For instance, if only one of the five has such religious or moral beliefs, can that affect the permissibility of diverting the trolley? This is a tough question, but, fortunately, I don't think there is a real need to consider it here.

life can probably, although not with certainty, be saved if his leg is amputated; if it is not amputated, however, he will definitely die. Thomson points out that it is to Edward's advantage for him to undergo the amputation, even if, in fact, he ends up dying, because the surgery would increase his chances for survival. Similarly, Thomson claims, it is to the advantage of all six in Trolley if an agreement is made at 8 A.M. to turn the trolley, as it would increase everyone's chances for survival. There is a slight asymmetry between the case of Edward and Trolley, as Edward stands to lose nothing if the surgery is performed - without it he will surely die, but one of the six stands to lose a lot, i.e. his life, if I turn the trolley. But I think we should nevertheless accept Thomson's conclusion that it is in a very real sense advantageous for all six to reach an agreement at 8 A.M. that the trolley be diverted, should the need arise.

There is one last detail to be cleared up. Thomson is very careful not to have her criterion be about what someone would have said had he or she been asked. As she points out, if an amputation is necessary and sufficient to save a patient's life and the patient is unconscious, it hardly matters if, should the patient be woken up, he would be so disoriented and in so much pain that he would either not be able to understand any questions posed to him, or would not be able to respond intelligibly, and so no consent would result. What matters in general is not what one or another person would have said had he or she been asked, but the facts which make it the case that it would have been rational

for the person to consent if given the chance.

Of course, we have to be extra cautious here, since if, for instance, Edward does get a chance to voice his opinion in the amputation case, and, especially, if he is in his right mind while doing so; regardless of whether or not it is reasonable for him to say what he does, it seems that we may have to stop what we are doing should he request it⁵⁸. I will not dwell on this point, as there are more important issues to be discussed, but it seems to me that it follows that if someone has, at some point, especially prior to his situation's being determined by random events or procedures, explicitly expressed the desire not to have his chances of survival maximized by a sacrificing arrangement, it may be the case that we cannot use the hypothetical consent approach to justify killing him in order to save five others, should the appropriate circumstances arise.

We are now ready to discuss the merits of Thomson's solution. The first issue that I would like to raise is that even actual consent at the time immediately prior to the action in question doesn't necessarily make the action morally permissible. Thomson concedes as much in saying that "it might be true of a man that he would consent to your slitting his throat in that he has the mad idea that he killed Cock Robin, and deserves throat slitting for it; the fact that he would consent does not make it permissible

⁵⁸Another problem is that it seems that if the six actually discuss the possibility of a runaway trolley situation's arising at 8 A.M. and reject the alternative of diverting the trolley, it should no longer matter whether or not it was reasonable for them to reject it. An actual rejection should, I think, nearly always trump grounds for rational hypothetical consent.

for you to proceed"⁵⁹. The reason for this is that, as I've mentioned before, what's important is not the consent itself, but the underlying facts which would make consent reasonable under the circumstances. Obviously, it is not reasonable for the person in question to consent to having his throat slit, and, equally obviously, it is not morally permissible for me to slit his throat. But let's consider a more sane potential victim.

Sacrifice: Suppose that there is a tribe somewhere which is suffering as a result of severe drought. The elders, as well as the entire tribe, are convinced that the tribe's chances for survival are slim unless the Gods are appeased by an act of human sacrifice. So, the entire adult population of the tribe agrees to draw lots to see who gets to be the sacrifice. Let's suppose that everyone wholeheartedly wants the sacrifice to proceed, including even the unlucky one who loses the drawing, and his or her entire family.

I think that despite the consent, it is still morally wrong for the sacrifice to take place. And, especially, I believe that if the tribe's religion calls for an outsider to perform the ritual in order to make it legitimate, and I happen to be in the neighborhood and am asked to assist, it would be morally wrong for me to oblige this poor misguided tribe⁶⁰.

Now, one may respond that even though the people in this example are not certifiably mad, unlike the alleged murderer of Cock Robin, the facts that would make it reasonable for the one

⁵⁹Judith Jarvis Thomson, The Realm of Rights (ROR), Cambridge, MA - London, Harvard University Press 1990; p.188.

⁶⁰Some people believe that the same is true with regard to the death penalty, i.e. that it is not morally permissible to execute murderers, even if all of society believes that it is beneficial for them to be killed, and even the murderers themselves consent to their own executions.

who lost the drawing to consent to the sacrifice are still lacking - it is not really to the benefit of anyone that he or she die.

But, first of all, that is not necessarily the case - perhaps the belief that Gods have been appeased would restore long lost hope to the entire tribe and make them persevere in spite of the drought. And, secondly and more importantly, I do not see any rational grounds for discounting the moral relevance of the religious beliefs of this tribe, while taking notice of religious beliefs of Christian Scientists and Jehovah's Witnesses, as Thomson does in her discussion of amputation cases. I think that if Edward's being a Christian Scientist is sufficient grounds for it not to be the case that it would be reasonable for him to consent to the amputation, and so a sufficient reason for the surgeon not to proceed if he is aware of the fact that Edward is a Christian Scientist, then in Sacrifice we have a case where it is reasonable for the one to consent to being killed and yet we may not kill him even if he provides actual on-the-spot consent. Of course, I do not wish to place too much weight on this point alone, as the problem may, for instance, be easily, although not altogether comfortably, remedied by retracting the claim that we may not proceed with a life-saving surgical procedure without consent, if we are aware of the fact that the patient is a devout Christian Scientist.

Next, let's consider the following problem. Let's assume that it is reasonable for a father to consent to have his kidney transplanted into his dying daughter, at the very least, provided that that is the only way to save the child's life, and that the

father's life will not be endangered by the transplant. But suppose the father is unconscious and cannot consent; he has also never expressed consent in the past. I think it may be, and, in fact, probably is, morally impermissible to take his kidney. But, even more importantly, even if the father agrees in advance, but then at the last moment changes his mind and refuses to proceed, it is morally impermissible for us to knock him out and perform the transplant anyway. In this case, it is reasonable for the father to have given his consent earlier, and, in fact, he has given it. Furthermore, it is still, presumably, reasonable for him to consent to the procedure at the time when it is supposed to take place, but, because he refuses to consent, we cannot proceed.

To me this suggests that in cases where prior consent is justified, but at-the-time consent is no longer advantageous, if the "victim" protests, we may have to refrain from whatever we were otherwise about to do⁶¹. By Thomson's admission, it is very likely that the one in Trolley would, given the chance, protest at 3 P.M. But the fact that in actuality he does not get a chance to object should not matter, as, according to Thomson, if we know that Frank, who will survive if and only if we amputate his leg, would protest against the amputation if he were conscious and in a clear state of mind, say, because he is a Christian Scientist, we may not proceed, even though Frank does not actually get a chance to protest. As I've already indicated, to me that suggests a

⁶¹At the very least, in cases where the would-be "victim" has not greatly and irreversibly benefited at somebody else's expense in exchange for giving prior consent in the first place. I thank Judith Thomson for pointing out this concern to me.

problem with turning the trolley in Trolley based on the reasonableness of prior consent.

It could be replied that what happens in the father-daughter kidney transplant case is that at first the father considers the life of his daughter to be more important to him than his kidney, and so it is actually on balance good for him if the transplant takes place; but later he changes his mind and finds his kidney to be more important to him than the life of his child, so perhaps it is no longer on balance good for him that we proceed; or, perhaps, it never was on balance good for him that the transplant be performed, and he had just been mistaken for a while. I don't think that any of these are good descriptions of what is going on in this situation. But I can offer a different example, where it should be relatively clear what is on balance good for each participant -

Balloon: Suppose that there is a hot air balloon adrift in bad weather with six people onboard. If all of them stay in the balloon, it will crash and five of the six will die⁶². They are all aware of this, although they don't know who would be the lucky one. They are also aware of the fact that should one of them jump, or be pushed, out of the basket, the balloon will be able to maneuver to safety because of reduced weight.

We seem to be dealing with the exact same spread in Balloon as in Trolley. If all six agree to roll a fair die, with the loser being committed to jumping to his death, each of them increases his chances of survival from 1/6 to 5/6. Suppose they agree to do

⁶²I realize that this sounds somewhat strange, but I'd rather not be forced to come up with a crazy reason why it is that exactly one of the six will survive the crash.

this, but then the one who loses the roll of the die backs out and refuses to jump. I think that if the five push him overboard anyway, it would be a morally impermissible killing at best, murder at worst. Many would, perhaps, disagree with me and argue that it is only fair that the one forfeit his right not to be killed in case he loses the roll of the die. He did, after all, agree to the arrangement of his own free will, knowing that while it would increase his chances for survival, it might still result in his death. I disagree⁶³. But regardless of that, even if actual prior consent can override the one's right not to be killed, it seems clear that mere hypothetical consent, or grounds for such, cannot do so, despite the advantages of consent in terms of probabilities of survival.

In further support of this claim, consider a variation of Balloon -

Ground Control: Suppose that I am on the ground and have access to buttons which operate levers on the balloon with which I can eject any person from the basket. Suppose that if five of the six are to be saved, action has to be taken within 10 seconds, so there is no time for me to discuss the situation with them, even though there is a radio connection between me and the balloon. But I have a list of their names in alphabetical order and a six-sided die. I can quickly roll it, pick the corresponding name off the list, and press the appropriate button.

