

# Dimensions of Moral Agency



# Dimensions of Moral Agency

Edited by

David Boersema

**CAMBRIDGE**  
**SCHOLARS**  

---

**P U B L I S H I N G**

Dimensions of Moral Agency, Edited by David Boersema

This book first published 2014

Cambridge Scholars Publishing

12 Back Chapman Street, Newcastle upon Tyne, NE6 2XX, UK

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

Copyright © 2014 by David Boersema and contributors

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

ISBN (10): 1-4438-6692-X, ISBN (13): 978-1-4438-6692-7

# TABLE OF CONTENTS

<i>Introduction: Dimensions of Moral Agency</i> .....	1
David Boersema, Pacific University	

## **Part One: Moral Agency**

<i>Chaos and Constraints</i> .....	14
Howard Nye, University of Alberta	

<i>The Myth of the Mental (Illness)</i> .....	30
Sarah Vincent, University of Memphis	

<i>The Role of Love in Descartes' Meditations</i> .....	39
Chiara Bandini, San Francisco State University	

<i>Emotions, Ethics, and Equality: Humanity (Ren) as "True Moral Feeling"</i> .....	61
Mary I. Bockover, Humboldt State University	

<i>Epictetus' Serenity Meditation</i> .....	97
Morgan Rempel, University of Southern Mississippi	

<i>Species Egalitarianism and Respect for Nature</i> .....	108
Pierson Tse, University of Calgary	

<i>Berkeley on Human Freedom and Moral Responsibility</i> .....	117
Brandon Bowen, University of Utah	

## **Part Two: Moral Agency and Society**

<i>Aristotle, Virtue, and the Wrong Kind of Reason</i> .....	126
Noell Birondo, Wichita State University	

<i>The Green Problem: How Green Consumption Creates Overconsumption</i> .....	140
Colton Markham, Pacific University	

<i>Habermas and Practical Moral Discourse</i> .....	157
Kari Middleton, Independent Scholar	
<i>Happiness as Resistance: Moral Obligations of the Oppressed</i> .....	164
Kameron Johnston St. Clare, University of Oxford, St. Hilda's College	
<i>Beyond the Dilemma of Desire Satisfaction in Well-Being</i> .....	174
Ryan Michael Murphy, San Francisco State University	
<i>The Ethics of Debt Today: Hegelian Reflections on Abusive Lending and the Financial Crisis</i> .....	182
Kate Padgett Walsh, Iowa State University	
<i>Prostitution Law and Paternalism</i> .....	194
Jeffrey A. Gauthier, University of Portland	
Contributors.....	203

# INTRODUCTION: DIMENSIONS OF MORAL AGENCY

DAVID BOERSEMA

## **Modern Moral Philosophy**

Modern moral philosophy is multi-dimensional in its taxonomy. At one level, philosophers have the tri-partite division of normative ethics, metaethics, and applied ethics. Normative ethics concerns focus on theories of right action and how one ought to live. Standard philosophical theories at this level include those that are goal-based, or consequentialist, such as utilitarianism, with its emphasis on right action as determined by what has utility in maximizing happiness. They also include theories that are duty-based, such as deontologism, with its emphasis on right action as determined by externally set duties or obligations; religious ethics, for instance, are deontological, in the sense of right actions being decreed by a divine being. Yet another form of normative ethics is rights-based, with its emphasis on right action as determined by the securing and exercise of basic (human) rights and a recognition of a moral agent's inherent dignity as a moral agent. Most normative ethical views acknowledge all three components (goals, duties, rights). For example, utilitarianism can and does acknowledge the importance of securing a moral agent's rights as an important factor in maximizing happiness; likewise, deontologism can and does recognize the importance of a principle of utility and of the consequences of actions in the formulation of one's obligations.

The focus of metaethics, unlike that of normative ethics, is on the very nature of moral values rather than on enunciating theories of right action. For example, the question of whether moral values are objective or subjective is a metaethical one. What distinguishes moral values from other sorts of values, such as, say, aesthetic values, is a metaethical concern. Rather than asking, or stating, what right action consists in, metaethics asks what it means for an action to be moral. Where normative ethics speaks to the difference between an action as moral or immoral, metaethics speaks to the difference between an action as moral or amoral.

Other metaethical concerns include what counts as being a moral agent; that is, the kind of being to which morality applies. Much of the moral philosophy of the first half of the 20<sup>th</sup> century was primarily focused on metaethics; for instance, the view that came to be known as emotivism, in which it was claimed that moral judgments are essentially emotional responses. (This was sometimes referred to as the “Boo-Hooray” theory, in which claiming that an act was wrong boiled down to a negative emotional response while claiming that an act was right boiled down to a positive emotional response.) Another point of concern for metaethics is how one can know whether or not an action is right or wrong. Are there moral facts that could determine this? Yet another topic within metaethics is the distinction between what is right and what is good. Both are moral concepts, but they are not the same thing. *Good* is contrasted with *bad* and *right* is contrasted with *wrong*. Good has to do with what is beneficial to, or in the interests of, something, while right has to do with an appropriate choice among options. The nature of the two, and their relevance to one another, is a focus of metaethical concern, as well as of practical social morality (e.g., to what extent and in what ways civil laws should range over the good as well as over the right).

Applied ethics, as the term suggests, is applying normative ethical theories to specific contexts. Such contexts include areas such as bioethics, media ethics, business ethics, sports ethics, etc. The application is often taken as demonstrating how a particular normative moral theory applies to these various contexts; for instance, how a given utilitarian perspective would approach the topic of abortion or capital punishment or euthanasia. However, the “application” can go in the other direction. That is, decisions about abortion, etc., can be and often are taken as points of assessment about normative moral theories, in the sense that assessments about the contextually-situated judgments are used to evaluate the practicability of a given normative moral theory. If, say, a rights-based normative theory failed to elucidate what to do in particular moral conflicts or problems, this can be, and often is, taken as a reason to question or even reject that normative theory. Quite simply, a normative moral theory that failed to provide (reasonable) guidance for moral decision-making is a weak theory, just as a scientific theory that failed to account for specific empirical results is a weak theory.

