JUST DESERTS

Debating Free Will

DANIEL C. DENNETT
GREGG D. CARUSO

CONTENTS

"Just Deserts is a delight: a sharp and interesting discussion of punishment, morality, choice, and much else. It hits the sweet spot; it's wonderfully clear and accessible – perfect for a newcomer to the free will debates – but also deep and subtle, with plenty to engage experts in the field."

Paul Bloom, Brooks and Suzanne Ragen Professor of Psychology, Yale University, and author of *Against Empathy*

"What it means to make a choice, to deserve praise or blame, to do the right thing – these are all at stake in the debate over free will. Here you will find two different viewpoints, elaborated and defended by true masters. Given the sharpness of both interlocutors, neither has anywhere to hide; a wide spectrum of important points is laid out for careful consideration."

Sean Carroll, author of *The Big Picture: On the Origins of Life, Meaning, and the Universe Itself*

"This is a very lively, engaging, and thoughtful debate between two well-informed and insightful philosophers. It is written in a very accessible style, and students and even scholars in other disciplines or sub-fields of philosophy will learn from it and find themselves drawn in. It does not just re-hash traditional debates, but pushes the frontiers outward. Highly recommended."

John Martin Fischer, Distinguished Professor of Philosophy at UC Riverside

"A philosophical debate in the grand style. Caruso and Dennett play in the philosophical equivalent of a three set tennis championship where the prize is whether free will exists or not and what this means for reward, punishment, and the criminal law. Serve, volley, amazing gets, overheads, long rallies, a few trick shots, several match

points. Really smart play from two philosophers at the top of their games."

Owen Flanagan, James B. Duke Distinguished University Professor, Duke University

"*Just Deserts* made me think philosophy should never be done alone, but with a partner of equal strength and opposing views, that the best of it should be made available to the public, and that it should leave readers with an appreciation of the depth and difficulty of the questions but no easy answers. It is a stirring discussion of a difficult issue, that distils the best of what has been said for both sides. I can think of no discussion of free will and desert that gets to the heart of the issues so effectively. It reminds you just how important and difficult and vitally alive philosophical debate can be."

Jenann T. Ismael, Professor of Philosophy at Columbia University, and author of *How Physics Makes Us Free*

"This is a spirited and enlightening debate between an influential defender of compatibilism about freedom, responsibility, and determinism (Dennett) and an astute defender of a hard incompatibilist or free will skeptical position (Caruso). The book breaks new ground on many issues; and it has made clearer to me than anything else I have ever read on the subject how central is the issue of "just deserts" to age-old debates about free will, moral responsibility, and determinism."

Robert Kane, University Distinguished Professor Emeritus of Philosophy and Law, University of Texas at Austin

Daniel C. Dennett is Co-Director of the Center for Cognitive Science and the Austin B. Fletcher Professor of Philosophy at Tufts University. His books include *Content and Consciousness* (1969), *Brainstorm*s (1978), *Elbow Room* (1984), *The Intentional Stance* (1987), *Consciousness Explained* (1991), *Darwin's Dangerous Idea* (1995), *Kinds of Minds* (1996), *Freedom Evolves* (2003), *Breaking the Spell* (2006), and *From Bacteria to Bach and Back: The Evolution of Minds* (2017). He is a leading defender of compatibilism, the view that determinism can be reconciled with free will, and is perhaps best known in cognitive science for his concept of intentional systems and his multiple drafts model of human consciousness.

Gregg D. Caruso is Professor of Philosophy at SUNY, Corning, and Honorary Professor of Philosophy at Macquarie University. He is also the Co-Director of the Justice Without Retribution Network at the University of Aberdeen School of Law. His books include *Free Will and Consciousness* (2012), *Rejecting Retributivism: Free Will, Punishment, and Criminal Justice* (2021), *Exploring the Illusion of Free Will and Moral Responsibility* (ed. 2013), *Neuroexistentialism: Meaning, Morals, and Purpose in the Age of Neuroscience* (co-ed. with Owen Flanagan, 2018), and *Free Will Skepticism in Law and Society: Challenging Retributive Justice* (co-ed. with Elizabeth Shaw and Derk Pereboom, 2019). He is a leading proponent of free will skepticism, which maintains that who we are and what we do is ultimately the result of factors beyond our control, and because of this we are never morally responsible for our actions in the basic desert sense – i.e. the sense that would make us truly deserving of blame and praise, punishment and reward.