Suppose that once I roll the die and pick Fred's name off the

⁶³Although the situation can get much more complicated if the one actually accepts some form of highly valuable and irreplaceable, at least under the circumstances, compensation for participating in the arrangement, and then still refuses to jump after receiving his compensation and losing the roll of the die. Thanks to Judith Thomson for directing my attention to this possibility.

list, I can hear Fred saying over the radio that he would not want to risk being sacrificed in order to improve his chances for survival; furthermore, even if he had agreed beforehand, he would have refused to jump had he lost. Can I push the button? I think not. Can I push the button if Fred remains silent throughout, but I have no reason to believe that he would not have protested, had I had the time to tell him what I am about to do? Once again, I think not. Yet, in terms of probabilistic reasons grounding reasonable hypothetical consent, Ground Control is eerily similar to Trolley. It seems that even if we can bypass the lack of current consent in some tradeoff cases where prior consent is present, we can only do it if the prior consent is actual.

Another problem is presented by the Fat Man case. I am assuming that Thomson believes that throwing the fat man off the bridge is impermissible even if he is also a track workman, and if, in addition, his assignment was determined by drawing lots earlier in the day - he got the bridge assignment and five others got the main track assignment. Furthermore, it doesn't seem to be a crucial feature of the case that the one on the bridge is fat. We can either stipulate that all six track workmen are fat, or, alternatively, that even a thin man's body is sufficient to halt the trolley's progress. None of the above-mentioned stipulations seem to change anything that's morally relevant. Yet on Thomson's account, they appear to do so. For under the circumstances, it seems that it would be reasonable for all six to agree, prior to drawing lots, that should the threat arise, the one on the bridge will be thrown off into the path of the runaway trolley, so that

the other five may be saved. Thus it seems that contrary to Thomson's own intuitions, it would be morally permissible, based on her criterion, to throw the one off the bridge⁶⁴.

And, finally, let's go back to Transplant-like cases. Let's begin by taking a look at

Two People Car: Suppose that there are two people in a car that ends up being forced off the road. The resulting injuries are not related to facts about who was sitting where, or to any other decisions made by either of the passengers. Suppose that Fred and Mary are both unconscious, but Fred's life is not in danger, while Mary has suffered massive internal damage and will surely and quickly die unless one of Fred's kidneys and one of Fred's lungs are transplanted into her.

It is probably impermissible for me to operate, considering that Fred's permission cannot be obtained. At the very least, if I proceed, it seems that once Fred finds out what I have done, he will be justified in being upset with me and demanding compensation for the violation of his rights. Yet, if Fred and Mary were asked 5 minutes before the accident what should be done if one of them would die unless a lung and a kidney from his or her unconscious companion were transplanted, it would have been reasonable for both of them to consent to the procedure, as such consent would have raised their chances of survival from 50% to 100%.

Now let's take a look at a variation of Two People Car -

Six People Car: Suppose that, instead of two people, there are six people in the car and five of them have suffered severe internal injuries, which will prove fatal unless organs from the

⁶⁴Kamm expresses similar concerns about Thomson's proposed solution in chapter 6 of Morality, Mortality: Volume II.

one, who is also unconscious, but whose life is not in danger, are transplanted into them, thereby killing him⁶⁵.

Once again, it seems that there are grounds for rational prior consent here. But I think it is clear that I cannot proceed with the transplant. Thomson claims that if we imagine a hypothetical world in which organ failure came upon us suddenly, without warning, without any relation to our lifestyles, or any other choices we have made, and, furthermore, should fresh organs be provided to the "victims", the probability of their failing would be no higher than the probability of any other person's organs' failing; we would likely find that it is permissible to save the five in that world's equivalent of Transplant⁶⁶. I think that Thomson is wrong about this, and that my car accident examples, which seem morally equivalent to that other-worldly Transplant case, support me on this point.

If it can be shown that the other-worldly Transplant case is morally equivalent to Trolley, then so much the worse, I think, for the opinion that it is permissible to redirect the trolley in Trolley. If, on the other hand, there are other features of the cases that make other-worldly Transplant, and our-worldly, but slightly modified, Fat Man, different from Trolley, than it is those features, and not the ones underlying the would-be

⁶⁵I should also add the stipulation that medicine has advanced to the point of being able to rule out organ rejection, and therefore the five will survive as surely as the one will die.

⁶⁶Perhaps I can even help the plausibility of this claim by offering the use of my cloning machine instead of the usual gory cutting and pasting.

reasonableness of hypothetical prior consent, that would need to form the basis of the solution to the Trolley Problem. Thus, even though it is clear that consent-related issues could be morally relevant - after all, if everyone involved, including the one, asks me to kill the one to save the five, it might, under some circumstances, be morally permissible for me to do so, even if it would not be permissible without such a request; I believe that Thomson's hypothetical prior consent approach fails to provide a satisfactory resolution of the Trolley Problem.

Section V.

Back to the Killing/Letting Die Distinction.

As we have seen during the discussion of proposed solutions in sections II-IV, all of them suffer from very serious deficiencies. None of them are generalizable to the entire array of Trolley-Problem-related cases. And most of them rely on distinctions whose moral significance is dubious at best, and which are drawn in very controversial ways to begin with. Should we be looking for a different solution? Well, it certainly would be preposterous to maintain that it is morally permissible to save the five in Trolley, although not in Transplant, but there is absolutely no explanation for this moral phenomenon. So, it seems that if there is no solution to the Trolley Problem to be found, that can only be because the premises of the problem are mistaken,

and either saving the five is permissible in both cases, or in neither one. I will now argue that we have ample reason to think that there can't be a satisfactory solution to the Trolley Problem.

I want to point out that while everyone who has written on the Trolley Problem believes that it is permissible to divert the trolley in Trolley, they all have to draw the line of moral permissibility somewhere short of cases clearly morally equivalent to Transplant. In her early work, Thomson drew the line at introducing a different threat, as in Avalanche or Lethal Fumes. Costa draws the line between Loop and Modified Loop. Kamm's line separates the lone-option equivalent of pushing B in Four Buttons, i.e. Avalanche, from the lone-option equivalents of pushing C or D in Four Buttons. The list could go on for a while. But as we have already seen, none of the above-mentioned red lines are defensible.

Of course, there must be a threshold somewhere if there is a solution to the Trolley Problem. What strikes me as plausible, however, is that everyone who has written on the Trolley Problem experiences great, even if sometimes subconscious, moral discomfort regarding all the cases that we have looked at, in which the agent can kill one in order to save five. Sooner or later, as the cases get more and more objectionable (or only seem to), that discomfort just can't be ignored any longer, and the philosopher arrives at the conclusion that saving the five is morally impermissible in the case being considered, and all cases that are as bad or worse. But by constructing more complicated

examples, or comparison cases such as Three Buttons or Four Buttons, we can see that there is no room for the line of moral permissibility to be drawn there, at least for the authors listed above. In "Thoughts on the Trolley Problem"⁶⁷, John Martin Fischer even constructs a sequence of cases which, he claims, shows that Fat Man and Trolley are morally on a par. I disagree with his argument, but I would not be surprised if a better one could be constructed. I think we should seriously consider the possibility that no line needs to, or can be, drawn in the first place, because saving five by killing one is already impermissible in Trolley.

So, why do I think that the killing/letting die distinction has a role to play in Trolley-Problem-related examples? In the process of her discussion of the Health-Pebble example in "Killing, Letting Die and the Trolley Problem", Thomson makes the following claim: "There is no Principle of Moral Inertia: there is no prima facie duty to refrain from interfering with existing states of affairs just because they are existing states of affairs"⁶⁸. The implication is that there is no overriding moral pressure on us to avoid altering pre-existing distributions of benefits and threats, even if the benefits are potentially life-saving, or if the threats are potentially lethal. While I certainly believe that in some cases we might have no moral duty to "refrain from interfering with existing states of affairs just

⁶⁷Unpublished, printed in EPR.

⁶⁸EPR, p.72.

because they are existing states of affairs", I strongly disagree with Thomson's claim as a general principle. In order to explain why, I will first introduce Thomson's case -

Health-Pebble: There are six men on a beach who are all dying. Five are standing together, and the sixth is on another part of the beach. There is a magical Health-Pebble, capable of curing whichever ailments you have, floating towards the beach. If I do nothing, it will float towards the one and save his life. If I deflect it, it will instead reach the five, and they will be saved, as each of them, unlike the one who needs the entire pebble, only requires one-fifth of the pebble in order to survive.

Thomson argues that, assuming that the one has no more of a claim on the pebble than the five do, it is morally permissible for me to divert the pebble to the five, and thus ensure the allegedly better distribution of it. It is therefore reasonable to add some further assumptions to the case: the pebble does not belong to anyone, I am under no obligation towards any of the six, none of the six are responsible for their, or anybody else's, being involved in this predicament, the six did not determine their locations on the beach by participating in some sort of a lottery with full knowledge that a health-pebble was on the way that at most five of them will be able to use, etc., etc. But, Thomson argues, and this is where that claim that I quoted a bit earlier comes in, the one does not have any more of a claim on the pebble simply in virtue of the fact that he would gain possession of it if I do not intervene.

Thomson proceeds, by analogy, to explain the permissibility of diverting existing threats, as in Trolley, although not of

creating new ones, as in Transplant, or of otherwise doing something to the one in order to produce a better distribution of the good or the threat, such as pushing the one on the beach out of the way, or killing him in order to prevent him from using the pebble. I have to agree that I feel much less of a discomfort at the thought of redirecting the health-pebble than I feel at the thought of diverting the trolley. But the reason for that is that I do not believe Trolley and Health-Pebble to be analogous, even in respects other than that the pebble is a good and the trolley is a threat. If the pebble is unowned and no one has any claim on it, I can, if I want to, take possession of it and make it *my* pebble. After that, I should have the right to distribute that pebble any way I want to⁶⁹.