Besides the tri-partite division of moral concerns, there is another dimension of moral philosophy. It is sometimes portrayed as principle-based ethics vs. character-based ethics (or, virtue ethics). Principle-based ethics emphasizes the enunciation and justification of principles of right action, such as maximizing happiness, performing one’s duties or securing



a moral agent's rights. Principles of action are what matter, with right action being those actions that flow from, are consistent with, or are implied by the identified principle(s). Character-based ethics, however, emphasizes the enunciation and justification of particular moral virtues and vices, such as honesty, humility, courage, loyalty, compassion, etc. Good character is what matters (with right action being those actions that flow from, are consistent with, or are implied by the exercise of moral virtues). Character-based ethics places the emphasis of morality on being a good person, with the assumption that a good person will do what is right; principle-based ethics places the emphasis on right action, with the assumption that even good people can have moral conflict and problems, such that being good will not necessarily inform one of what is the right thing to do vis-à-vis someone else.

Yet another approach to the multi-dimensionality of moral theory is that of considering the metaphysics of morality, the epistemology of morality, and the axiology of morality. The metaphysics of morality refers to the "what" of morality; the epistemology of morality the "how"; and the axiology of morality the "why". That is to say, the metaphysics of morality concerns things such as what moral action is (as opposed to immoral or amoral action), what sorts of things can be moral agents (such as a human, capable of behaving in moral ways) or moral patients (such as non-human animals, capable of being acted upon in moral ways), what constitutes a moral good, etc. Moral epistemology concerns how we know, or believe, what is good or right. Can there be moral knowledge, as opposed to mere opinion? Does the concept of moral knowledge imply that there are moral facts, analogous to scientific knowledge implying that there are empirical facts? What would constitute moral truth or falsehood? Finally, the axiology of morality concerns questions and issues of the value of behaving morally. In a phrase: why be moral? Axiological questions about morality apply not only to normative ethics (again: why be moral?), but also to metaethics (what good is moral theory or moral theorizing?). These various approaches to morality can and do include topics such as the psychology of morality (how we cognitively process behavior that we deem as moral), evolutionary ethics (understanding ethical behavior and theory via the lens of having adaptive survival value), moral agency and identity (how we understand who we are as moral beings), and other lens via which we engage in the world morally.

These emphases on the what, how, and why of morality, however, all ultimately relate to the issue of "who". That is to say, they matter only to the extent that they relate finally to some moral agent. After all, what do we want a moral theory (or principle or virtue) to do, if not provide

guidance to behavior? In the final analysis, the desiderata of moral theory and moral philosophy are not merely to be descriptive, but to be exhortative—to help agents engage in the world. It is moral agency, then, that at the end of the day is the appropriate concern of moral theory and moral philosophy. Moral knowledge, principles of moral justification, and motivations for moral action: these concerns and more matter only because there are moral agents. Explicitly addressing the nature of moral agents, then, is fundamental to any worthwhile moral theorizing.

### **Papers in this volume**

In his essay “How Not to Solve Ethical Problems”, Harvard philosopher Hilary Putnam quipped, “When a philosopher ‘solves’ an ethical problem for one, one feels as if he had asked for a subway token and had been given a passenger ticket valid for the first interplanetary passenger-carrying spaceship instead.” The papers that comprise this volume belie and help to rectify this problem. All but two of the papers included herein were presented at the 65<sup>th</sup> Northwest Philosophy Conference, which was held at Pacific University, in Oregon, USA, in October 2013. They range over a variety of dimensions of moral theory and demonstrate the rich and fecund soil of contemporary moral philosophy.

### ***Moral Agency***

The first topic is moral agency itself; not in a removed, abstract sense, but with respect to lived, experienced agency (such as matters of love, emotion, illness, etc.) The field of moral agency is wide-ranging and overlaps with matters of personhood, free will, belief and knowledge as bases for action, and responsibility, among others. Some philosophical work on moral agency has fairly theoretical emphases, which is true of the papers by Howard Nye and Sarah Vincent. This is not to say that there are not definite practical aspects to them, but their emphasis is on the conditions and criteria of agency. Nye (“Chaos and Constraints”), for instance, argues that all plausible theories of agent-centered constraints against causing harm are undermined by the likelihood that our actions will make the world drastically different than it would have been. Theories that impose constraints against intended harming but none against foreseen harming have unacceptable implications for choices between more and less harmful ways of securing greater goods. Theories that impose constraints against *proximally* caused harm but none against *distally* caused harm have similarly unacceptable implications. This leaves as

plausible only theories that impose constraints against some distally caused harm. Nye argues that, given the dramatic distal effects our actions are likely to have, these theories entail that any way we could live our lives involves unjustified killing, and that any version of them that is strong enough to be plausible entails that we are morally required either to allow ourselves to waste away or kill ourselves.

For her part, Vincent, in her paper, “The Myth of the Mental (Illness)”, wrestles with the issue of moral agency from the perspective of matters related to mentality and mental illness, particularly by engaging with the work of the well-known anti-realist about mental illness, Thomas Szasz. With new research indicating that as much as 25% of the (U.S.) population has been diagnosed with a mental illness, there is perhaps good reason to take the challenge implicit in Szasz and be critical about our concept, *mental illness*. In her essay, Vincent sketches Szasz’s two most provocative papers, detailing his reasons for arguing that mental illness does in fact not exist. She then proceeds to highlight both where she thinks Szasz’s argument is productive and where she thinks it is both empirically and philosophically dubious. Still, she concludes that his argument is best left behind, as an antiquated take on a burgeoning field of medicine. But to avoid stopping short, she goes on to propose what she sees as a more promising alternative to Szasz’s view that there is a myth around mental illness that still takes some of his concerns seriously. There is a myth indeed, but it surrounds the “mental” rather than the “illness”. With new developments in embodied cognition, she urges a revisiting of the question of mental illness from this perspective, to correctly diagnose the problematic myth that must be confronted by the psychiatric community, and to explore what the myth of the mental means for mental illness, including implications for concerns about agency generally, including moral agency.

Related to concerns of both moral agency and moral character are matters that pertain to the nature and scope of our non-cognitive moral experiences. Among such matters are love and emotions. These matters are the focus of essays by Chiara Bandini and Mary I. Bockover. In “The Role of Love in Descartes’ *Meditations*” Bandini presents a reading of the *Meditations* through Descartes’ conception of love and argues that, for Descartes, love supports and facilitates the meditative process by which ideas are rendered clear and distinct in the mind. To ground her position, Bandini reviews William Beardsley’s views on this issue and offers an alternative interpretation of love, one that is consistent with Descartes’ work. Such an interpretation, for Bandini, provides support and justification

that some emotions, most notably love, play a role in the search after truth, extendable to moral truth.