Just Deserts

Debating Free Will

Daniel C. Dennett and Gregg D. Caruso

polity

First published in 2021 by Polity Press

Polity Press
65 Bridge Street
Cambridge CB2 1UR, UK

Polity Press
101 Station Landing
Suite 300
Medford, MA 02155, USA

ISBN-13: 978-1-5095-4577-3

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

Library of Congress Cataloging-in-Publication Data

Names: Dennett, D. C. (Daniel Clement) author. | Caruso, Gregg D., author.
Title: Just deserts : debating free will / Daniel C. Dennett and Gregg D. Caruso.
Description: Medford : Polity Press, 2021. | Includes bibliographical references and index. | Summary: “An eye-opening debate on the philosophy and psychology of free will and what they tell us about our societies”-- Provided by publisher.
Identifiers: LCCN 2020029874 (print) | LCCN 2020029875 (ebook) | ISBN 9781509545759 (hardback) | ISBN 9781509545766 (paperback) | ISBN 9781509545773 (epub)
Subjects: LCSH: Free will and determinism. | Free will and determinism--Psychological aspects. | Civilization, Modern--21st century.
Classification: LCC BJ1461 .D4276 2021 (print) | LCC BJ1461 (ebook) | DDC 123/.5--dc23
LC record available at https://lccn.loc.gov/2020029874
LC ebook record available at https://lccn.loc.gov/2020029875

Foreword

Derk Pereboom

This exchange between Daniel Dennett and Gregg Caruso on free will, moral responsibility, and punishment is intense and engaging, and will captivate any reader who is interested in the cutting edge of the contemporary free will debate. It has much to offer newcomers and seasoned veterans alike. This exchange serves as an excellent introduction and at the same time provides details about the contested positions not available elsewhere.

Caruso is an incompatibilist about free will and determinism. If determinism is true, then there are factors beyond our control, events in the distant past and natural laws, that causally determine all of our actions, and incompatibilists maintain that this would rule out free will. Incompatibilists divide into those who hold that determinism is false and that we have free will – the libertarians – and those who hold that determinism is true and we lack free will – free will skeptics. Dennett affirms compatibilism about free will and determinism, and he contends that we do have free will. Caruso argues that we would lack free will if our world is deterministic, but also if it were indeterministic, say in the way some interpretations of quantum physics propose. Caruso and Dennett are thus situated on opposite sides of a traditional divide – Dennett is a compatibilist and affirms free will, Caruso is an incompatibilist and a free will skeptic.

There is also a conceptual issue about how “free will” should be defined, which divides Dennett and Caruso. Dennett is well known for recommending that we should use “free will” so that it refers to a kind of free will *worth*

wanting. The kind of free will worth wanting is a capacity for rational response to stimuli in our natural and social environment, which has developed in our species in its evolutionary history and matures in individuals as they become adults. This is clearly a valuable capacity, and I think that Dennett's proposal is defensible.

But a question that arises about Dennett's characterization is whether it works for dividing the sides in the debate. Because few believe that Dennett's notion of free will is incompatible with determinism, his definition results in a challenge for defining compatibilism so that it's controversial. Caruso, by contrast, defines free will as the control in action required for attributions of desert in its *basic* form, as do a number of other participants in the current debate. In the basic form of desert, someone who has acted wrongly deserves to be blamed and perhaps punished just because she has acted for morally bad reasons, and someone who has acted rightly deserves credit or praise and perhaps reward just because she has acted for morally good reasons. Such desert is basic because these desert claims are fundamental in their justification; they are not justified by virtue of further considerations, such as anticipated good consequences of implementing them. According to Caruso, in order to facilitate the free will debate, so that there are substantial numbers of participants on each side, free will should be defined as the control in action required for basically deserved praise and blame, reward and punishment.

Dennett and Caruso disagree about the prevalence of the notion of basic desert. By contrast with Dennett, Caruso believes it is widespread. He supports his view by using a thought experiment that derives from Immanuel Kant, in which there are no good consequences to be achieved by punishing a wrongdoer (1785: Part II). Here is a version of such an example. Imagine that someone on an isolated

island brutally murders everyone else on that island, and that he is not capable of moral reform, due to his inner hatred and rage. Add that it is not possible for him to escape the island, and no one else will ever visit because it's too remote. There is no longer a society on the island whose rules might be determined by a social contract aimed at good consequences, since the society has been disbanded. Do we have the intuition that this murderer still deserves to be punished? If so, then punishment would be basically deserved if the example in fact does eliminate the options for non-basic desert, as it seems to.