Most would say that the best thing for me to do is to give the pebble to the five, although I have argued in my previous chapter that the best thing to do in such cases might be to flip a coin, thus giving all six an equal chance for survival. But, regardless of that, if I take possession of the pebble and become, so to speak, its rightful owner, the direction in which the pebble was headed, which determines who would have come into possession of it if I did not, stops being a factor in determining the moral

⁶⁹Although I do believe, as I have argued in the previous chapter, that this is compatible with it being the case that I ought not distribute it in some ways. For instance, supposing that the pebble is only useful for another half-hour, and neither I nor anybody else I could get the pebble to in the next 30 minutes, other than the six on the beach, have any ailments worth mentioning, it is probably morally wrong of me to give it to no one and keep it as part of my pebble collection.

permissibility of actions; just as if I find an envelope with \$1,000,000 in cash, which has just miraculously appeared on the sidewalk and is unowned, should I pick it up, the fact that there is a person walking right behind me who would have taken the envelope if I had not, becomes a non-factor in determining the moral permissibility of my actions concerning the money, even if the \$1,000,000 can save that person's life⁷⁰.

The crucial point here is that while I might, under certain circumstances, such as when the good in question can save somebody else's life, need a good reason in order for it to be morally permissible for me to keep the good for myself⁷¹, even though I have the right to do so, I don't need much of a reason for taking possession of an unowned good in the first place. And if I do take possession of an unowned good, it no longer matters who would have come to own it if I hadn't⁷². But this does not seem to be

⁷⁰Clearly, since the envelope is unowned, I also have every right to kick it to a different spot, or pick it up and throw it on a different sidewalk, even if that changes who will find it.

⁷¹Such a reason is, of course, a lot easier to find with regard to one million dollars than with regard to a health-pebble, the peculiar usefulness of which will be exhausted within 30 minutes.

⁷²A concern may be voiced here about cases in which I can't take possession of the pebble, but can only deflect it. One can conjure up a complex deflection mechanism operated by a push of a button from thousands of miles away. I think that in many, if not all, cases, such deflection activity, provided it violates no one's rights in itself, can be viewed as taking temporary possession of the good - what I am doing is as right or as wrong as literally picking up an unowned good and then, say, handing it over to somebody else of my own free will, even though I have a right not to. But the less plausible such an interpretation appears to be in a particular case, the more it seems that I need a good reason for diverting the good in question - the kind of serious reason that in possession cases I would need if I were to keep the good for myself and let all involved die, and the kind of serious reason that I believe I would need in order to divert

the case in examples where the action in question is diverting a threat, such as a trolley or a nuclear missile. The threat is not an unowned something that I can make mine, even temporarily, and thus acquire the right, let alone the moral permission, to distribute it in any way I see fit. It is unclear what it would mean for me to take possession of a threat like this in the first place, and even if it were, people certainly don't have a right or a moral permission to distribute threats (as opposed to goods) that they own any way they wish. I have the right to dispense my \$1,000,000 any way I see fit, subject to certain constraints such as, perhaps, not using them to fund terrorist activities, but I don't have the right to dispose of the contents of my vial of anthrax just any old way that strikes my fancy⁷³. Thus I believe that the analogy between Trolley and Health-Pebble breaks down, and that some sort of a principle of moral inertia is in effect at least for some threat cases.

The explanation for this phenomenon lies, I think, in the moral significance of the killing/letting die distinction for cases such as Trolley. In order to illustrate why I believe that the distinction is morally relevant in Trolley, I first need to introduce the following variant of Trolley -

One-One Trolley: This case is just like Trolley, but there is

a mortal threat from one group to another. I will discuss this need for reasons in more detail slightly later on.

⁷³On one, I believe quite common, view, it is always supererogatory to give away a good that you rightfully own, but there can and should be very strict moral rules regarding the distribution of threats, even, or perhaps especially, if the threat emanates from something you own.

only one person trapped on each piece of track. Let's call the one who is on the main piece of track, and who will be killed if I do nothing, John. And let's call the one who is on the spare piece of track, and who will die if I divert the trolley, Mike.

As I've just argued, the health-pebble analogy gives us no reason to believe that the alternative of killing Mike is morally equivalent to the alternative of letting John die, even though the two outcomes are clearly equivalent - in both cases one man dies and, by an assumption that we make for all of these cases, John deserves death no more and no less than Mike does. Furthermore, I believe that I owe John's family no explanation for not turning the trolley, beyond relating the circumstances of the case to them and pointing out that I would have needed to kill Mike in order to save John. If they are able to think rationally at this point, there is nothing further for them to inquire about - there is no reason why I should have substituted the death of another for John's death, provided that I am not responsible for the threat, and, if I had done nothing, John would have died.

However, what am I to tell Mike's family if I do divert the trolley? Simply explaining the setup won't do - they will, quite understandably, want to know what gave me the right to substitute Mike's death for John's, given that objectively the outcomes are equally bad, and that John is the one who would have died had I done nothing. I don't think there is anything that I can say, given that people are not interchangeable the way chairs are - I can substitute the destruction of one chair for the destruction of another, at least provided that both chairs belong to me, but it doesn't seem morally permissible for me to do the same with

people⁷⁴, unless I have one hell of a good reason for it, even if both outcomes are in some objective sense equally bad. So, my two alternatives appear to be completely asymmetric - I can only turn the trolley if I have a very good reason to do so⁷⁵, but I do not need a reason to refrain from turning the trolley⁷⁶.

Should I then opt for flipping a coin? After all, I did say that the two outcomes are equally regrettable, so why not be fair and give John and Mike an equal chance for survival? But there would be nothing fair about it, since the two alternative courses of action that are open to me are not, it seems, morally equivalent, even though the outcomes themselves are. It is unclear why John should be entitled to any possibility of Mike's death being substituted for his own. Besides, assuming, as we always do in these cases, that all other things are equal, the universe seems to have already conducted a coin toss and placed John in the path of the trolley - why would flipping a coin make things any more fair⁷⁷?

⁷⁴Perhaps the same applies even to animals.

⁷⁵Perhaps I can also turn the trolley just for fun, provided that it harms no one, and that it is not a violation of anyone's rights, such as those of the owner of the trolley, to turn it for no good reason at all - in cases where no harm can be foreseen, I am certainly not morally required to maintain the status quo for its own sake.

⁷⁶As I've suggested in footnote 72, the same asymmetry appears to apply to distribution-of-a-good cases where the "taking possession" interpretation does not seem at all plausible, if such cases do, indeed, exist.

⁷⁷Heidi Malm gives an argument similar to the one I give regarding One-One Trolley about a different, although relevantly similar, group of cases, e.g. Michael Tooley's "Diabolical Machine" case, in "Killing, Letting Die, and Simple Conflicts" (1989), reprinted in EPR.

In conclusion, I think that it is impermissible for me to either to divert the trolley or to flip a coin in order to determine whether or not to divert it; and I believe that what makes the moral difference is that I cannot kill in order to substitute the deaths of some for the deaths of others, unless I have a very good reason to do so. One final note: I don't even see the Health-Pebble case as a *bona fide* case of conflict between killing and letting die, since it is morally permissible for me to take possession of the pebble, and then the choice is clearly a choice between letting five die and letting one die.

The conclusion that I have just reached still does not entitle me to claim that diverting the trolley is impermissible in Trolley. After all, there may be features of the case that render the killing/letting die distinction morally inert, or, as the explanation of my opponents is more likely to go, the good of preserving four extra lives outweighs the stringency of the injunction not to kill in order to substitute the deaths of one group for the deaths of another - the increment of good provides a sufficient reason to change who lives and who dies. I do not find either of these alternatives appealing. While I do not outright rule out the possibility that the killing/letting die distinction may be morally inert in some, but not all, cases⁷⁸, I do not see what could possibly make it morally inert in Trolley, but not One-One Trolley.

Also, as I have argued in the previous chapter, I do not agree

⁷⁸This possibility will be discussed in more detail in subsection 2 of section VI, as well as in section VII.

that numbers play a role in morality. In particular, I do not believe that the outcome of having saved five lives is in itself morally preferable to the outcome of having saved just one. Thus to me Trolley is morally on a par with One-One Trolley. But this is a very controversial and unpopular position, which I would rather not rely on as a premise. Fortunately, even philosophers who believe that numbers matter considerably in morality do not necessarily think that the numerical asymmetry in itself entitles me to kill a smaller group in order to save a larger one. As Judith Thomson says regarding Transplant in her latest published discussion of the Trolley Problem⁷⁹, "the increment of good the surgeon would produce is insufficiently great" to justify the violation of the one's rights that would be involved in killing him to save the five. This applies equally well to Trolley, as the alleged increment of good is the same.

I agree that the increment of good in itself, even if it exists, cannot justify killing one in order to save five in any case where all other things are equal⁸⁰. Aside from the fact that Trolley and Transplant are symmetric in that respect, but acting so as to save the five is deemed permissible in only one of the two cases, I think that the injunction against killing innocents

⁷⁹Chapter 7 of ROR, e.g. on p.176.

⁸⁰Thomson leaves the door open for the possibility that killing some to save others solely because of the increment of good at stake may be permissible sometimes if that increment is enormous, e.g. if the disparity in numbers is great enough. I do not necessarily disagree, despite my view on numbers in morality - see my discussion of this issue in section VII of chapter 2.

in order to substitute their deaths for the deaths of others is too basic and stringent for the above-mentioned increment to be able to override it⁸¹. Thomson proceeds to suggest a reason why killing the one is permissible in Trolley, i.e. the hypothetical prior consent approach that we discussed in section IV. I, on the other hand, think that there is not and cannot be a reason for me to be morally permitted to kill the one in Trolley to save the five.