Continuing the conception of the relational nature of morality, Bockover argues that the “feeling” relevant to understanding emotion is an irreducible unity of affect and cognition, in her essay, “Emotions, Ethics, and Equality: Humanity (*Ren*) as ‘True Moral Feeling’”. These emotionally relevant feelings (what she calls ERFs) are cognitive, but cannot be equated with belief; they entail belief, but are not entailed by it and, so, must be distinguished on conceptual grounds. Nor, for Bockover, do ERFs require the experience of specific bodily sensations and, so, are not a combination of cognition and affect, either. In the West, she claims, emotion has historically been misconceived because reason and affect have been treated as independent and often mutually exclusive faculties. Bockover goes on to tie her notion of emotion to a new way of thinking about (gender) ethics by arguing that feeling is moral only when one feels truly that, for example, a wrong exists that dehumanizes a person or group. This intensional affect links us to the humanity of others, and more specifically can give rise to concerns about the parity and fairness (or lack thereof) that can accompany imbalances of power. She further argues that an ethic that focuses on diversity alone cannot adequately account for the good life because it cannot show how equality is linking to human flourishing in general. This moral feeling of equality appeals to a deeper “universal” human reality that becomes key to challenging arbitrary or unjust designations of power and privilege that benefit some at the expense of others.

Along with the concerns about conditions for moral agency, as touched on in the previous papers, Morgan Rempel (“Epictetus’ Serenity Meditation”) suggests revisiting Stoic philosophy as a serious vehicle for applying theoretical explorations about moral agency to very practical concerns of getting along in the world. He claims that one of the things that recommend the comparison of Alcoholics Anonymous and Stoic philosophy is that, like A.A., Stoicism offers real-world guidance for the art of living, the ultimate goal of philosophical theorizing about moral agency. More precisely, it provides practical guidance for living a flourishing, purposeful life of enduring serenity. As his paper demonstrates, a version of A.A.’s basic goal of “peace, patience, and contentment” was articulated by Stoic philosophers centuries ago. Rempel proposes that the parallel examination of several key aspects of these two traditions will serve as a reminder both of Stoicism’s practical message of personal transformation and empowerment, and the enduring, therapeutic wisdom at the heart of Alcoholics Anonymous and the Serenity Prayer.

In a further paper herein that focuses directly on questions of moral agency, Pierson Tse's "Species Egalitarianism and Respect for Nature", attention is shifted from "mere" human agency to placing concerns about such agency within a larger, global and environmental scope. Tse considers agency within the context of our relationships with other species. While recognizing that other species might not be moral agents (that is, capable of acting morally or immorally) and are considered moral patients (that is, capable of being acted upon morally or immorally), he rejects the attitude that is often characterized as: Humans aren't the only species in the world; they just act as if they are. Against such a position, Tse investigates and expounds upon the alternative position of species egalitarianism. Species egalitarianism is the view that all living things have equal moral standing, and thus command equal respect. There is considerable debate over whether or not species egalitarianism is true, and Tse argues that the truth of the matter is not something that can be proven empirically. However, he adds, respect for nature requires the belief in species egalitarianism. Consequently, Tse argues that one should support species egalitarianism. Particularly if long-term sustainability is our desired goal, the only way to reach it is to adopt belief in species egalitarianism.

Unquestionably, an important topic connected to agency that has generated much attention among moral philosophers is moral responsibility. The issue has a long philosophical pedigree, going back at least as far as Plato. We commonly expect people to be accountable for their actions; we hold them to be praiseworthy for good behavior and blameworthy for bad. Nevertheless, questions about moral responsibility have been part and parcel of philosophical analysis from the outset—under what conditions might agents not be (morally) responsible for their actions; when are seemingly bad behaviors excused or at least excusable; how is moral responsibility related to legal responsibility; must moral responsibility entail that a person has committed an action as opposed to a person refraining from acting? This concern about moral responsibility is the focus of Brandon Bowen, who approaches it via a historical investigation of the work of George Berkeley and relevant metaphysical and epistemological underpinnings of moral agency. Bowen argues that in a recent book on Berkeley, George Dicker (*Berkeley's Idealism*) asserts that Berkeley's idealism leads to a lack of moral responsibility. On Dicker's reading of Berkeley, claims Bowen, one is unable to effect action that results in a consequence through a causal chain of events due to the idealist nature of the world. One's will is not linked to action. If one is unable to effect any real change in the world, one could not possibly be held responsible for it. Dicker concludes that according to Berkeley we

have neither human freedom nor moral responsibility. Bowen argues that this conclusion rests on a misinterpretation of Berkeley's view of human freedom and the implications of his idealism. Berkeley's view of human freedom is less about freedom of action and more about the freedom to will. The mind, according to Berkeley, is not material, and is, therefore, not subject to common notions of causation that provide a foundation for the theory of human determinism. Moreover, moral responsibility rests on the will directly, not the ability of the will to effect physical change, though Berkeley does not deny the ability to effect change via the will. Consequently, according to Berkeley a person is both free to will and can be morally responsible.

### *Moral Agency and Society*

Moral agents do not live in vacuums. We are social beings and the sociality of moral agency is central to any serious concerns regarding moral theory and moral practice. One specific issue inherent in the sociality of morality is: Why be moral? That is, why should any given agent be concerned with moral behavior, much less moral principles or justifications? Moral agents, after all, are individuals. Granted, individual humans are social—they are born into and live within social contexts—but why should those individual agents behave in certain ways and not in other ways? The question of “why” has been interpreted descriptively and prescriptively. In the descriptive sense, the question looks for empirical data to account for, say, what motivates agents to behave in certain ways and not in other ways. Or, again, in a descriptive sense the question looks for empirically relevant data to give an account for the prevalence or desirability of certain moral practices and principles over others. For instance, there might be greater survival value for individual organisms if they behave in certain ways and not in others. The point is that understanding “why” in a descriptive sense is to seek explanations for moral behavior and values. However, the question of “why” has also been interpreted in a prescriptive sense; that is, in the sense of seeking justification, not merely explanation. What would justify, say, capital punishment or, utilitarian principles, or enforcement of human rights? For many, the only legitimate or plausible answer to such questions is via the descriptive accounts just mentioned. What justifies, for example, the prohibition against incest, under this view is that such prohibition contributes to the survival of the group. Or what justifies, say, the practice of capital punishment is that it deters future crime. However, these sorts of accounts, for many, are not satisfactory. Why should I—this particular

moral agent—care whether or not there is group survival value to the prohibition of incest or to the permission of capital punishment?