But on Dennett's side, do we want to define "free will" so that anyone who rejects basic desert counts as denying free will, or so that anyone who denies that we have the control in action required for attributions of basic desert counts as denying free will? Perhaps enough of the role that the concept "free will" has in our thought and practice would survive the rejection of basic desert and the control in action required for it. We have many concepts that we've retained even though we've revised how they are characterized, say due to scientific advance.

Dennett contends that enough of the role of the concept "free will" would indeed survive the rejection of basic desert because we have a notion of non-basic desert that can do the work we want. Practice-level justifications for blame and punishment invoke considerations of desert, while that desert is not basic because at a higher level the practice is justified by good anticipated consequences, such as deterrence of wrongdoing and moral formation of wrongdoers. On Dennett's account, our practice of holding agents morally responsible in this non-basic desert sense should be retained because doing so would have the best overall consequences relative to alternative practices.

One might object that penalties and rewards justified by anticipated consequences on Dennett's model do not really qualify as genuinely *deserved*, since on such a view they ultimately function as incentives. In reply, citing the type of analogy Dennett provides in the exchange, it seems legitimate to say that someone who commits a foul in a sport deserves the penalty for that foul. But such sports-desert isn't basic – it's instead founded in considerations about how the particular sport works best. Similarly, suppose penalties for criminal behavior are justified on deterrence grounds, by the anticipated good consequence of safety. Imagine that lawyers and judges consider only backward-looking reasons to convict and punish, while their practice is justified on forward-looking grounds that lawyers and judges never consider or invoke. Arguably, it then would make sense for the lawyers and judges to think of the penalties as deserved.

Accordingly, the exchange between Dennett and Caruso involves substantive issues, and some conceptual and verbal issues as well. The conceptual issues are important, and their resolution depends on whether the role of the relevant concepts can be retained. Neither Dennett nor Caruso contends that the role of the concept of "basic desert" in justifying actual practice is worth preserving. But Dennett argues that "desert" and its role should be retained, while Caruso disagrees. Throughout the exchange, separating the verbal and conceptual from the substantive issues is a challenge, as it is generally in philosophy. Caruso and Dennett take it on in the classic way, by regularly prodding each other to clarify terms.

My sense is that Caruso's and Dennett's positions are substantively quite close on the basics of the free will debate, but that they do differ on other matters, such as the value of manipulation arguments for incompatibilism, the discussion of which is especially intense. They also diverge

on recommendations for treatment of criminals, despite both agreeing that current American practice requires serious reform. But it is not clear whether they differ on this issue because Dennett endorses justifications in terms of desert, while Caruso rejects them, or for some other reason. The reader will enjoy sorting out these issues in this valuable and timely dialogue.

Preface

The genesis of this book can be traced back to May of 2018, at a rooftop bar in Beirut, Lebanon, where the two of us first met and spent an enjoyable evening eating, drinking, and debating our respective views on free will during a conference on moral psychology at the American University of Beirut. We stayed in touch after that conference and eventually decided to work out our differences in the form of a conversation or debate, which resulted in a published exchange in *Aeon Magazine* on October 4, 2018 under the title: “Just Deserts: Can we be held morally responsible for our actions? Yes, says Daniel Dennett. No, says Gregg Caruso. Reader, you decide.” After that exchange was published, Pascal Porcheron from Polity Books approached us about continuing our conversation and expanding it into a book. We both agreed to the project immediately, since we have mutual respect for each other and thought there would be great value in continuing our conversation. The result is this book. It begins with a brief introduction in which Caruso discusses the problem of free will and defines some terminology. This is designed to aid readers unfamiliar with the problem of free will and to provide a brief summary of the various positions on the issue. The introduction is followed by three separate exchanges. The first is an edited and expanded version of our initial *Aeon* exchange, while the second and third are new and appear here for the first time.