It seems to me to be an obvious moral fact that I need one hell of a good reason to cause someone to die by any means, who would not have died, at least not at the time he did and the way he did⁸², without my interference, which is what I inevitably do if I kill someone; and I believe that merely avoiding letting a larger group perish is, all other things being equal, never a sufficient reason. Unfortunately, all I can offer in support of this opinion, which goes against the grain of what the majority believes, are the admittedly intuition-based arguments above, and

⁸¹This isn't to say that it is never morally permissible to kill one in order to save five, but only that the special circumstances, such as that the one is in the process of trying to kill the five, which would need to obtain in order for that to be the case, would also, in a relevantly similar situation, make it morally permissible to kill one in order to save just one, which I think is clearly not the case regarding Trolley's variant One-One Trolley. It is, however, possible that a similar claim is not true if we are talking about killing one in order to save millions, or the whole human race other than the one.

⁸²At least not through the same causal process, which is something we need to add because of cases of overdetermination. Of course, in overdetermination cases the issue might not just be how many people die, but whether I can kill someone, who would otherwise be killed anyway at the same time, in order to save the lives of several others. That is a hard question, which, fortunately, is irrelevant to this discussion because it does not arise in pure tradeoff situations.

the systematic failure of all attempts to come up with moral grounds to distinguish Transplant from Trolley sufficiently for the purpose of pronouncing saving the five in the latter, but not the former, morally permissible. I will spend the rest of this paper trying to remove various obstacles to the plausibility of my claim.

How is it possible, as I maintain, that nearly everyone has the moral intuition that saving the five is permissible in Trolley, whereas that is not the case? I can offer two suggestions. The first is that even though most people who think about it nowadays discard utilitarianism, as well as other purely consequentialist doctrines, the subconscious pull of utilitarianism is still quite strong. When faced with a choice between five lives and merely one, I suggest that the utilitarian pull tends to overwhelm and seduce us, unless the action in question *prima facie* seems so abhorrent that sound moral judgment prevails anyway. The latter is, I believe, the case with Transplant, and even, to a slightly lesser extent, Fat Man, but, unfortunately, not Trolley. The *prima facie* differences between those cases and Trolley, as well as whether saving the five in Trolley is really morally better than saving the five in Transplant, or whether it only seems that way, I will discuss in a moment. For now, I just want to suggest that it is this consequentialist seduction that causes us to think that there is a Trolley Problem in the first place, but those of us who do not believe that numbers matter in morality, are, of course, to a large extent (larger than the rest of us) immune to the

utilitarian virus, and hence, perhaps, in a better position to see the Trolley Problem for what it really is - a very intriguing, provocative, and eye-opening non-problem.

So, what are the differences between Transplant and Trolley, in virtue of which saving the five in Transplant seems much worse than saving the five in Trolley? Most of us think that there are special moral injunctions against the harming of patients, especially healthy ones, by doctors, as opposed to mere bystanders. Also, as was mentioned in section IV, organ failure is often, if not almost always, the result of certain lifestyle choices a person makes. In addition, those who are dying of organ failure usually have advance warning and hence time to get their affairs in order. These do appear to make saving the five in Transplant look worse than saving the five in Trolley. But all this shows is that the original formulation of Transplant is poorly chosen. Everyone seems to acknowledge this, but, in keeping with tradition, Transplant is still always used as the example to elicit our negative intuitive response. I suggest using a bystander version of Six People Car, which is introduced and discussed towards the end of section IV, instead.

Is there anything else? Well, even though death by runaway trolley is probably quite gruesome, it might be much easier to imagine the gruesomeness of being cut up and having one's organs removed, and that may make Transplant seem worse, whether or not it actually is. In addition, in Transplant, your organs, something you have a right to and others don't, are physically removed without your permission in order to save the five. I

think all this means is that we should modify our bystander version of Six People Car to have the agent use my instantaneous organ-cloning machine, patent pending. Let's call this case **Six People Car Cloning.**

Does that remove all of the remaining differences? There is one more factor, briefly discussed at the very end of section II, which we have not eliminated as possibly having moral significance, although we ruled it out as a possible solution to the Trolley Problem by itself, and which perhaps differentiates Trolley and Six People Car Cloning. Even though in the latter case the one's organs are not physically removed, there may be a sense in which we are doing something to the one via doing something to his organs in Six People Car Cloning, but we are only doing something to the threat, i.e. the trolley, in Trolley⁸³. There are, however, cases, such as Handrail and various Mafia examples, i.e. situations in which the Mafia will kill five people if I do not kill someone at their request, in which this distinction is either even less clearly applicable, as in the former, or fails to explain the correct verdicts, as in the latter. And even if this difference does make saving the five in Transplant worse than saving the five in Trolley, it is clearly, as we've discussed before, insufficient to account for the one's being morally impermissible while the other is not.

⁸³This is controversial, however, as it is the cloning that causally or noncausally, depending on the interpretation, makes the one's organs fail. So there might be no way to classify this as actually doing something to the one, without classifying diverting the trolley as doing something to the one as well.

All of these factors, all but one of which (if not every one) can, and probably should be, contrary to accepted practice, equalized for prior to even stating the Trolley Problem, surely make saving the five in Transplant look worse than saving the five in Trolley. Is that really the case? It seems to me that it might be, although I could be falling prey to a psychological fallacy which I will dub "the Jury Error" - an error that consists in resolving uncertainty regarding the defendant's guilt by finding him or her guilty of a lesser charge, even if there is no way, based on the facts of the case, that the defendant can be guilty of the lesser charge without also being guilty of the original, more serious, charge. After all, as a person raised in our society and familiar with much of the literature on the Trolley Problem, I certainly have very few doubts about the impermissibility of saving the five in Transplant, but I can't help having at least some initial doubt regarding the impermissibility of saving the five in Trolley. And it could conceivably be that initial doubt, rather than the facts of the two cases, that makes me think that saving the five in Transplant is morally worse than saving the five in Trolley. Still, I tend to believe that the former action is at least slightly less palatable than the latter, even though I believe that both are, in fact, morally impermissible.

I have staked out a major role in my response to the Trolley Problem for the killing/letting die distinction. I do not believe that it is necessary, even if it were possible within the scope of this paper, to produce a precise definition of the distinction,

as, fortunately, it is clear enough in all the Trolley-Problem-related cases that have come up, what is a killing and what is a letting die. However, as I will, in the next two sections, need to defend the distinction's moral significance from some rather influential arguments, in order to show that it can fill the shoes I need it to fill, certain issues that concern defining the killing/letting die distinction will have to be addressed and clarified.

Section VI.

A Few General Remarks on the Killing/Letting Die Distinction.

In this section I will address a pair of attempts that have been made to discredit the killing/letting die distinction as a morally significant one.

Subsection 1: Jonathan Bennett and the Quantitative Approach.

First, I want to discuss Bennett's attack on the moral relevance of the distinction in his book The Act Itself. Bennett sets out to analyze the role of the making/allowing distinction in morality, but doesn't want to rely on our intuitions about pairs of cases in order to do it. He has doubts about the reliability of our moral intuitions and, furthermore,

"however well such a pair [of cases] succeeds in eliciting certain intuitions, with however little help from extraneous factors such as cost, knowledge, etc., we *still don't know what*

is going on. To know what place making/allowing should have in morality, we need a good analytic understanding of what distinction it is - an account of it in terms that are clear, objective, and deeply grounded in the natures of things. Clear so that we can think effectively, objective so that we can think communally, and deeply grounded so that the issue about moral significance will not be trivialized."⁸⁴

Once we have a "good analytic understanding" of a distinction, it seems to follow that we should be able to analytically examine it and determine whether or not it can matter in morality. That is, in fact, what Bennett does with the making/allowing distinction, declaring it obviously morally inert as soon as he has analyzed it. Clearly, if the making/allowing distinction is morally inert, then so is the killing/letting die distinction, which is nothing more than a particular instance of it. I believe that Bennett is wrong, and I think it will be very illuminating to consider what appears to be the main reason why he's wrong. This will not require looking at either Bennett's analysis or the major criticisms of it in much detail, however, we will need to consider a brief and somewhat oversimplified sketch of Bennett's account.

Before presenting Bennett's analysis of the making/allowing distinction, we need to get a few preliminaries out of the way, so that we can properly understand what the analysis means. First, for Bennett the making/allowing distinction is a distinction that applies to facts, rather than acts, e.g. for him it is the fact that I turn the trolley that is a making, not the act of my turning it. Bennett gives an extensive argument for why it is better to draw the distinction that way, but I see no reason to

⁸⁴Jonathan Bennett, The Act Itself, Clarendon Press, Oxford, 1995; pp.78-79.

discuss it here. Second, the only kind of fact that can be a making or an allowing is a fact about somebody's movements, including, of course, the "null movement" of remaining at rest. Third, Bennett employs a very broad notion of causation. Any fact that is in any way relevant to whether or not another fact obtains, is, on his account, causally related to it. For instance, if a pot of water in my kitchen boils over, for each person who could have gotten to it in time to prevent the contents of the pot from spilling out, had they known about it and tried hard, the fact that they didn't turn off the stove, or remove the pot from it in time is causally relevant to the fact that my kitchen is now a mess; in fact, on Bennett's view, the former is a cause of the latter. There is a lot that can be said back and forth regarding all three of these preliminaries, as all of them are quite controversial, but I think we should just play along with Bennett here, because I believe that it is more than likely that the main idea behind his view can be expressed even without the framework he adopts in The Act Itself.