This concern points to the broad background issue of moral realism; that is, the issue of whether or not there are moral facts about the world in the same sense in which there are non-moral facts about the world. Are there, independent of the beliefs or values that individuals might have, facts about the world that correspond to the goodness or badness of an action or a principle? If lying for personal gain is wrong, what makes it wrong (or, for that matter, if it is right, what makes it right?), and, again, why should I, as a particular individual, restrain myself from doing what is wrong or commit to doing what is right? While moral realists say that there are (objective) moral facts about the world, others have argued that the only means we have of determining appropriate behavior—and so, of answering why should I be moral—is via social agreement. Why, and in what ways, I should be moral can only be determined, they say, not by the application of moral principles to specific cases, but rather by a bottom-up approach that identifies morality within the specific contexts in which moral dispute and problems arise. For instance, as American legal theorist, Alan Dershowitz, remarked, rights come from wrongs. Both the concept and the content of civil and human rights did not wait to be discovered, but emerged from how people interacted with one another. This does not in any way imply that there are not relevant empirical facts to how and why moral values and justification of those values emerge, but it is to imply that such facts do not settle the moral matter.

How can—and, for that matter, how ought—one answer the question of why be moral? In “Aristotle, Virtue, and the Wrong Kind of Reason” Noell Birondo suggests that while many discussions in metaethics focus on the nature of ethical reasons and, especially, on the nature of reasons for action, recent discussions have begun to focus instead on reasons for holding various ethically relevant attitudes. Birondo examines one such position, generated from the work of Pamela Hieronymi. The type of reasoning deployed in Hieronymi’s discussion, and in similar contemporary discussions, contrasts sharply, says Birondo, with reasoning deployed on related topics by, for instance, Aristotle. In analyzing and evaluating such current discussion, Birondo claims that the problem is not the wrong kind of reason, but rather the wrong kind of reasoning.

Colton Markham also addresses the issue of why we should be moral, but in a more concrete way, focusing on two approaches—namely, a consequentialist approach and a virtue ethics approach—to what he takes as the problem of overconsumption. In this case, the question of why be moral comes down to why I should consume in certain ways rather than in

other ways, or consume certain products rather than other products. Markham argues that green consumption—that is, the practice of consuming certain kinds of products, based on the belief that we can adequately address and perhaps even solve current environmental problems through consumption of those more earth-friendly products—cannot be accounted for or justified on the basis of consequentialist ethics. For Markham, the true source of climate change is not what we consume, but how much we consume. Consumption-based solutions, including green consumption, are ineffective, and further contribute to overconsumption, because, he says, we mistakenly believe that we can consume more products precisely because those products are earth-friendly. It is only via the cultivation of certain moral habits and virtues that we will adequately address this problem, for Markham, and, more broadly, it is only a virtue-ethics based perspective that will adequately answer the question of why be moral.

The issue of answering the question of why be moral is also one that Kari Middleton investigates. She considers the views of Jürgen Habermas and his claims that the only plausible approach to such concerns is what he calls practical moral discourse. Middleton considers Habermas's position and offers a critique motivated by the moral realist work of John McDowell, but in the end comes down, with some reservation, on the side of Habermas and the bottom-up consensus-generated approach to answering why we should be moral.

Related to virtue ethics, and which particular behaviors are virtuous, is the arena of the broad moral question of living a good life. A long-standing theme within moral theory, and a fundamental focus of normative ethics, is guidance on living a good life and achieving happiness, stretching as far back in Western philosophical moral theory to Aristotle's treatment of *eudaimonia*. The issue of happiness and its relation to living a good life is the subject of the four essays by Kameron Johnston St. Clare, Ryan Michael Murphy, Kate Padgett Walsh, and Jeffrey A. Gauthier. In what could seem to be a surprising take on happiness, St. Clare argues that while victims of oppression have a moral duty to resist their oppressors, in some cases this obligation is met by achieving a state of happiness, and hence undercutting the force of oppression. Indeed, happiness itself is, says St. Clare, a form of resistance.

Where St. Clare's notion of happiness is a form of internal peace, the essay by Murphy ("Beyond the Dilemma of Desire Satisfactionism in Well-Being") also looks at happiness and well-being, but from alternative approaches. Murphy uses the lens of desire satisfaction as the means of considering well-being. Desire satisfaction—the claim that what is best for a person is to satisfy desires—has long been a topic of philosophical



debate, including recently by Robert Nozick, Harry Frankfurt, Derek Parfit, L.W. Sumner, and others. Murphy suggests that the disparate views and claims by philosophers about this issue result from pursuing different theoretical objectives. So, desire satisfaction formulated as a theory of well-being is at odds with desire satisfaction articulated as a welfarist theory. Murphy concludes by claiming that the topic is still very much unsettled and alive, but that these two conceptions are fruitful toward attaining a fuller understanding of desire satisfaction, and, by implication, of what constitutes a good life.

The final two essays are the most immediately focused on the sociality of moral agency, both in the sense of our very nature as (moral) agents, being ones that are inherently concerned with matters of interrelations with others, and also as being housed within moral matters that emerge because of our social interactions. In “The Ethics of Debt Today: Hegelian Reflections on Abusive Lending and the Financial Crisis”, Padgett Walsh specifically considers the economic conditions within which we live as constitutive of our moral practices and agency. In particular, she looks at the recent global economic downturn as a venue for evaluating the interrelations between individual moral agency and social practice and policy. The 2008 mortgage crisis, she argues, brought to light many ethically questionable lending and borrowing practices. As we continue to learn about what caused this crisis, it has become urgent that we think more carefully about conditions under which loans can be ethically offered and accepted, but also about when it might be morally permissible to default on debt. Padgett Walsh examines two standard philosophical approaches to assessing the ethics of debt and default. Both approaches, she claims, are impoverished because they focus only on individual borrowers and lenders. Both approaches thus overlook the real importance of broader social and economic factors that directly caused the crisis. Only by taking a wider view of the matter can we fully understand the moral dimensions of debt and default today.

