Ludger Jansen
Paul M. Näger   *Editors*

Peter van Inwagen

Materialism, Free Will and God

Springer

Münster Lectures in Philosophy

Volume 4

Since 1997 the Department of Philosophy at the University of Münster has hosted the Münster Lectures in Philosophy. This lecture series gives especially young researchers in philosophy and adjacent disciplines the opportunity to enter into an intellectual exchange with internationally and nationally renowned philosophers. Each volume of the series contains an evening lecture by the guest, critical contributions regarding the guest's work provided by the participating young researchers, and commentaries of the guest relating to these contributions.

More information about this series at http://www.springer.com/series/13531

Ludger Jansen • Paul M. Näger
Editors

Peter van Inwagen

Materialism, Free Will and God

Springer

Editors

Ludger Jansen
Faculty for Catholic Theology
Ruhr University
Bochum, Germany

Paul M. Näger
Department of Philosophy
WWU Münster
Münster, Germany

Preface

Peter van Inwagen is one of the most influential contemporary philosophers. Tackling perennial philosophical problems with a view on modern scientific results, his philosophy masterfully combines positions that have been considered irreconcilable: incompatibilism concerning free will, materialism, organicism, theism and realism concerning fictional entities. His arguments are witty, surprising and deep.

This book emerged from the 19th Münster Lecture in Philosophy in 2015. It includes van Inwagen's lecture on free will and ten papers from the associated colloquium, discussing central themes of his philosophy, plus replies to these contributions by van Inwagen. As an introduction to his philosophy that relates his work to other contemporary views, the book is of interest to graduate students and professionals in philosophy alike.

In his lecture *The Problem of Free Will Revisited* (Chap. 1), van Inwagen argues for the claim that the usual terminology in the debate is confusing and should be replaced. Especially the central term "free will" has been understood in many different ways – compare the libertarian's meaning of the word with that of the compatibilist – producing semantical *scheindebates*, or pseudo-debates, on what is its correct meaning. As a cure, van Inwagen proposes a new straightforward terminology, whose central term is "being able to", allowing him to reformulate the problem in a non-ambiguous way. His argument reveals that beyond the terminological debates about the correct meaning of "free will" lies a substantial paradox, which in the new terminology he calls the "culpability problem".

The contributions start with methodological issues. The first paper (Chap. 2) discusses Peter van Inwagen's epistemological claim that there is no convincing reaction to situations when persons disagree in their judgments, although being presented with the same evidence ("revealed peer disagreement"). The authors propose a solution by specifying how "rationality", "truth", "evidence" and "justification" should be viewed to interrelate.

Elaborating on Quine's views, Peter van Inwagen has developed a distinctive position on meta-ontological issues, which is the topic of Chap. 3. The authors provide a comprehensive summary of his meta-ontology – the first outline of this kind – and advert to two open questions that will need further attention.

Chapters 4, 5, 6, and 7 deal with several aspects of van Inwagen's ontological works. One of van Inwagen's most famous claims is that living organisms are the only composite objects that exist – a thesis that is sometimes called "organicism". In Chap. 4, the authors suggest to expand this ontology by also accepting composite objects in the case that physical bonding occurs. According to this proposal, besides living organisms, rocks, molecules and galaxies exist. It is another consequence of Inwagen's organicism that there are no material artefacts like chairs and houses. Chapter 5 critically examines his central arguments for this view and, positively, provides reasons for the case that artefacts exist.

Concerning fictional characters like Sherlock Holmes, van Inwagen holds the view that these characters do exist and that they are theoretical entities of literary criticism. Chapter 6 argues that the latter claim is doubtful because, among others, the analogy between fictional entities and other theoretical entities in scientific disciplines, e.g., electrons, is rather weak. A fictional Socratic dialogue between a proponent of van Inwagen's view of fictional characters and a critic both explores open questions and proposes possible answers (Chap. 7).