The following is an approximate statement of

Bennett's Making/Allowing Distinction (BMAD): Suppose that I do A at t⁸⁵, and that causes⁸⁶ an outcome O. Furthermore, the fact that I did A at t is a fact about my movements. Consider all the ways in which I could have moved at t, and call the set of all my possible movements at t my behavior space at t. If the subset of my behavior space at t that contains my movements, the facts about which would have caused O to occur, is larger

⁸⁵By "doing" A, of course, I don't mean to beg the question of whether the fact that I did A represents a positive act or an omission - "doing" A is supposed to cover both possibilities.

⁸⁶In Bennett's sense of "causes".

than the subset of my behavior space at t that contains my movements, the facts about which would have caused O not to occur, then the fact that I did A at t is an allowing of O. If the reverse is true, then the fact that I did A at t is a making of O.

Upon stating roughly this, Bennett immediately concludes that the making/allowing distinction has no moral significance. It does, indeed, seem unlikely, even though not necessarily outright impossible, that the moral status of an act, going back to the way we have been talking throughout this paper, can depend on how many other acts would have resulted in the same outcome, as compared with the number of acts that would have had the opposite effect.

Should we accept Bennett's analysis? Hardly! I think we should in general be worried by any analysis that starts out with "surely by A we can only mean B", where A is a term which is perpetually at the center of philosophical disputes, and B is something no one would ever possibly mean by A. I am inclined to agree that BMAD clearly lacks moral significance, but I think it is obvious that BMAD is not what we mean when we talk about the making/allowing distinction. I sympathize with Bennett's goal that, whenever possible, we should not argue about the significance of a distinction before arriving at an analysis of what the distinction is. But even the alternative that Bennett tries to avoid seems preferable to the alternative of substituting for a relatively clear⁸⁷, albeit only intuitive, distinction, an analysis which obviously fails to capture what we mean by the distinction.

⁸⁷At least in a large number of basic cases.

I do not want to go into too much detail regarding the problems with BMAD as an analysis of the making/allowing distinction. A good and detailed discussion can be found in Jeff McMahan's "A Challenge to Common Sense Morality".⁸⁸ The objection that I want to concentrate on is the complaint that BMAD ignores all background considerations, even the ones that are very relevant to the common-sense making/allowing distinction. As a result, Bennett misclassifies many cases as makings instead of allowings, and vice versa. For our current purposes I want to concentrate on cases like

Baby: I am a single parent and have a 3-month old daughter. I have every opportunity to feed her, but neglect to do so for days. She dies.

Since the number of my movements that would have resulted in my daughter's death is greater than the number of my movements that would not have, i.e. the feeding movements, BMAD would classify what I do in Baby as a case of allowing, on a par with

Starving Children: There are starving children in Africa. I could send money to feed them, but I do not. The children die.

Mind you, the issue is not whether Baby and Starving Children are morally on a par, as there are many differences between the two cases that are irrelevant to analyzing the making/allowing distinction. The question is whether or not the two cases are equivalent with regard to the distinction itself. I think that

⁸⁸In *Ethics* 108 (January 1998):394-418. I agree with much of what McMahan says about Bennett and BMAD, although not with McMahan's own analysis of the making/allowing distinction.

they are clearly not equivalent. It seems very reasonable, in light of what we normally seem to mean by "killing" and "letting die", to think that I merely let the children die in Starving Children, while I kill my baby daughter in Baby. The former is a case of allowing, and the latter is a case of making.

The problem seems to be that no brute counting up of possibilities, or the counting up of anything else for that matter, can capture the meaning of the making/allowing distinction. The reason for this may be that it is altogether wrong to expect the making/allowing distinction to be a merely physical distinction, which nevertheless makes a moral difference. In "Physician-Assisted Suicide: Two Moral Arguments"⁸⁹, Judith Thomson introduces and argues for a different approach to analyzing the distinction - an approach which already includes ethical terms in the analysis. She suggests that whether or not something I do is a letting die depends on whether or not I have a liberty-right to do that thing. This is the criterion that Thomson suggests:

Thomson's Killing/Letting Die Distinction (TKLD): My doing A is a letting die only if (i)the "victim" dies of a pre-existing threat to him or her, e.g. of a pre-existing medical condition; (ii)the victim loses only what he or she would have had with my aid; and (iii)I have a liberty-right to do A.⁹⁰

⁸⁹Cited in footnote 11.

⁹⁰This definition is, of course, incomplete as a definition of the distinction between killing and letting die, since it leaves open what a killing is (thanks to Sarah McGrath for directing my attention to this). We certainly wouldn't want to define a killing simply as a causing to die that's not a letting die, as on most, if not all, accounts of causation, there are many causings to die that are neither killings nor cases of letting die. But, as I've said earlier, for the

(i) and (ii) are quite uncontroversial. Obviously, if the victim was under no threat until and unless I did A, my doing A cannot be a letting die. Equally, even if the victim was under threat, but ended up dying of a clearly different threat, e.g. if I shoot a drowning cancer patient who has just swallowed a bottle of sleeping pills, my action cannot be a letting die. (ii) points out that my doing A can only be a letting die if A deprives the victim only of no more than he or she would have had with my assistance, e.g. if I fail to throw a life-jacket to John who is drowning, that can be a letting die, but if John is struggling to get to a lifejacket that someone else threw in for him, and I go by in a motorboat and remove it, leaving John to drown, that cannot be a letting die. On the other hand, (ii) also reflects the fact that if there is nothing I can do to help the victim, whatever I do cannot be a case of letting die, although it might not be a case of killing either.

But the key to TKLD is (iii), which introduces an ethical term into the analysis. It is important to point out that my having a liberty-right to do A is quite compatible with it being morally wrong for me to do A, e.g. in most cases I would have a liberty-right not to throw a life-jacket to John who is drowning, but, at the very least, if there is one nearby and it costs me nothing to throw it to John, it is morally wrong for me to just stand there and watch John drown. Is introducing (iii) into the analysis

purposes of this discussion we do not need a precise definition of the killing/letting die distinction, but only a rough, though workable, idea of how to draw it - which TKLD certainly seems to provide.

appropriate? Thomson is concerned with the apparent differences between

Doctor: I am a doctor and there is a terminally-ill patient in a vegetative state, who has previously expressed a desire to have his life-support equipment turned off should such a situation arise. After all the appropriate procedures were followed, I was authorized by the hospital and the patient's family to pull the plug. I do so. The patient dies.

and the following two cases -

Intruder: I am an intruder in the hospital. I see a terminally-ill patient in a vegetative state. I pull the plug. The patient dies;

and

Mistaken Doctor: This is just like Doctor except I mistakenly think that I have been authorized to disconnect the life-support equipment. I pull the plug. The patient dies.

It is important to note that all three cases meet condition (i), as the patient dies of a pre-existing terminal medical condition. My actions in Intruder, unlike Doctor and Mistaken Doctor, fail condition (ii), as the patient loses something he would have had independently of my aid. And only Doctor, but not Mistaken Doctor or Intruder, meet condition (iii). This, Thomson thinks, is as it should be - intuitively, Doctor is a case of letting die, but Intruder and Mistaken Doctor are cases of killing, the latter with the best intentions, the former perhaps with the worst. These three cases seem to provide yet another example of how Bennett's ignoring of relevant background considerations leaves him with an incorrect classification of cases.

I am not sure whether I agree with Thomson. Perhaps my actions in Doctor, as well as Intruder and Mistaken Doctor, should be classified as a killing, although, unlike in the other two cases, a morally permissible one. I am also not certain whether conditions (i) and (ii) are entirely satisfactory, or whether more conditions need to be added⁹¹. After all, Thomson's precise formulation of her criterion is intended to deal with only a narrow group of cases - those involving "assisted suicide"; and, perhaps, after we look at many other cases we will realize that (i) and (ii) need to be clarified or amended, or that more conditions need to be added. But I think that in any case, having (iii) as a part of the analysis of the killing/letting die distinction is crucial, as it allows for the proper handling of the differences between cases like Baby, and cases like Starving Children, which BMAD mishandles quite egregiously. I think it is very plausible that the reason why what happens in Baby, but not what happens in Starving Children, is a case of killing, is that all other things being equal, I have the liberty-right not to aid starving children in Africa, regardless of whether or not it is morally wrong for me to do so; but if I have a child, I have a duty to take care of her until she can take care of herself, which means that I do not have the liberty-right to stop feeding her

⁹¹If, in fact, it is true that my actions in Doctor should be classified as a morally permissible killing, that would, indeed, indicate that (i) and (ii) in their current form are insufficient to screen out at least some killings. Of course, even in this hypothetical scenario, it seems that the fact that my actions in Doctor satisfy (iii) would play a major role in making them, but not my actions in Intruder or Mistaken Doctor, morally permissible.

when she is 3 months old and let her starve to death.

I can also invoke condition (iii) to explain my intuition that my not rendering basic medical aid to someone who would die without it, is a letting die if I am a mere bystander, but a killing if I am a doctor, or, at the very least, if I am a doctor on duty. This is not uncontroversial, however, as it may be argued that even a doctor has the liberty-right to not render basic life-saving medical aid, even though it is surely morally wrong for him or her to do so. I disagree, but I find it reassuring, as far as my belief that something like condition (iii) is necessary for any analysis of the killing/letting die distinction is concerned, that it seems intuitively very plausible to me that such a debate would have relevance to whether or not it is a case of killing if a doctor refuses to provide life-saving medical aid to a patient. In fact, should someone offer me a convincing argument for the doctor's having the liberty-right in question, I would likely find myself convinced that the omission in question is a case of letting die rather than killing.