Finally, Gauthier, in his “Prostitution Law and Paternalism”, looks at the issue of moral agency within the context of social mores and legal practice. He notes that liberals and feminists have long criticized the paternalistic approach to prostitution found in most jurisdictions in the U.S. In his recent book *Prostitution and Liberalism*, Peter de Marneffe defends just such an intervention, arguing that the demonstrated harmfulness of a life of prostitution justifies paternalistic policies aimed at reducing the number of women who are involved in it. Although de Marneffe does not endorse the prohibitionist approach typical in the U.S., he argues that the best reasons for alternative approaches to the practice

(including some forms of regulated legalization) are necessarily paternalistic. In his essay, Gauthier questions de Marneffe's contention that the strongest reasons for state intervention with regard to prostitution are paternalistic in nature. Rather, he argues that reasonable state action toward prostitution is best understood not as a paternalistic intervention to remedy some moral or epistemological failure on the part of prostitutes, but rather as an attempt to advance the interests of vulnerable parties more generally concerning what they reasonably desire but could not otherwise ensure. Further, Gauthier argues that such an approach might favor abolitionist over regulatory policies, depending upon how the vulnerable class is defined.

**PART ONE:**  
**MORAL AGENCY**

# CHAOS AND CONSTRAINTS

HOWARD NYE

## Introduction

Chaos theory tells us that our world exhibits “sensitive dependence on initial conditions,” or that small changes at any point can lead to dramatically different outcomes. These have become known as “butterfly effects,” after Edward Lorenz’s vivid example of “the flap of a butterfly’s wings in Brazil set[ting] off a tornado in Texas.”<sup>1</sup> Laura Cannon suggested that the pervasiveness of these effects might be morally important when she “consider[ed] the plight of Lorenz’s butterfly,” and “wondered how a butterfly might feel if it had the mental capacities to comprehend Lorenz’s discovery. What sense of responsibility might it feel, knowing that its movement might be the cause of great suffering? Might some butterflies sit paralyzed on the branch in an attempt to avoid being the cause of such harm?”<sup>2</sup>

Cannon, following Samuel Scheffler (1995), claimed that the far-reaching effects of our economic decisions create trouble for commonsense notions of responsibility. But I believe that Cannon and Scheffler underappreciated the pervasiveness of the problem that butterfly effects pose for views according to which we have stronger moral responsibilities to avoid causing harm ourselves than we have to prevent harms that would occur in the absence of our interference. In this paper I argue that, given the harmful butterfly effects our actions are likely to have, all plausible theories of agent-centered constraints on harming entail that we *must* sit paralyzed—or kill ourselves—in order to avoid causing harm.<sup>3</sup> I believe that this result is extremely important for ethical theory. Shelly Kagan, Frank Jackson, and Michael Smith have argued that *absolute* constraints against inflicting serious harms on innocents regardless of the benefits of doing so lead to paralysis under conditions of risk.<sup>4</sup> But to many of us, the most plausible constraints on harming are not absolute. Even if inflicting harm is significantly harder to justify than failing to prevent harm, it seems that we should still be permitted to painlessly kill one individual to save the rest of the world’s billions from dying the most excruciating deaths imaginable. I shall argue, however, that the likelihood that our actions will

have dramatic butterfly effects undermines all plausible *non-absolutist* understandings of constraints, according to which it is harder but not impossible to justify inflicting serious harms.<sup>5</sup>

## 1. An Overview of the Argument

My argument begins with the observation that for any way we might sustain our lives, we will have to perform some set of actions, *A*, of which it is reasonable to believe that some members will have butterfly effects. As chaos theory shows, it is not just the consumption choices of westerners in a global economy that can lead to dramatically different outcomes. The most non-intrusive existence that could sustain our lives—say that of a hermit who expends minimal effort tending his garden before dutifully returning to the fetal position—will run a far greater risk of causing a dramatic cascade of events than Lorenz’s butterfly. The minimal life-sustaining actions our hermit must perform, repeated millions of times over the course of his life, will make it virtually certain that his actions will somewhere make things dramatically more different than they would have been had he not performed them. In fact, it is reasonable to believe that his actions over the course of his life will have many such effects.

Because butterfly effects result in such dramatic events as tornadoes, it is reasonable to believe that at least some of these effects of *A* will make some individuals worse off and others better off than they would have been had *A* not been performed. Call the former the “butterfly effect harms of *A*,” or *BH(A)*, and call the latter the “butterfly effect benefits of *A*,” or *BB(A)*. It is, in particular, reasonable to believe that for any acts *A* that could sustain our lives, there will be at least some deaths in *BH(A)* and some life-savings in *BB(A)*. Such are the results of causing and preventing such momentous events as tornadoes. There is, however, no reason to believe that *BB(A)* will tend to be either greater or less than *BH(A)*. As such, it is reasonable to expect that *BB(A)* will on average be about equal to *BH(A)*.

If the beneficial upshots of our conduct were able to justify the harmful upshots, so long as the benefits were equal to or greater than the harms, then the expected butterfly effect benefits in *BB(A)* would exactly justify the expected butterfly effect harms in *BH(A)*, and we could—as it seems we should—treat the unpredictable butterfly effects of our actions as something we can ignore for practical purposes. But agent-centered constraints on harming hold that some harmful upshots of our conduct cannot be justified by equal or somewhat greater benefits. These constraints hold, for instance, that the benefits of saving five individuals

from dying of organ failure cannot justify the harms we would cause to one healthy individual by removing her organs and transplanting them into the five.<sup>6</sup> But if the harms our conduct inflicts on some cannot be justified by the equal or somewhat greater benefits it brings to others, there is a serious worry that for any course of action  $A$  that would be needed to sustain our lives,  $BH(A)$  cannot be justified by  $BB(A)$ .

One way to prevent butterfly effects from making trouble for agent-centered constraints on harming would be to claim that these constraints make it difficult to justify inflicting only those harms we *intend*. I will argue that this is implausible in Section 2, but for now I simply note that most proponents of agent-centered constraints hold—plausibly—that the benefits of saving five cannot justify certain ways of causing foreseen but unintended harms to one, like running her over if this is the only way to reach the five in time to save them from drowning.<sup>7</sup> Another way to prevent butterfly effects from making trouble for constraints on harming would be to claim that these constraints make it difficult to justify actions only if they are “proximate” causes of harm, or if we can foresee who the victims of these actions will be. I will argue that this is implausible in Section 3, but for now I simply suggest that most proponents of constraints will hold—plausibly—that somewhat greater benefits cannot justify certain ways of causing harm distally or to unknown victims. These would presumably include saving five by pulling a trigger that sets off an elaborate Rube-Goldberg device that kills one, or sets off a device that fires thousands of rifles at thousands of victims, an unknown one of which is loaded with live ammunition.