The papers in Chaps. 8, 9, and 10 treat the subject of van Inwagen's Münster lecture, the problem of free will and responsibility, which is another major theme in his work. He is famous for his claim that free will is incompatible with both determinism and indeterminism but nevertheless holds in a libertarian sense. Chapter 8 re-examines the prospects of reconciling free will with indeterminism by discerning a mental and a physical perspective on the problem. The second paper on the topic (Chap. 9) explores a neglected route in the problem, namely, the route of "fatalism" or "hard determinism". The author argues that the view is both consistent and plausible if one overcomes certain epistemological misunderstandings. Chapter 10 criticizes two of van Inwagen's claims on moral responsibility: The authors argue that in cases of radical memory loss, a person should not be held responsible for the actions of her predecessor; furthermore, there are cases in which internal or external factors provide mitigating circumstances for morally wrong actions (which van Inwagen denies).

The third major strand in van Inwagen's oeuvre is the philosophy of religion. As a theistic philosopher, he has presented a particular view on the central problem of evil, in his case by providing a defence story about how evil could have emerged. The authors of Chap. 11 cast doubt on this defence for presupposing implausible meanings of the concepts "God", "love" and "freedom".

The final chapter (Chap. 12) contains van Inwagen's reply to each of the contributions.

We thank all contributors for making the event and this book happen. First of all, this Münster lecture would not have been possible without the generous support of Springer. We are also grateful to Logan Pecinovsky for improving the style and language of the discussion papers. According to the idea of the Münster lectures to promote an exchange between young researchers and an internationally renowned philosopher, most of them were graduate students at the time of the colloquium,

writing their first research paper under the supervision of an experienced staff member. The book witnesses that such exchanges can be productive, stimulating and exciting. This is essentially also due to the fact that Peter van Inwagen did not cease to explain and defend his ideas, which is recorded in his detailed replies. His smart and entertaining way of discussing philosophical problems made it a pleasure to have him as our guest in Münster.

Bochum/Münster, Germany   Ludger Jansen
July 2017   Paul M. Näger

Contents

About the Editors

Ludger Jansen teaches philosophy at the Ruhr University Bochum and the University of Rostock. He holds degrees in both philosophy and theology and a doctorate in philosophy from the WWU Münster and habilitated in philosophy from the University of Rostock. His research covers wide areas of metaphysics and ontology, including biomedical and social ontology.

Paul M. Näger is currently an assistant professor for philosophy at the WWU Münster. He studied philosophy and physics in Munich, Oxford and Bremen and specialises in the philosophy of science and metaphysics. He is author of the book *Quantum Entanglement and Causation* (forthcoming, Springer) and a co-author of the textbook *The Philosophy of Quantum Physics* (German ed. 2015, English ed. forthcoming, Springer).

Chapter 1
The Problem of Free Will Revisited

Peter van Inwagen

In 1886 and 1887, the American philosopher George Santayana, then a very young man, took a leave of absence from his doctoral studies at Harvard University and spent those two years in Berlin, attending the lectures of various famous professors. He was particularly impressed by the lectures of the great psychologist Hermann Ebbinghaus. Ebbinghaus, for his part, took a fatherly interest in the young American. In his autobiography, Santayana relates the following anecdote.

> [Ebbinghaus] even asked me to his house, showed me his first fat baby[1] and talked about William James, of course eulogistically, but with fundamental reservations, as for instance, on the question of free will and responsibility, on which he said "*Das hat er eigentlich nicht durchgedacht.*" (Santayana 1963, 258)

If Ebbinghaus was alluding to "The Dilemma of Determinism", the only thing James wrote on free will that I have read, I must agree with his judgment. I reproduce this anecdote, however, because I think that Ebbinghaus's words are no less applicable to the authors of a very high proportion of the essays and books devoted to the free-will problem that have been published during the last 25 or 30 years than they were to James's lecture.