D. D. and G. C.

Introduction

Gregg D. Caruso

The problem of free will has real-world implications for our self-understanding, our interpersonal relationships, and our moral and legal practices. The assumption that we have free will lurks behind the justification of many of our everyday attitudes and judgments. For instance, when someone morally wrongs us, not only do we experience resentment and moral anger, we typically feel that we are *justified* in doing so, since we assume that, absent any excusing conditions, people are free and morally responsible for what they do and are therefore appropriate targets for such responses. We also typically assume that when individuals "act of their own free will," they *justly deserve* to be praised and blamed, punished and rewarded for their actions since they are morally responsible for what they do. Similar assumptions are made in the criminal law. The US Supreme Court, for instance, has asserted: "A 'universal and persistent' foundation stone in our system of law, and particularly in our approach to punishment, sentencing, and incarceration, is the 'belief in freedom of the human will and a consequent ability and duty of the normal individual to choose between good and evil'" (United States v. Grayson, 1978). But does free will really exist? What if it turns out that *no one* is ever free and morally responsible in the relevant sense? What would that mean for society, morality, meaning, and the law? Could society properly function without belief in free will? These are just some of the questions to be debated in this book.

To begin, it's important to introduce some key terms and positions. First, we can say that *free will*, as contemporary philosophers tend to understand it, is the control in action

required for a particular kind of moral responsibility. More specifically, it's the power or capacity characteristic of agents, in virtue of which they can *justly deserve* to be blamed and praised, punished and rewarded for their actions. Understanding free will as linked to moral responsibility in this way, anchors the philosophical debate in something comparatively concrete and undeniably important to our lives. As Manuel Vargas notes: "This is not a sense of free will whose only implication is whether it fits with a given philosopher's particular speculative metaphysics. It is not a sense of free will that is arbitrarily attached to a particular religious framework. Instead, it is a notion of free will that understands its significance in light of the role or function it plays in widespread and recognized forms of life" (2013: 180).

Contemporary theories of free will can be divided into two general categories: those that endorse and those that are skeptical of the claim that human beings have free will. The former category includes *libertarian* and *compatibilist* accounts of free will, two general views that defend the claim that we have free will but disagree on its nature or its conditions. The latter category consists of a class of *skeptical views* that either doubt or deny the existence of free will. The main dividing line between the two pro-free will positions, libertarianism and compatibilism, is best understood in terms of the traditional problem of free will and determinism. *Determinism*, as it is commonly understood, is the thesis that at any given time only one future is physically possible (van Inwagen 1983: 3). Or put differently, it's the thesis that facts about the remote past in conjunction with the laws of nature entail that there is only one unique future (McKenna and Pereboom 2016: 19). *Indeterminism*, on the other hand, is the denial of this thesis – it's the claim that at some time more than one future is physically possible. The traditional *problem of free*

will and determinism therefore comes in trying to reconcile our intuitive sense of free will with the idea that our choices and actions may be causally determined by factors over which we have no ultimate control, that is, the past before we were born and the laws of nature.

Historically, libertarians and compatibilists have reacted to this problem in different ways. *Libertarians* (not to be confused with the political view) acknowledge that if determinism is true, and all of our actions are causally determined by antecedent circumstances, we would lack free will and moral responsibility. Yet they further maintain that at least some of our choices and actions must be free in the sense that they are not causally determined. Libertarians therefore reject determinism and defend an indeterminist conception of free will in order to save what they maintain are necessary conditions for free will – the *ability to do otherwise* in exactly the same set of conditions and/or the idea that we remain, in some important sense, the *ultimate source/originator* of action. *Compatibilists*, on the other hand, set out to defend a conception of free will that can be reconciled with determinism. They hold that what is of utmost importance is not the absence of causal determination, but that our actions are voluntary, free from constraint and compulsion, and caused in the appropriate way. Different compatibilist accounts spell out requirements for free will differently but widely endorsed views single out responsiveness to reasons, self-control, or connection of action to what one would reflectively endorse.