If this is right, then any plausible theory of agent-centered constraints on harming will hold that the infliction of certain unintended distal harms cannot be justified by equal or somewhat greater benefits. I will argue in Section 4 that because of the drastic nature of butterfly effects, it is reasonable to believe that some of these difficult-to-justify harms will be among the butterfly effect harms in  $BH(A)$ . Because each benefit in  $BB(A)$  is needed to justify a corresponding harm of equal magnitude in  $BH(A)$ , this will mean that some deaths in  $BH(A)$  will remain unjustified by life-savings and other reasonably expected benefits in  $BB(A)$ . So all plausible theories of constraints on harming will hold that for any way we could sustain our lives, the butterfly effects of our actions can be expected to kill others in ways that cannot be justified by the lives they will save. Moreover, I observe in Section 5 that on any plausible theory of constraints on harming, the benefits of saving oneself and  $N$  individuals are insufficient to justify killing  $N$  other innocent individuals in ways that are difficult to justify. If this is right, then all otherwise plausible theories

of agent-centered constraints on harming entail that we are morally required either to allow ourselves to waste away or kill ourselves. This, I argue, undermines the plausibility of agent-centered constraints on harming.

## 2. Constraints Against Only Intended Harm?

Once we know about butterfly effects, we can foresee with reasonable certainty that if we act to sustain our lives, we will make some individuals worse off than they would have been. As I mentioned, one way to deny that this makes trouble for agent-centered constraints on harming is to insist that these constraints make it difficult to justify causing only those harms we intend, while harms we merely foresee can be justified by the equal or slightly greater benefits of causing them. Although a few authors have suggested something like this view of agent-centered constraints,<sup>8</sup> I think they fail to appreciate how unattractive it really is. If we read “intending harm” literally, then you need to intend a victim to suffer the harm of death only if her dying plays a causal role in what you aim at. But then a prohibition against causing only intended harm would permit you to save five individuals dying of organ failure by harvesting the organs of one healthy individual while she is alive, since her dying as a result of your removing her organs would be a byproduct that plays no role in saving the five.<sup>9</sup> This would seem to undermine the entire motivation for believing in agent-centered constraints on harming.

If, on the other hand, we interpret “intending harm” a bit less strictly, as something like intending a harmful effect on someone’s body or intending someone to instantiate a property that ends up harming her, then you must intend harm in harvesting the organs of one to save five, although you need not intend harm in driving over one individual trapped in the road to save five others from drowning.<sup>10</sup> But because it seems about as abhorrent to knowingly run over one to save five as to harvest her organs to save them, this understanding of constraints against only causing intended harm, which prohibits the latter but permits the former, also seems to undermine the entire motivation for believing in agent-centered constraints on harming.

To appreciate the absurdity of such a theory of constraints, consider the following cases:

*Less Harmful Transplant.* You have two ways of saving five individuals from dying of organ failure: (1) remove the organs of one healthy individual and transplant them into the five, or (2) run over four healthy

individuals who are trapped in the road that you would need to drive over to get organ-failure-preventing drugs to the five.

*Less Harmful Terrorism.* The only way to save five from being killed by a cannon is to drop bombs that will have two effects: (a) destroy the cannon's ammunition and (b) demoralize the terrorists operating the cannon into surrendering by killing some of four innocent bystanders they care about, where both (a) and (b) would be sufficient by itself to save the five. Suppose that if you (1) drop your bombs with an intention of killing one bystander, a mind-reading demigod will shield the other three from your bombs and you will kill only one, but if you (2) drop your bombs with the intention merely of destroying the ammunition, the demigod will leave the three unshielded and you will kill all four.<sup>11</sup>

A theory according to which there are constraints against inflicting only intended harms (which are strong enough to make it wrong to harvest one individual's organs to save five) would in these cases tell us to take option 2 and save the five by killing four individuals instead of only one.<sup>12</sup> But the mere fact that by killing the four we could avoid having a problematic intention towards the one is a preposterously narcissistic justification for killing three additional individuals. If we are not permitted in these cases to inflict lethal harm on the one with the intention of doing so, then we cannot be permitted to do what we foresee will certainly kill the four either.<sup>13</sup>

### **3. Constraints Against Only Proximal or Identifiable Harm?**

I have thus argued that plausible theories of agent-centered constraints on harming must apply these constraints to some harms that are foreseen but unintended. As I mentioned, one might still deny that constraints on harming apply to harms caused by butterfly effects by holding that these constraints make it more difficult to justify actions only if they are "proximal" causes of harm, or one can foresee who the victims of the actions will be.<sup>14</sup> But it seems quite implausible that how proximally an action causes harm, or whether one can identify its victims, matters *in itself*, quite independently of this indicating a greater risk of causing harm. Consider:

*Greater Distal Harm.* You have two ways of saving five from drowning: (1) drive straight, which will kill one when you drive over a platform, the



depression of which will crush her, or (2) take an alternate road, which will kill four when you drive over a different platform, the depression of which will crush them. But while the four are located directly under the second platform, the depression of its top half will kill them by setting off an elaborate Rube-Goldberg device with hundreds of causal intermediaries that will eventually cause the downward movement of its bottom half, which will crush them.

*More Unknown Victims.* You have two ways to prevent the sadistic dictator Pedro from shooting five innocents: (1) shoot one other innocent yourself, or (2) press a button that will select four other innocents from a databank of everyone in the world and send reliable kill-bots after them, which you know with certainty will kill the four.

A theory of constraints that applied only to proximally caused harms or harms with known victims (and was strong enough to make it wrong to drive over one to save five) would in these cases tell us to take option 2 and save the five by killing four individuals instead of only one.<sup>15</sup> But the mere fact that we would kill the four by a longer sequence of causal intermediaries or that we don't know who they will be are ridiculous reasons to kill three additional individuals. If we are not permitted in these cases to proximally cause the death of the identifiable one, we cannot be permitted to distally cause the death of the possibly unidentifiable four either.

#### 4. Plausible Constraints Against Distal Harm

I have thus argued that plausible theories of agent-centered constraints on harming must hold that they apply to causing some merely foreseen distal harms to unknown victims. But these theories are directly vulnerable to the butterfly effect argument. I should emphasize that there are many forms such theories can take. Some will hold that it is difficult to justify harms so long as they counterfactually depend on events that constitute our actions.<sup>16</sup> These will make *all* the harms in *BH(A)* difficult to justify, and consequently unjustified by the roughly equal benefits in *BB(A)*.