One widespread consequence of not properly thinking through the question of free will and responsibility that is to be found in most of these current and recent essays and books is their almost universal adherence to the thesis that Harry Frankfurt's arguments have shown that moral responsibility does not require free will (in the sense of being able to act otherwise than one does). Another is the fact that in that literature one frequently encounters phrases like "libertarian free will" and "the compatibilist conception of free will" – phrases that make no sense whatever. But I have discussed those two confusions (so I believe them to be) extensively on various other occasions, and in the remainder of this lecture I will mention the first only briefly and in passing and say nothing at all about the second.

---

[1] The "fat baby" can only have been Julius Ebbinghaus (1885–1981), later to become a well-known philosopher.

P. van Inwagen (✉)
Department of Philosophy, University of Notre Dame, Notre Dame, IN, USA
e-mail: Peter.VanInwagen.1@nd.edu

My topic is rather the effect that confused thinking has had on the language in which the free-will problem is framed and discussed. In my view, this language, this family of interwoven technical terms, has, as a consequence of this confused thinking, been corrupted – has in fact become hopelessly corrupt. I intend, as I proceed, to make it clear what I mean by this statement.

To begin. Too many philosophers of the present day treat the phrase "free will" as if it were like such philosophically important words as "knowledge" and "cause" and "existence". These are words that have important applications in everyday life and which may therefore reasonably be supposed to express concepts that may be explicitly formulated and clarified and analyzed. If this is so in the case of, for example, "cause", it makes sense for two philosophers who have proposed incompatible accounts of causation to dispute about which of them (if either) has got the *right* account of causation. The right account of causation – whatever other features it may have – will be an account that in some sense corresponds to the concept of causation that is enshrined in natural language. But nothing like this is possible in the case of "free will" because that phrase is a pure term of philosophical art. That is, in the work of a given philosopher, it means whatever that philosopher says it means – it is thus less like "knowledge" and "cause" and "existence" than it is like "realism" or "quantifier" or "quale". Let us take "realism" as our example of a pure philosophical term of art. I ask you to imagine a heated debate between two philosophers, Philomena and Philocles, that proceeds as follows.

> *Philomena*: Philocles, your work on realism rests on a fundamental misunderstanding of the concept of realism. You say that realism is the doctrine that propositions have mind-and-culture-independent objective truth-values. But that is not what realism is. Realism is nothing like that. Realism is the doctrine that universals exist, and its opposite is *nominalism*.
> *Philocles*: No, Philomena, I'm right and *you're* fundamentally confused. Realism is exactly what I said it is, and its opposite is *anti-realism*.

At this point, a third philosopher, Philotas, breaks into the discussion and says:

> Excuse me, but I couldn't help overhearing what you two were saying, and I'm sorry to have to tell you that you're *both* wrong – both fundamentally confused about what realism is. Realism is the doctrine that, in addition to minds and sensations, there is a material world that exists independently of anyone's perceptual awareness of it. The opposite of realism is *idealism*.

Obviously, this debate makes no sense whatever – for the simple reason that "realism" is a pure philosophical term of art and, in any philosophical book or essay, it means whatever the author of that book or essay says it means. It is my contention that debates about the meaning of "free will" or about "what free will is" make no more sense than this imaginary debate about what realism is.

Now you may want to tell me that what I'm saying is obviously wrong, since we do use the words "free will" in everyday life, and we understand them as well as we understand any words we use. I reply that there is a sense in which it is true that we

use the words "free will" in everyday life (and use them in a perfectly intelligible way), but that this fact does not contradict what I have been saying. In everyday life – in the business of getting and spending, falling in love and raising families, voting in elections, consulting doctors and lawyers, and so on – when we use the words "free will" it is almost always inside the phrase "of one's own free will" (where "one's" represents the position of a possessive pronoun). And there's no real dispute about what that phrase means: if, for example, you're asked in a court of law whether you did something-or-other of your own free will, you're being asked whether you acted under duress – under any sort of coercion. (Were you for example threatened with certain untoward consequences if you did not act that way?) And everyone knows that people sometimes do things when they believe that their not doing them would have no untoward consequences. It is obvious that none of the disputes about "free will" that are so notable a part of the intellectual landscape, past and present, are about whether people ever do things without being coerced. Suppose someone asks:

How can we possibly have free will if God foresees everything we do?