It is important to note that the kind of analysis that offers an ethics-based interpretation of the killing/letting die distinction is not, by itself, sufficient to resolve dilemmas such as the Trolley Problem, even though it makes it clear why the distinction is at least sometimes morally relevant. For one thing, in cases more complicated than Trolley and Transplant, we may find ourselves in a debate regarding whether or not the agent has a liberty-right to perform the action or omission in question - with a condition like (iii) onboard, ethical debates can start

one level earlier. Also, an analysis like TKLD does not settle the question of exactly how and in which cases the killing/letting die distinction is morally relevant, it only points out a more proper way of drawing the distinction in the first place. But I think that especially after considering Bennett's account of the making/allowing distinction, if we believe that the killing/letting die distinction may have some moral significance, the appeal of an analysis like as TKLD, which incorporates an ethics-based condition such as (iii), should be obvious. Nothing else, it seems, can give us even a remote chance of drawing the line where intuitively it should be drawn.

Subsection 2: The Contrast Strategy.

The second major type of argument against the moral significance of the killing/letting die distinction that I want to discuss in this section involves pairs of cases which are purportedly equalized for everything but the killing/letting die factor. Some philosophers have argued, based on such pairs of cases, in which it seems *prima facie* that the actions in question in the two cases are morally equivalent, that the killing/letting die distinction lacks moral significance, for if a distinction plays a moral role, a properly equalized pair of cases differing only with regard to that distinction would have to exhibit some moral difference between the actions in the two cases. I will discuss a few such pairs of examples, as well as whether or not the actions in the members of these pairs are, in fact, morally equivalent, as some authors have claimed, in section VII. What I want to address now

is a serious general flaw in such arguments against the moral significance of various distinctions.

In "The Additive Fallacy" (1989)⁹², Shelly Kagan attacks the strategy behind what he calls *contrast arguments* - arguments for general conclusions regarding the moral significance of a distinction, based on one or several pairs of cases purportedly equalized for all potentially morally relevant factors other than the distinction in question⁹³. Many problems can arise when we use contrast arguments - we can become mired in debates regarding whether or not the pair of cases is properly equalized, whether or not it is even in principle possible to equalize for everything but the factor in question, whether or not the author's moral intuitions are on target, and even, last but not least, whether or not our moral intuitions can ever be trusted without an independently supported moral theory to back them up. But these problems will threaten to come up any time we try to use intuitions about cases to further our ethical knowledge, and it seems that to a certain extent, if we want to get anywhere in ethics, we will need to trust at least a few basic moral intuitions that people tend to agree on, unless reasons for suspicion are presented. The problem that I want to concentrate

⁹²*Ethics* 1998, reprinted in EPR.

⁹³Such a strategy is utilized, at least to a certain extent, by, for instance, Michael Tooley in "An Irrelevant Consideration: Killing Versus Letting Die" (1980), James Rachels in "Active and Passive Euthanasia" (1975), Richard Trammell in "Saving Life and Taking Life" (1975), Bruce Russell in "On the Relative Strictness of Negative and Positive Duties" (1977), all reprinted in EPR, as well as undoubtedly by many other authors writing on the subject.

on, and which Kagan discusses in detail, is the problem of getting from our intuitions about a pair of examples, which lead us to believe that a particular distinction does or does not make a moral difference there, to a general conclusion about whether or not the distinction matters in all other cases.

Anyone trying to argue, for instance, that the killing/letting die distinction is in general morally inert, based on just one pair, or even several pairs, of contrast cases, seems to presuppose something like the principle that Kagan dubs the

Ubiquity Thesis: If a variation in a given factor makes a [moral] difference anywhere, it makes a difference everywhere⁹⁴.

Indeed, without such a principle, the arguments in question cannot get off the ground. But why should we believe that the Ubiquity Thesis is true? Kagan suggests that the belief arises from thinking, quite plausibly, that the fundamental principles of morality have to be universal. Many are tempted to make the leap from this rather uncontroversial claim to the conclusion that a variation in a particular factor cannot ever make a moral difference unless it makes one everywhere, i.e. the Ubiquity Thesis. Kagan gives us a good reason to think that this line of reasoning at the very least *may* be false, by offering the following analogy:

"The presence or absence of oxygen has a role in determining chemical reactions. This role is presumably universal - that is, the fundamental laws of chemistry do not vary from case to case. Yet, obviously enough, the particular effects of oxygen's

⁹⁴EPR, p.258.

presence or absence do vary: in some cases, for example, the presence of oxygen makes a difference to whether or not a compound burns; but it would be a mistake to think that it must make this difference in every case. This is because making a difference to whether compounds burn is not, strictly speaking, part of the role of oxygen in the laws of chemistry. Rather, it is a consequence of that role in particular cases. The mere fact that the role of oxygen in the laws of chemistry is universal does not imply that some particular kind of effect must be universal..."⁹⁵

Thus someone who argues that the killing/letting die distinction never matters in morality because it has purportedly failed to make a moral difference in a particular pair of equalized cases, may be on a par with someone who maintains that the presence of oxygen never plays a role in whether or not compounds burn because in a particular controlled experiment the presence of oxygen has failed to make such a difference.

Kagan suggests that the most natural view for someone who believes in the truth of the Ubiquity Thesis to hold, is the view that incorporates the

Additive Assumption: "The moral status of an act is a function of the various (genuinely relevant) factors that come into play in the given situation ... The overall status of an act is the sum of the contributions made by individual factors ... [Furthermore] the size of a given factor's contribution is determined solely by the value of that factor: variations in the given factor will affect the nature of its contribution; but variations in the *other* factors will *not* affect the contribution made by the factor in question."⁹⁶

This is a very simple and elegant hypothesis about how morality works. Kagan also suggests that what adds plausibility to the Additive Assumption is that people often think about whether or

⁹⁵EPR, p.259.

⁹⁶EPR, pp.259, 261.

not an action ought to be performed by, so to speak, adding up all the moral reasons for and against the action, and balancing them against each other to see whether the positives outweigh the negatives or vice versa.

It is important to note that it is not logically necessary for the Additive Assumption to be true in order for the Ubiquity Thesis to hold. However, I think we have to agree with Kagan that it is completely unclear what someone using the contrast strategy to argue for a general conclusion regarding the moral significance of a distinction, and thus committed to the Ubiquity Thesis, could possibly have in mind as a justification for his strategy, unless he believes that something very much like the Additive Assumption is true. But, Kagan adds, we have been given no reason to believe that the Additive Assumption is true. In fact, he argues, anyone who holds many of our common-sense intuitions about some particular cases will have to agree that the Additive Assumption is false.

Kagan suggests, for instance, that if we accept the assumption that the moral status of an act is determined by some sort of a function of the values of the morally relevant factors involved, the additive model fails to properly capture the roles that seem to be played by the factor of whether or not the agent is acting in self-defense, and the factor of guilt or innocence of a person whose suffering is at stake. Kagan proposes that a multiplicative model might better explain how these two factors affect the moral status of acts. For instance, the self-defense factor can be supposed to take values of 0 in cases of self-defense, and 1 in

all other cases, with the coefficient then multiplied by the value of the killing/letting die factor. This, unlike the additive model, would seem to explain Kagan's intuition that the action of pushing someone who is trying to kill you off a cliff to his death, is morally on a par with not warning such a person that he is about to fall off a cliff, and letting him fall to his death, despite the differing contributing values of the killing/letting die factor.

I will return to the question of whether Kagan is right about self-defense and other "guilty victims" situations in section VII. But, regardless of whether or not he is right about those cases, Kagan appears to have given us some reasons to believe that the Additive Assumption is false, and hence that, most likely, the Ubiquity Thesis is false as well. He has also pointed out that we have been given absolutely no reason, either depending on, or independent of, the Additive Assumption, to believe that the Ubiquity Thesis is true - aside from the fact that perhaps ethics would be a lot more simple and manageable if morality worked the way the Ubiquity Thesis claims it does, there isn't much, at least up to this point, to be said in favor of the hypothesis.

Thus there appear to be absolutely no grounds for taking contrast arguments in favor of general conclusions, such as that the killing/letting die distinction never matters in morality, seriously⁹⁷. This means that we should pay a lot more attention to

⁹⁷Of course, neither can we take seriously arguments for the conclusion that the killing/letting die distinction always makes a difference, if they are based only on our intuitions about one or several pairs of cases.

how various factors that we believe might be morally relevant interact with each other, and why they interact the way they do; instead of just evaluating our intuitions regarding roles of various factors considered in isolation. This should not, however, be taken to mean that according to Kagan, all kinds of contrast arguments should be banned from ethics altogether⁹⁸.

Contrast arguments can be very helpful, in fact, in some cases, perhaps even indispensable, as long as we don't expect to reach unwarranted general conclusions by using them. We can certainly expect to establish, via a properly constructed contrast argument, that a certain distinction is morally relevant in at least some cases, or that it is not morally relevant in at least some cases. Also, even though we have no reason to expect that a given morally relevant factor will always make the same, or even any, contribution to the moral status of an act, if the contribution is different in two cases, or the factor makes a difference in one situation but not in another, there has to be a reason, i.e. the presence or absence of some other feature or features that interact with the given factor in some particular way. So, if no such difference between two cases appears to be present, it seems quite justified, until and unless reasons to believe that there is such a difference are pointed out, to presume that the distinction in question will make the same contribution to the moral status of acts in the cases being considered. This, for instance, is the reason why I feel

⁹⁸Heidi Malm seems to have misinterpreted Kagan that way in her "In Defense of the Contrast Strategy" (Unpublished), printed in EPR.

justified in claiming that since the killing/letting die distinction appears to be morally relevant in One-One Trolley, it is also morally relevant in a similar way in Trolley - the presence of five people on the main piece of track instead of just one doesn't seem to be the kind of factor that should be expected to "neutralize" the contribution of the killing/letting die factor to the moral status of diverting the trolley.