Other theories of constraints against merely foreseen, distally caused harm will hold something more like the view that it is difficult to justify a harm if one's actions *produce* it, or there is a continuous, transitive chain of causal events linking one's action to the harm.<sup>17</sup> Since some harms in *BH(A)* will not be produced by A, these theories may allow some harms in *BH(A)* to be justified by equally great benefits in *BB(A)*. But it is

reasonable to believe that for any act  $A$  that will sustain our lives, *some* lethal harms in  $BH(A)$  will be produced by  $A$ —this, after all, is the way butterflies’ wings kill the victims of the tornadoes they set off. Moreover, since the expected benefits in  $BB(A)$  are equal to the expected harms in  $BH(A)$ , for  $BB(A)$  to justify  $BH(A)$ , each harm in  $BH(A)$  must be justified by a corresponding benefit of equal magnitude in  $BB(A)$ . In particular, each harm in  $BH(A)$  that is produced by  $A$  must be justified by a corresponding benefit of equal magnitude in  $BB(A)$ —since all other benefits in  $BB(A)$  are already needed to justify the harms in  $BH(A)$  that are *not* produced by  $A$ . So, since theories which posit constraints against producing distal harm entail that these harms in  $BH(A)$  that are produced by  $A$  cannot be justified by equal or somewhat greater benefits, they too will entail that the benefits in  $BB(A)$  cannot justify the harms in  $BH(A)$ .

The same logic applies to more elaborate theories of constraints against causing distal harm. Some, for instance, will have “distributive exemptions” for actions that cause harm to some individuals with the same materials or forces that would have caused harm to others had they not been performed.<sup>18</sup> While some lethal harms in  $BH(A)$  may be caused in this way, the pervasive nature of butterfly effects makes it reasonable to believe that there will be other lethal harms in  $BH(A)$  that are *not* caused in this way (like deaths caused by tornadoes that wouldn’t have formed had one not acted), making it impossible for  $BH(A)$  to be entirely justified by  $BB(A)$ . Other theories may hold that benefits cannot easily justify harms if elaborate explanatory relations accumulate between them. For instance, Frances Kamm proposes that a benefit cannot easily justify causing a harm if “something—[the] means [to the benefit]—brings along with it causes [an effect on a victim that harms her] either directly or by overlapping with the direct cause of [this effect].”<sup>19</sup> But *whatever* one takes the distally caused harms to be that are difficult to justify, the drastic nature of butterfly effects makes it reasonable to believe that some of these will be in  $BH(A)$ , making it impossible for  $BB(A)$  to entirely justify  $BH(A)$ . For instance, it is reasonable to believe that some of your actions will cause lethal disasters, but that some aspects of these disasters will cause lives to be saved later on—which fits Kamm’s criterion for actions, the lethal harms of which cannot be justified by their life-saving benefits.

## 5. Is this an Argument for Universal Suicide?

I have thus argued that, given the likelihood that our actions will have butterfly effects, all plausible theories of agent-centered constraints on harming entail that for any way we could sustain our lives, it is reasonable

to expect that it will involve at least  $N \geq 1$  instances of difficult-to-justify killing, which are not justified by the corresponding  $N$  lives that it can be reasonably expected to save.<sup>20</sup> Now any theory of agent-centered constraints strong enough to make it wrong to kill one individual in the difficult-to-justify way to save five others will entail that we are not permitted to kill  $N$  individuals in the difficult-to-justify way to save  $N+1$  others. Moreover, any plausible theory of constraints on harming will apply them to cases where we would be one of the beneficiaries of the harming—and consequently will not permit us to kill  $N$  individuals in the difficult-to-justify way to save  $N$  other individuals and ourselves. If, for instance, we are morally prohibited from harvesting the organs of one person to transplant into two others, then surely we remain morally prohibited from doing so if we are one of the two who need organs. So any theory of agent-centered constraints on harming that is strong enough to be plausible will not permit us to perform a set of acts that would sustain our lives, even though this would save our lives and the lives of  $N$  others, if it would in a difficult-to-justify way kill a different group of  $N$  individuals. Since, as I have argued, on any plausible theory of agent-centered constraints, the butterfly effects of our actions make it overwhelmingly likely that any set of acts that would sustain our lives *will* involve at least  $N$  instances of killing that are difficult-to-justify and consequently unjustified by the fact that it will save our lives and those of  $N$  others, any otherwise plausible theory of agent-centered constraints will not permit us to sustain our lives. It will require us either to let ourselves waste away or kill ourselves.<sup>21</sup>

If I am correct that, because of the likely butterfly effects of our actions, all otherwise plausible theories of agent-centered constraints on harming require us either to allow ourselves to waste away or kill ourselves, what should we conclude? Should we conclude that there are no agent-centered constraints on harming, or should we conclude that we are in fact morally required to let ourselves waste away or kill ourselves? A moral theory should not be dismissed simply because it entails that, given our contingent circumstances, we are all morally required to let ourselves die or kill ourselves.<sup>22</sup> But there seems to be something absurd about the view that we must waste away or kill ourselves simply because of the harms that would be wrought by the unpredictable butterfly effects of our actions, when we can reasonably expect these butterfly effects to prevent comparable amounts of harm, and we can live our lives in ways that are predictably beneficial to others and can consequently be expected to do more good than harm on the whole.<sup>23</sup> The implausibility of the idea that the unpredictable harms our lives are likely to cause cannot be justified by

the fact that that our lives are likely to prevent even greater harms seems to illustrate the direct implausibility of the view embodied in otherwise plausible constraints on harming—namely that such factors as whether a harmful upshot of our conduct was produced by our actions or would have occurred in their absence make a significant intrinsic moral difference.<sup>24</sup> So I think we should continue to believe that we are *not* morally required to waste away or kill ourselves, and conclude from my argument that there are no agent-centered constraints on harming.

The great irony is that it is the view that killing is worse than letting die, rather than the view that letting die is just as bad as killing, that seems to make morality too demanding.