Or, again, suppose someone asks:

How can we possibly have free will if a being with infinite power of calculation and a knowledge of the laws of mechanics and of the forces that particles exert on one another and a knowledge of the position and momentum of every particle of matter at any given time could calculate the position and momentum of every particle of matter at any other time?

Or, finally, suppose someone asks:

How can we possibly have free will if the Libet experiments show that physical conditions sufficient for our so-called voluntary bodily movements exist prior to our conscious decision to make those movements?

It is obvious that none of these three speakers is asking how it could be that our acts are uncoerced in the circumstances they have specified – for those circumstances don't even *seem* to be incompatible with the absence of coercion.

The ordinary meaning of the phrase "free will" – or "*freier Wille*" or "*libre arbitre*" or what have you – simply does not explain its use in contexts like these. (Incidentally, everything I say about "free will" will apply to any closely related words or phrases – such as "freedom", "free action", "free choice", "acted freely", …).

Now all this would be of no philosophical consequence if philosophers all meant the same thing by "free will". And it would be of no philosophical consequence if several meanings had been given to "free will" but it was well known what these alternative meanings were and it was always or at least normally clear in any given context which of them a philosopher was using. (Something like this is the case with philosophers' use of "possibility": there is epistemic possibility, there is metaphysical possibility, there is physical possibility – and so on.) And there was a time in the history of analytical philosophy when it came very close to

being the case that all the philosophers of that school meant the same thing by "free will". From the mid-sixties till the mid-eighties of the last century, almost all philosophers writing in English meant the following by "free will": agents have free will just to the extent that they perform free actions, and an action is free just in the case that the agent who performs it "could have done otherwise" – that is, had the ability to do something else instead, the ability to perform some other, alternative action. I speak of analytical philosophers writing in English, but it is instructive to look at the collection *Seminar: Freies Handeln und Determinismus* (Pothast 1978) for the use of German university students. It is true that most of the essays in this volume were German translations of essays originally composed in English, but the translators did not hesitate to translate "free will" as "*freier Wille*" and "free action" as "*freies Handeln*".

Unfortunately, it is no longer the case that "free will" and "free action" have single, uniform meanings. For example, some philosophers now say that for an act to be free, it must be the case that the agent who performed that act had open to him alternative possibilities for action in whatever sense of "having alternative possibilities open to him" it is that is required for the agent to be morally responsible for that action. (And this is meant not simply to be a true – and philosophically substantive – statement about free action. It is meant to be a *definition* of "free action".) Other philosophers say that for an agent to have free will is for that agent to be the ultimate source of his acts. Contrast these definitions with the old definition, so to call it – that is, that an action is free just in the case that the agent who performs it was able to do something else instead. Note how much clearer the old definition is: it essentially depends on the ordinary, everyday word "able" – a word understood by anyone who speaks English, just as "*können*" is understood by anyone who speaks German and "*pouvoir*" is understood by anyone who speaks French and "*posse*" is understood by anyone who has learned a little Latin. To understand the other definitions, however, one must understand phrases like "alternative possibilities" and "ultimate source", phrases whose meaning is far from clear. But my main objection to this – as I see it – corruption of the terminology of the free-will problem is not merely that it constitutes a regression from relative clarity to relative obscurity. My main objection is that it has entirely obscured the problem. This corruption of terminology has had the consequence that it is no longer clear what the problem of free will *is*. But I do not think that at this point you will be able to understand why I say this. I will give my reasons for affirming this thesis partly by examining some of the language used by philosophers who claim to be discussing "free will" and partly in the form of a commentary on a certain philosophical problem.