In contrast to these pro-free will positions are those views that either doubt or outright deny the existence of free will and moral responsibility. Such views are often referred to as skeptical views, or simply *free will skepticism*. In the past, the leading form of skepticism was *hard determinism*: the view that determinism is true and incompatible with

free will – either because it precludes the ability to do otherwise (*leeway incompatibilism*) or because it is inconsistent with one's being the ultimate source of action (*source incompatibilism*) – hence, no free will. For hard determinists, libertarian free will is an impossibility because human actions are part of a fully deterministic world and compatibilism fails to reconcile determinism with free will. Hard determinism had its classic statement in the time when Newtonian physics reigned supreme and was thought to be deterministic. The development of quantum mechanics, however, diminished confidence in determinism, for the reason that it has indeterministic interpretations. This is not to say that determinism has been refuted or falsified by modern physics, because a number of leading interpretations of quantum mechanics are consistent with determinism. It is also important to keep in mind that even if we allow some indeterminacy to exist at the micro-level of the universe, say the level studied by quantum mechanics, there may still remain *determinism-where-it-matters* – i.e. at the ordinary level of choices and actions, and even the electrochemical activity in our brains. Nonetheless, most contemporary skeptics tend to defend positions that are best seen as distinct from, but as successors to, traditional hard determinism.

Many contemporary free will skeptics, for instance, maintain that while determinism is incompatible with free will and moral responsibility, so too is *indeterminism*, especially if it is limited to the sort posited by certain interpretations of quantum mechanics. Others argue that regardless of the causal structure of the universe, we lack free will and moral responsibility because free will is incompatible with the pervasiveness of *luck*. Others still, argue that free will and ultimate moral responsibility are incoherent concepts, since to be free in the sense required for ultimate moral responsibility, we would have to be

causa sui (or “cause of oneself”) and this is impossible. Here, for example, is Nietzsche on the *causa sui*:

> The *causa sui* is the best self-contradiction that has been conceived so far; it is a sort of rape and perversion of logic. But the extravagant pride of man has managed to entangle itself profoundly and frightfully with just this nonsense. The desire for “freedom of the will” in the superlative metaphysical sense, which still holds sway, unfortunately, in the minds of the half-educated; the desire to bear the entire and ultimate responsibility for one’s actions oneself, and to absolve God, the world, ancestors, chance, and society involves nothing less than to be precisely this *causa sui* and, with more than Baron Munchhausen’s audacity, to pull oneself up into existence by the hair, out of the swamps of nothingness. (1886/1992: 218–219)

The one thing, however, that all these skeptical arguments have in common, and what they share with classical hard determinism, is the belief that our choices, actions, and constitutive characters are ultimately the result of factors beyond our control – whether that be determinism, chance, or luck – and because of this we lack the kind of free will needed to hold agents morally responsible in the relevant sense.

List of Useful Definitions

Determinism: The thesis that facts about the remote past in conjunction with the laws of nature entail that there is only one unique future.

Compatibilism: The thesis that free will can be reconciled with the truth of determinism – i.e. it is possible for

determinism to be true and for agents to be free and morally responsible in the relevant sense.

Incompatibilism: The thesis that free will cannot be reconciled with determinism – i.e. if determinism is true, free will is not possible.

Libertarianism: The thesis that incompatibilism is true, that determinism is false, and that some form of indeterminist free will exists.

Free Will Skepticism: The thesis that no one has free will, or at the very least, that we lack sufficient reason for believing that anyone has free will.

Hard Determinism: The thesis that incompatibilism is true, that determinism is true, and therefore no person has free will.

Hard Incompatibilism: The thesis that free will is incompatible with both determinism and indeterminism – i.e. that free will is incompatible with *both* causal determination by factors beyond the agent's control *and* with the kind of indeterminacy in action required by the most plausible versions of libertarianism.

Hard Luck: The thesis that regardless of the causal structure of the universe, we lack free will and moral responsibility because free will is incompatible with the pervasiveness of *luck*.

Basic-Desert Moral Responsibility: For an agent to be morally responsible for an action in this sense is for it to be theirs in such a way that they would deserve to be blamed if they understood that it was morally wrong, and they would deserve to be praised if they understood that it was morally exemplary. The desert at issue here is *basic* in the sense that the agent would deserve to be blamed or praised just because they have performed the action, given an understanding of its moral status, and not, for example,

merely by virtue of consequentialist or contractualist considerations (Pereboom 2014: 2).

Consequentialism: The view that normative properties depend only on consequences – i.e. whatever produces the best aggregate set of good outcomes or makes the world best in the future.

Contractualism: The thesis that moral norms and/or political authority derive their normative force from the idea of a contract or mutual agreement.