## References

- Aboodi, Ron, Adi Borer, and David Enoch. 2008. "Deontology, Individualism, and Uncertainty: A Reply to Jackson and Smith." *The Journal of Philosophy*, 105: 259-272.
- Bennett, Jonathan. 1981. "Morality and Consequences." In S. McMurrin (ed.), *The Tanner Lectures on Human Values*, Salt Lake City: University of Utah Press, 46-116.
- . 1995. *The Act Itself*. Oxford: Clarendon Press.
- Cannon, Laura. 2003. "The Butterfly Effect and the Virtues of the American Dream." *Journal of Social Philosophy*, 34: 545-555.
- Draper, Kai. 2005. "Rights and the Doctrine of Doing and Allowing." *Philosophy and Public Affairs*, 33: 253-280.
- Donagan, Alan. 1977. *The Theory of Morality*. Chicago: University of Chicago Press.
- Finnis, John, Germain Grisez, and Joseph Boyle. 2001. "'Direct' and 'Indirect': A Reply to Critics of Our Action Theory." *The Thomist*, 65: 1-44.
- Fitzpatrick, William. 2006. "The Intend/Foresee Distinction and the Problem of 'Closeness.'" *Philosophical Studies*, 128: 585-617.
- Foot, Philippa. 1967. "The Problem of Abortion and the Doctrine of Double Effect." *Oxford Review*, 5: 5-15. Reprinted in P. Foot, *Virtues and Vices*, Oxford: Oxford University Press, 19-31.
- . 1984. "Killing and Letting Die." In J.L. Garfield and P. Hennessey (eds.), *Abortion: Moral and Legal Perspectives*, Amherst: University of Massachusetts Press, 177-185.
- Hall, Ned. 2004. "Two Concepts of Causation." In J. Collins, N. Hall, and L. Paul, *Causation and Counterfactuals*, Cambridge, Massachusetts: the MIT Press, 225-276.

- Hanser, Matthew. 1999. "Killing, Letting Die, and Preventing People from Being Saved." *Utilitas*, 11: 277-295.
- Hart, H.L.A. 1967. "Intention and Punishment." *Oxford Review*, 4. Reprinted in H.L.A. Hart, *Punishment and Responsibility*, Oxford: Clarendon Press.
- Hart, H.L.A., and Tony Honoré. 1985. *Causation in the Law*. 2nd ed. Oxford: Clarendon Press.
- Hilborn, Robert. 2001. *Chaos and Nonlinear Dynamics: An Introduction for Scientists and Engineers*, 2nd edition. Oxford: Oxford University Press.
- Howard-Snyder, Frances. 2011. "Doing vs. Allowing Harm." In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Winter 2011 Edition,  
 URL = <<http://plato.stanford.edu/archives/win2011/entries/doing-allowing/>>
- Jackson, Frank and Michael Smith. 2006. "Absolutist Moral Theories and Uncertainty." *The Journal of Philosophy*, 103: 267-283.
- Kagan, Shelly. 1989. *The Limits of Morality*. Oxford: Oxford University Press.
- Kamm, Frances. 2007. *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*. New York: Oxford University Press.
- Lorenz, Edward. 1972. "Predictability; Does the Flap of a Butterfly's Wings in Brazil Set Off a Tornado in Texas?" *American Association for the Advancement of Science*, 139<sup>th</sup> meeting.
- Mack, Eric. 1985. "Three Ways to Kill Innocent Bystanders: Some Conundrums Concerning the Morality of War." *Social Philosophy and Policy*, 3: 1-26.
- McMahan, Jeff. 1993. "Killing, Letting Die, and Withdrawing Aid." *Ethics*, 103: 250-279.
- . 2009. "Intention, Permissibility, Terrorism, and War." *Philosophical Perspectives*, 23: 345-372.
- Mikhail, John. 2007. "Universal Moral Grammar: Theory, Evidence, and the Future." *Trends in Cognitive Science*, 11: 143-152.
- . 2011. *Elements of Moral Cognition: Rawls' Linguistic Analogy and the Cognitive Science of Moral and Legal Judgment*. New York: Cambridge University Press.
- Moore, Michael. 2009. *Causation and Responsibility: An Essay in Law, Morals, and Metaphysics*. Oxford: Oxford University Press.
- Norcross, Alastair. 2003. "Killing and Letting Die." In R.G. Frey and C.H. Wellman (eds.), *A Companion to Applied Ethics*, Oxford and Boston: Basil Blackwell, 451-463.

- Pruss, Alexander. 2013. "The Accomplishment of Plans: a New Version of the Principle of Double Effect." *Philosophical Studies*, 165: 49-69.
- Quinn, Warren. 1989a. "Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing." *Philosophical Review*, 98: 287-312.
- . 1989b. "Actions, Intentions, and Consequences: The Doctrine of Double Effect." *Philosophy and Public Affairs*, 18: 334-351.
- Ross, W.D. 1930. *The Right and the Good*. Oxford: Clarendon Press.
- Rickless, Samuel. 2011. "The Moral Status of Enabling Harm." *Pacific Philosophical Quarterly*, 92: 66-86.
- Russell, Bruce. 1977. "On the Relative Strictness of Negative and Positive Duties." *American Philosophical Quarterly*, 14: 87-97.
- Scanlon, T.M. 2008. *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge: Harvard University Press.
- Scheffler, Samuel. 1995. "Individual Responsibility in a Global Age." *Social Philosophy and Policy*, 12: 219-236.
- Shaw, Joseph. 2006. "Intentions and Trolleys." *Philosophical Quarterly*, 56: 63-83.
- Thomson, Judith. 1976. "Killing, Letting Die, and the Trolley Problem." *The Monist*, 59: 204-217.
- . 1985. "The Trolley Problem." *Yale Law Journal*, 94: 1395-1415.
- . 2008. "Turning the Trolley." *Philosophy and Public Affairs*, 36: 359-374.
- Vihvelin, Kadri and Terrance Tomknow. 2005. "The Dif." *The Journal of Philosophy*, 102: 183-205.

## Notes

---

<sup>1</sup> The example is from Lorenz 1972. For a systematic introduction to chaos theory, see Hilborn 2001.

<sup>2</sup> Cannon 2003, 145.

<sup>3</sup> The basic idea of this argument was suggested to me by Allan Gibbard. I am also grateful to John Ku for many extremely helpful discussions of it. But any problems with the argument as I develop it here should be attributed entirely to me.

<sup>4</sup> That is, under conditions where we can reasonably assign probabilities to the various possible outcomes of our conduct, but we cannot know these outcomes with certainty. See Kagan 1989, 87-91 and Jackson and Smith 2006.

<sup>5</sup> Besides applying only to absolute constraints on harming, another limitation of Kagan, Jackson, and Smith's arguments is that they allow defenders of constraints to avoid the conclusion that constraints require paralysis by applying constraints only in cases where the probability of causing harm exceeds a certain threshold (see Kagan 1989, 90n5; Jackson and Smith 275-278; and Aboodi et al. 2008. Jackson and Smith argue that this response is problematic, to which Aboodi et al.