In any case, since we do have reasons to believe that the killing/letting die distinction makes a difference in determining the moral status of acts in at least some cases, it seems clear that contrast arguments cannot really threaten the plausibility of the claim that the killing/letting die distinction is a morally relevant one. Exactly when the distinction makes a moral difference is the topic of the next and final section.

Section VII.

When does the Killing/Letting Die Distinction Make a Moral Difference?

Why think that the killing/letting die distinction sometimes makes a moral difference? First, as discussed in section V, it seems that we need a very good reason in order to change who lives and who dies by substituting one group of people for another, even if the group we kill to save the other is smaller. Second, the idea behind Thomson's analysis of the distinction is very

plausible, and appears to capture our intuitions about what is a killing and what is a letting die quite well. If an analysis like Thomson's, incorporating a condition concerned with liberty-rights of the agent with regard to the potential victims, is correct, it is clear why the distinction would make a moral difference at least in some situations. And, finally, there are pairs of cases where it seems obvious that the killing/letting die distinction is making a moral difference, such as

Lifejacket: John is drowning, but I can save him by throwing him his lifejacket which he had forgotten on the beach. I don't throw John his lifejacket. John drowns;

and

Taken Lifejacket: John is drowning, but somebody threw him his lifejacket, and he is about to get to it. I swim by in a boat and take John's lifejacket before he can reach it. John drowns.

Clearly, my actions in Taken Lifejacket are considerably worse, from a moral standpoint, than my actions in Lifejacket, although I think it is relatively uncontroversial that both actions are morally impermissible.⁹⁹ But, as I mentioned in subsection 2 of section VI, pairs of cases have also been offered in the literature as examples of situations in which the killing/letting die distinction fails to make a moral difference. I think we should now take a look at a few such pairs, and see whether they

⁹⁹Also, for a recent very interesting and thought-provoking argument in favor of the moral significance of the making/allowing distinction, see David McCarthy's "Harming and Allowing Harm" in *Ethics* 110, July 2000, pp. 749-779.

have been properly analyzed.

Consider the following two situations -

Child Crushing: "Smith walks into a room and discovers that a machine that has been set up to crush the child inside has malfunctioned. Smith knows that he could restart the machine by pushing a nearby button, and he does so solely because he is curious to see how flat a person can be;"

and

Child Crushing 2: "Jones walks into a room and discovers that a machine that has been set up to crush the child inside is about to do so. Jones knows he could stop the machine by pushing a nearby button, but he does not do so solely because he is curious to see how flat a person can be."¹⁰⁰

Here is an excerpt of what Malm has to say about these cases:

"Were killing a person in itself worse than letting a person die, then we should see a difference in the moral status of Smith's and Jones' behaviors. But we do not. The acts are clearly both wrong, and Jones, it seems to me, deserves as much moral disapprobation as Smith."¹⁰¹ I agree with Malm - the acts are both morally wrong, and perhaps Jones does deserve exactly as much moral disapprobation as Smith. Why, however, this is supposed to show that Smith's actions in Child Crushing are morally on a par with Jones' actions in Child Crushing 2, I fail to see.

The fact that two actions are morally impermissible certainly doesn't imply that they are morally equivalent. Neither are two actions necessarily morally on a par if the agents of those actions are equally to blame or to praise. As discussed in

¹⁰⁰Heidi Malm, "Killing, Letting Die and Simple Conflicts", EPR, p.135.

¹⁰¹Malm, EPR, p.135.

section II, there are factors which are relevant to whether or not the agent is to blame, and, if so, to what extent, which are, nevertheless, irrelevant to the moral status of the act itself - for instance, the agent's intentions. If the intentions are horrific enough, as surely they are in both Child Crushing and Child Crushing 2, it certainly seems possible for two actions to be on a par as far as the blameworthiness of the agents who performed them is concerned, but not equivalent with regard to the goodness, badness, or moral permissibility of the actions themselves.

Disentangling questions regarding the moral status of agents from questions regarding the moral status of acts should help us see the situation more clearly. However, it might still not become entirely obvious what's really going on, since the badness of the intentions in the two cases may have what has been referred to in the literature as a "sledgehammer effect" - equally, or even just nearly equally, horrific intentions, can, in some cases, overwhelm our intuitions and prevent us from seeing the difference between the moral statuses of the acts; especially if not only the intentions, but the acts themselves, are very bad, and so the difference between their moral statuses may appear to be so small as to be negligible in comparison with the badness of both. Keeping the possible presence of the sledgehammer effect in mind, as well as the necessity of evaluating actions and agents separately, it seems to me to be a very plausible view to hold that even if Smith and Jones are equally blameworthy, their actions are still not morally on a par - Jones' actions are

somewhat worse.

The same explanation, I think, applies to the following pair of cases -

Push: "Gertrude pushes Bertrand into a deep pit, hoping that the fall will kill him so that she will inherit the family fortune";

and

Warning: "Seeing that Bertrand is about to fall into a deep pit, Gertrude deliberately refrains from warning him, hoping that the fall will kill him so that she will inherit the family fortune"¹⁰².

If anything, the explanation seems more plausible here, as well as in this non-comparison example -

Poison: "Two sons are looking forward to the death of their nasty but very wealthy father. Tired of waiting, they decide, independently of one another, to kill their father. The one puts some poison in his father's whiskey, and is discovered doing so by his brother, who was just about to do the same thing. The latter then allows his father to imbibe the deadly drink, and refrains from administering an antidote which he happens to have."¹⁰³

Surely, says Tooley, who appears to be attempting the exact kind of fallacious contrast argument against the significance of the killing/letting die distinction that was discussed in subsection 2 of section VI, the actions of the two brothers are morally on a

¹⁰²Shelly Kagan, "The Additive Fallacy", EPR, p.254; Similar examples are discussed in a similar context by Nancy (Ann) Davis, Judith Jarvis Thomson, as well as undoubtedly many others.

¹⁰³Michael Tooley, "An Irrelevant Consideration: Killing Versus Letting Die", EPR, p.108.

par. But it seems more plausible to me, both in comparing Push with Warning, and in considering Poison, that the agents in question are equally blameworthy, and can to an equal degree be judged to be evil from their actions and intentions, but the actions themselves are still not morally equivalent.

Are there any other pairs of cases, equalized for everything but the killing/letting die distinction, in which the distinction appears, at least at first glance, to be morally inert? What about the two examples which Kagan uses to illustrate the inadequacy of the Additive Assumption -

Active Self-Defense: "In order to defend myself against the aggressor, I push him into a pit, expecting the fall to kill him";

and

Passive Self-Defense: "In order to defend myself against the aggressor, I refrain from warning him about the pit into which he is about to fall, and I expect the fall to kill him"¹⁰⁴?

Assuming that both situations involve legitimate instances of self-defense, i.e. that the aggressor was about to cause grievous harm to me, sufficiently serious to justify doing something that would result in his death, certainly my actions are morally permissible in both cases. I am also equally blameless in Active Self-Defense and Passive Self-Defense. Does this necessarily mean, as Kagan believes, that the two actions are morally on a par? I don't think so. Even though I can't appeal to the

¹⁰⁴Shelly Kagan, "The Additive Fallacy", EPR, p.262.

sledgehammer effect to hedge my bets here¹⁰⁵, I find it plausible that there is once again at least a slight difference in the moral statuses of the acts themselves, even though there is perhaps no difference at all in the way those acts reflect on me¹⁰⁶.

Another very famous pair of cases is the comparison between passive and active euthanasia. I am inclined to believe, provided that proper consent and authorization are present, that it is morally permissible for a doctor to perform either one - at the very least, I am at a loss to come up with a scenario in which one kind of euthanasia is morally permissible, but the other is not. Once again, this does not have to mean that the killing/letting die distinction is morally inert here. For one thing, to make an obvious point, as should already be evident from the self-defense cases, it is not necessary, even for someone who believes that the killing/letting die distinction always makes a moral difference, to believe that all killings are morally impermissible. And it is possible that the same kind of analysis as the one I suggested for the self-defense cases applies here. But also, as I've already mentioned in subsection 1 of section VI, I am not convinced that passive euthanasia situations are cases of letting die in the first place. If they are not, then both passive and active

¹⁰⁵Although the fact that the aggressor is intending to do me grievous harm may have a sledgehammer-effect-like influence in favor, rather than against, the two actions in question.

¹⁰⁶It may be possible, however, that should I have both options of dealing with the aggressor available to me, even though either one would be morally permissible, choosing the passive option would reflect slightly better on my moral character.

euthanasia cases involve killing, which would render the euthanasia controversy irrelevant to the discussion at hand.

Another potentially interesting pair of examples concerns military situations -

Shelling: I am a general in command of my nation's armed forces. The only way to win a brutal war which threatens hundreds of thousands of additional lives, and possibly the very existence of my country, is to carpet-bomb a certain area, as a decoy, then proceeding to attack the enemy from the other direction. Unfortunately, the carpet-bombing would kill 300 of my troops, which I don't have the time to pull back. I proceed anyway.

and

Sitting Ducks: I am a general in command of my nation's armed forces. The only way to win a brutal war which threatens hundreds of thousands of additional lives and possibly the very existence of my country, is to allow the enemy to carpet-bomb a certain area without warning and calling back 300 of my troops, in order to confuse the enemy about which direction my army's attack will be coming from. I do not call back the troops.

This is a hard pair of cases. For one thing, I am not extremely comfortable with the moral permissibility of my actions in either situation. However, I believe that if my actions are permissible in one of these cases, they are permissible in the other; and, furthermore, I am quite likely equally blameworthy or blameless in both scenarios, even though I am inclined to think that there is a moral difference between the moral statuses of the two acts themselves - what I did in Shelling being at least slightly worse than what I did in Sitting Ducks.