Deontology: The view that the morality of an action should be based on whether the action itself is right or wrong under a clear set of rules, rather than based on the consequences of the action.

Exchange 1
Debating Free Will and Moral Responsibility

Caruso: Dan, you have famously argued that freedom evolves and that humans, alone among the animals, have evolved minds that give us free will and moral responsibility. I, on the other hand, have argued that what we do and the way we are is ultimately the result of factors beyond our control, and that because of this we are never morally responsible for our actions, in a particular but pervasive sense – the sense that would make us *truly deserving* of blame and praise, punishment and reward. While these two views appear to be at odds with each other, one of the things I would like to explore in this conversation is how far apart we actually are. I suspect that we may have more in common than some think – but I could be wrong. To begin, can you explain what you mean by "free will" and why you think humans alone have it?

Dennett: A key word in understanding our differences is "control." Gregg, you say "the way we are is ultimately the result of factors beyond our control" and that is true of only those unfortunates who have not been able to become autonomous agents during their childhood upbringing. There really are people, with mental disabilities, who are not able to control themselves, but normal people can manage under all but the most extreme circumstances, and this difference is both morally important and obvious, once you divorce the idea of *control* from the idea of *causation.* Your past does not control you; for it to control you, it

would have to be able to monitor feedback about your behavior and adjust its interventions – which is nonsense.

In fact, if your past is roughly normal, it contains the causal chains that turned you into an autonomous, self-controlling agent. Lucky you. You weren't responsible for becoming an autonomous agent, but since you are one, it is entirely appropriate for the rest of us to *hold you responsible* for your deeds under all but the most dire circumstances. As the American country singer Ricky Skaggs once put it: "I can't control the wind, but I can adjust the sails." To suppose that some further condition should be met in order for you or anyone else to be "truly deserving" of praise or blame for your actions is to ignore or deny the manifest difference in abilities for self-control that we can observe and measure readily. In other words, the rationale or justification for excusing someone, holding them not deserving of criticism or punishment, is their deficit in this competence. We don't try to reason with bears or babies or lunatics because they aren't able to respond appropriately. Why do we reason with people? Why do we try to convince them of conclusions about free will or science or causation or anything else? Because we think – for good reason – that in general people are reasonable, are moved by reasons, can adjust their behavior and goals in the light of reasons presented to them. There is something indirectly self-refuting in *arguing* that people are not moved by reasons! And that is the key to the kind of self-control which we are justified in treating as our threshold for true desert.

Caruso: I don't disagree with you that there are important differences between agents who have the kind of rational control you highlight and those who lack it. Such a distinction is undeniable. A normal adult who is responsive to reasons differs in significant ways from one who is suffering from psychopathy, Alzheimer's, or severe mental illness. I have no issue, then, with acknowledging various

degrees of “control” or “autonomy” – in fact, I think you and other compatibilists have done a great job highlighting these differences. My disagreement has more to do with the conditions required for what I call “basic-desert” moral responsibility. As a free will skeptic, I maintain that the kind of control and reasons-responsiveness you point to, though important, is not enough to ground basic-desert moral responsibility – the kind of responsibility that would make us truly deserving of blame and praise, punishment and reward in a purely backward-looking sense.

Consider, for example, the various justifications one could give for punishing wrongdoers. One justification, the one that dominates our legal system, is to say that they *deserve* it. This *retributive* justification for punishment maintains that punishment of a wrongdoer is justified for the reason that he/she *deserves* something bad to happen to them just because they have knowingly done wrong. Such a justification is purely backward-looking. For the retributivist, it is the basic desert attached to the criminal’s immoral action alone that provides the justification for punishment. This means that the retributivist position is not reducible to consequentialist considerations that try to maximize good outcomes in the future, nor in justifying punishment does it appeal to wider goods such as the safety of society or the moral improvement of those being punished. I contend that retributive punishment is never justified since agents lack the kind of free will and basic-desert moral responsibility needed to ground it.

While we may be sensitive to reasons, and this may give us the kind of voluntary control you mention, the particular reasons that move us, along with the psychological predispositions, likes and dislikes, and other constitutive factors that make us who we are, themselves are ultimately the result of factors beyond our control. And this remains true whether those factors include determinism,

indeterminism, chance, or luck. This is not to say that there are not other conceptions of responsibility that can be reconciled with determinism, chance or luck. Nor is it to deny that there may be good forward-looking reasons for maintaining certain systems of punishment and reward. For instance, free will skeptics typically point out that the impositions of sanctions serve purposes other than giving criminals what they basically deserve: it can also be justified by its role in incapacitating, rehabilitating and deterring offenders. My question, then, is whether the kind of desert you have in mind is enough to justify retributive punishment? If not, then it becomes harder to understand what, if anything, our disagreement truly amounts to, since forward-looking justifications of punishment are perfectly consistent with the denial of free will and basic-desert moral responsibility. And, if you are willing to reject retributivism, as I think you might be, then I'm curious to know exactly what you mean by "desert"– since it's debatable whether talk of giving agents their *just deserts* makes any sense devoid of its backward-looking, retributive connotations.

Dennett: You grant that the distinction I make between people who are autonomous and those who are not (because of various limits on their abilities to control themselves) is important, but then say that it is not enough for "the kind of desert" that would "justify retributive punishment." I too reject retributivism. It's a hopeless muddle, and so is any doctrine of free will that aspires to justify it. But that doesn't mean there is no "backward-looking" justification of punishment.

It's quite straightforward. On Monday you make me a promise, which I accept in good faith, and rely on when I adjust my own activities. On Friday, I discover you have broken your promise, with no excuse (what counts as an excuse has been well explored, so I will take that on

without further notice). I blame you for this. My blaming you is of course backward-looking: “But you *promised me*!” Autonomous people are justly held responsible for *what they did* because all of us depend on being able to count on them. It is for this reason that among their responsibilities is preserving their status as autonomous agents, guarding against the usurpation or manipulation of their own powers of discernment and decision. So, we can blame them for being duped, for getting drunk, etc. When we blame them, we are not just diagnosing them, or categorizing them; we are holding them *deserving* of negative consequences. If this isn’t “basic desert” then so much the worse for basic desert. What is it supposed to add to this kind of desert?

The fact is – and I invite you to consider whether it is a fact – that autonomous people *understand* that they will be held to account and have tacitly accepted this as a condition for their maintaining their freedom in the political sense. I take this to be all the grounds we need for justifying the imposition of negative consequences (under all the usual conditions). The difference between the madman who is physically restrained and removed to quarantine for the sake of public safety, and the deserving culprit who is similarly restrained and then punished, is large, and it is a key feature of any defensible system of government. The culprit has the kind of desert that warrants *punishment* (but not “retributive” punishment, whatever that is).

As I have argued before, we can see this rationale in a simpler domain of human activity: sport. The penalty kicks and red cards of soccer, the penalty box of ice hockey, the ejection of players for flagrant fouls, etc., all make sense; the games they enable would not survive without them. The punishment (consider the etymology of “*penalty*”) is relatively mild because “it’s only a game,” but if the transgression is serious enough, large fines can be assessed, or banishment from the game, and, of course,

criminal prosecution for assault or cheating also lurks in the wings. Free will skeptics should consider if they would abolish all these rules because the players don't have real free will. And if they would grant a special exemption for such penalties in sport, what principle would they cite for not extending the same policies to the much more important game of life?

You also say "the particular reasons that move us, along with the psychological predispositions, likes and dislikes, and other constitutive factors that make us who we are, themselves are ultimately the result of factors beyond our control." So what? The point I think you are missing is that autonomy is something one *grows into,* and this is indeed a process that is *initially* entirely beyond one's control, but as one matures, and learns, one begins to be able to control more and more of one's activities, choices, thoughts, attitudes, etc. Yes, a great deal of luck is involved, but then a great deal of luck is involved in just being born, in being alive. We human beings are well designed to take advantage of the luck we encounter, and to overcome or deflect or undo the bad luck we encounter, to the point where we are held responsible for not taking foolish chances (for instance) that might lead to our losing control. There is no incompatibility between determinism and self-control.

Caruso: Well, I'm glad to know that you reject retributivism along with "any doctrine of free will that aspires to justify it." This point of agreement is significant since it entails that major elements of the criminal justice system are unjustified. I'm curious to know, however, with what exactly you would replace retributive legal punishment, and to what extent you reject the status quo. I ask because, though you claim to reject retributivism, you go on to defend a backward-looking conception of blame and punishment grounded in the idea that offenders are