Also, even though it is even less likely in Sitting Ducks than

in passive euthanasia cases, I am not 100 percent convinced that my actions in *Sitting Ducks* qualify as a letting die. Certainly, a general, unlike a doctor, does not and cannot take an oath to do soldiers under his command no harm, but he is, nevertheless, responsible for their lives in a very significant way, in which a mere bystander would not be. Thus I am not entirely convinced that, as a general, I have any more of a liberty-right to leave my men to be killed as a decoy when I have a chance to pull them back, than I do to shell them myself. So, once again, I am not even absolutely certain that this pair of examples is relevant to the issue at hand.

I have not, thus far, encountered any two cases that are definitely equalized for everything but the killing/letting die distinction, and in which it is clearly true that the actions of the agents are morally on a par. The analyses in the literature of purported pairs of such cases are badly confused, as they lump together the moral statuses of actions and the moral statuses of the agents that perform them. In all of the pairs that I am aware of, there is some alternative analysis available that appears to me to be at least as plausible as the one that claims that the killing/letting die distinction is not making a moral difference.

Thus I see no basis, at this point, for believing that the killing/letting die distinction is ever completely morally inert, although it is quite likely that even if it is never entirely extraneous, how big of a role it plays depends on other morally relevant factors present. Of course, even if I am wrong about this, and the killing/letting die distinction is sometimes morally

irrelevant, in light of the fact that we have no reason to accept the Ubiquity Thesis, that is not a problem for my view regarding Trolley, as long as the killing/letting die distinction does make a significant moral difference in cases similar to Trolley as far as the other morally significant factors present. And, as I have tried to make clear with my discussion of One-One Trolley, it appears that we have every reason to believe that the killing/letting die distinction is very relevant in Trolley-Problem-related cases. Thus I don't see any grounds emerging from this discussion for doubting my verdict for the Trolley case, and, more generally, no grounds for believing that, all other things being equal, it is ever morally permissible to kill one person in order to save five.

BIBLIOGRAPHY.

Bennett, Jonathan. "Whatever the Consequences." *Analysis*, January 1966, reprinted in ETHICS Problems and Principles, ed. John Martin Fischer and Mark Ravizza, Harcourt Brace Jovanovich College Publishers, New York, 1992, pp. 93-105.

Bennett, Jonathan. The Act Itself. Clarendon Press, Oxford, 1995.

Bennett, Jonathan. Events and Their Names. Hackett Publishing Company, Indianapolis/Cambridge, 1988.

Boorse, Christopher & Sorensen, Roy A. "Ducking Harm." *The Journal of Philosophy*, LXXXV, vol. 3 (March 1988), pp. 115-134.

Broome, John. "Kamm on Fairness." *Philosophy and Phenomenological Research*, vol. LVIII, No.4, December 1998, pp. 955-962.

Chang, Ruth (ed.). Incommensurability, Incomparability and Practical Reason. Harvard University Press, Cambridge, 1997.

Costa, Michael J. "The Trolley Problem Revisited." *Southern Journal of Philosophy*, vol. 24(4), 1986, pp. 437-450.

Costa, Michael J. "Another Trip on the Trolley." *Southern Journal of Philosophy*, vol. 25(4), 1987, pp. 461-466.

Davis, Nancy (Ann). "The Doctrine of Double Effect: Problems of Interpretation." *Pacific Philosophical Quarterly*, 1984, Vol. 65, No.2, pp. 107-123.

Dickey, Ken. "Why I May Count the Numbers." In ETHICS Problems and Principles, ed. John Martin Fischer and Mark Ravizza, Harcourt Brace Jovanovich College Publishers, New York, 1992, pp. 243-250.

Fischer, John Martin. "Thoughts on the Trolley Problem." In ETHICS Problems and Principles, ed. John Martin Fischer and Mark Ravizza, Harcourt Brace Jovanovich College Publishers, New York, 1992, pp. 308-317.

Foot, Philippa. "Euthanasia." *Philosophy and Public Affairs*, vol. 6, N.2, 1977 (winter), pp. 85-112.

Foot, Philippa. "The Problem of Abortion and the Doctrine of Double Effect." In Moral Problems: A Collection of Philosophical Papers, Second Edition, ed. James Rachels, Harper & Row, New York,

1975, pp. 59-70.

Hanna, Robert. "Morality *De Re*: Reflections on the Trolley Problem." In ETHICS Problems and Principles, ed. John Martin Fischer and Mark Ravizza, Harcourt Brace Jovanovich College Publishers, New York, 1992, pp. 318-336.

Jamieson, Dale (ed.). Singer and His Critics. Blackwell's Publishers, Oxford, 1999.

Kamm, Frances Myrna. "Equal Treatment and Equal Chances." *Philosophy and Public Affairs*, vol. 14 (1985), pp. 177-194.

Kamm, Frances Myrna. Morality, Mortality, Volume I: Death and Whom to Save From It. Oxford University Press, New York, 1993.

Kamm, Frances Myrna. Morality, Mortality, Volume II: Rights, Duties, and Status. Oxford University Press, New York, 1996.

Kamm, Frances Myrna. "Feminine Ethics: the Problem of Distance in Morality and Singer's Ethical Theory." In Singer and His Critics, ed. Dale Jamieson, Blackwell's Publishers, Oxford, 1999, pp. 162-208.

Kamm, Frances Myrna. "Reply to Critics." *Philosophy and Phenomenological Research*, vol. LVIII, No.4, December 1998, pp. 963-974.

Kamm, Frances Myrna. "Harming Some to Save Others." *Philosophical Studies*, vol. 57(3), November 1989, pp. 227-260.

Kamm, Frances Myrna. "Toward the Essence of Nonconsequentialism." Unpublished manuscript (2000).

Kagan, Shelly. "The Additive Fallacy." *Ethics* 99(1), October 1988, pp. 5-31.

Kavka, Gregory S. "The Numbers Should Count." *Philosophical Studies*, October 1979, vol. 36(3), pp. 285-294.

Malm, Heidi. "Killing, Letting Die and Simple Conflicts." *Philosophy and Public Affairs*, vol. 18, N.3, 1989, pp. 253-258.

Malm, Heidi. "In Defense of the Contrast Strategy." In ETHICS Problems and Principles, ed. John Martin Fischer and Mark Ravizza, Harcourt Brace Jovanovich College Publishers, New York, 1992, pp. 271-277.

McCarthy, David. "Harming and Allowing Harm." *Ethics* 110, July 2000, pp. 749-779.

McMahan, Jeff. "A Challenge to Common Sense Morality." *Ethics*

108, January 1998, pp. 394-418.

McGrath, Sarah. "Doing/Allowing: What It Is, Whether It Matters." Unpublished manuscript (2000).

Montmarquet, James A. "On Doing Good: The Right and the Wrong Way." *The Journal of Philosophy*, vol. 79(8), August 1982, pp. 439-455.

Nagel, Thomas. The View from Nowhere. Oxford University Press, New York, 1986.

Nagel, Thomas. "Moral Luck." In Thomas Nagel, Mortal Questions, Cambridge University Press, Cambridge, 1991.

Parfit, Derek. "Innumerate Ethics." *Philosophy and Public Affairs*, vol. 7, N.4, 1978, pp. 285-302.

Quinn, Warren S. "Actions, Intentions, and Consequences: The Doctrine of Double Effect." *Philosophy and Public Affairs*, vol. 18, N.4, 1989, pp. 334-351.

Quinn, Warren S. "Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing." *The Philosophical Review* 98 (1989), pp. 287-312.

Rachels, James. "Active and Passive Euthanasia." *The New England Journal of Medicine*, 1975, reprinted in ETHICS Problems and Principles, ed. John Martin Fischer and Mark Ravizza, Harcourt Brace Jovanovich College Publishers, New York, 1992, pp. 111-116.

Russell, Bruce. "On the Relative Strictness of Negative and Positive Duties." *American Philosophical Quarterly*, 1977, reprinted in Killing and Letting Die, ed. Bonnie Steinbock, Englewood Cliffs, N.J., Prentice-Hall, 1979.

Sartorio, Carolina. "Deflecting Trolleys." Unpublished manuscript (2000).

Singer, Peter. "Famine, Affluence, and Morality." *Philosophy and Public Affairs*, vol. 1, N.3, 1972 (Spring), pp. 229-243.

Taurek, John. "Should the Numbers Count?" *Philosophy and Public Affairs*, vol. 6, N. 4, 1977, pp. 293-316.

Thomson, Judith Jarvis. "Rights and Deaths." *Philosophy and Public Affairs*, vol. 2, N.2, 1973, pp. 117-145.

Thomson, Judith Jarvis. "The Trolley Problem." *The Yale Law Journal*, vol. 94, 1985, pp. 1395-1415.

Thomson, Judith Jarvis. "Killing, Letting Die, and the Trolley

Problem." *The Monist*, vol. 59, 1976, pp. 204-217.

Thomson, Judith Jarvis. Realm of Rights. Harvard University Press, Cambridge, Mass., 1990.

Thomson, Judith Jarvis. "Physician Assisted Suicide: Two Moral Arguments." *Ethics* 109, 1999, pp. 497-518.

Tooley, Michael. "An Irrelevant Consideration: Killing Versus Letting Die." In Killing and Letting Die, ed. Bonnie Steinbock, Englewood Cliffs, N.J., Prentice-Hall, 1979.

Trammell, Richard. "Saving Life and Taking Life." *The Journal of Philosophy*, LXXII, vol. 5 (March 13, 1975), pp. 131-137.