I will begin with the commentary on the problem – a problem that I regard as one of the great, central problems of philosophy. I will first state the problem and then proceed to the promised commentary. *Some* writers have called this problem "the problem of free will". It is, in fact, closely modeled on a problem that I myself once proposed as a referent for the name "the problem of free will" (van Inwagen 2008). You will notice, however, that the adjective "free" will not occur in my statement of the problem – nor will the noun "freedom" or the adverb "freely'" My statement of the problem will be, so to speak, "free"-free. If I sometimes *mention* the f-word,

generally in the course of quoting the writings or discussing the work of others, I shall at any rate not *use* it again. (And I will not use the noun "will" again, either.) You will also notice that there will be no mention of the moral responsibility of agents for their actions in my statement of the problem. I shall indeed speak of certain states of affairs *being the fault of* various agents, or, alternatively, of those agents' *being to blame for* those states of affairs, and fault and blame certainly bear some intimate relation to moral responsibility, whatever the words "moral responsibility" may mean. I take it that a statement like

> Manfred is morally responsible for his wife's unhappiness

can be nothing more than a philosopher's needlessly technical way of saying "Manfred is to blame for his wife's unhappiness" or "It's Manfred's fault that his wife is unhappy". (I'm not much interested in what a philosopher might mean by saying that someone – Manfred, say – was "morally responsible" for some *good* state of affairs. I suppose it would be something like "Manfred gets the credit for it".) So I propose to speak of fault and blame instead of moral responsibility. And note that if one has decided to speak of fault and blame (as opposed to moral responsibility), and if that decision leads one to reflect on how judgments of fault and blame are framed, one will realize that the objects of those judgments are states of affairs and not actions. That is to say, the things that can be said to be someone's fault (or for which someone can be blamed) are states of affairs that are in some way or other, causally related to that person's actions or inactions, not the actions or inactions themselves. The statement "It's Henry's fault she was elected; he cast the deciding vote" makes perfect sense. But what could a statement like "It's Henry's fault that he raised his hand when the chair said, 'All in favor?'" mean? In the unlikely event that someone did say that, I'd have to cast about for an interpretation; I suppose I'd decide that the speaker must have regarded some consequence of the hand-raising as a bad thing and had chosen a rather puzzling way of saying that that bad thing was Henry's fault.

Finally, I will not use the treacherous – because radically ambiguous – phrase "could have done otherwise". (I will, in fact, not use "could have" at all.)

So much for the words and phrases that will not occur in the statement of the philosophical problem I have promised you. Now the problem.

1.1 Four Theses

I begin by presenting four theses. I am not affirming these theses (and that is certainly to my credit – for two of them are mutual contradictories); I am rather putting them before you so that I can refer to them in the statement of the problem. Two of the theses – the two contradictories – have familiar names: "determinism" and "indeterminism". I'll state the other two first, however. These theses have no standard or customary names, or at least none that do not involve the words and phrases I have promised not to use, so I shall have to invent my own names for them. In

order to ensure that these names are not tendentious, I will call them simply "Thesis One" and "Thesis Two". Here is Thesis One:

> On at least some occasions when a human agent is trying to decide between two or more incompatible courses of action, that agent is able to perform each of them.[2]

If, for example, the following story is true, there has been one such "occasion". Early last January, Sally was admitted both to the Julliard School of Music to study piano and to the Harvard Law School. At that time she wanted very much to become a concert pianist (for reasons of personal fulfillment). She also wanted very much to become a lawyer (in this case, her reasons were moral and social). She spent the month of January trying to decide whether to study piano at Julliard or law at Harvard (or, more immediately, whether to accept the Julliard offer and decline the Harvard offer or to accept the Harvard offer and decline the Julliard offer). At every moment during the course of these deliberations, she was able to do this:

> accept the Julliard offer and decline the Harvard offer,

and she was able to do this:

> accept the Harvard offer and decline the Julliard offer.

That is to say, at every moment in the course of her deliberations she had *both* those abilities. (Of course, having both the ability to do A and the ability to do B is not the same thing as having the ability to do both A and B.)

---

[2] I concede that "able to" has many senses. In the text, I alluded to the "radical ambiguity" of "could have done otherwise", and it is my firm opinion that "was able to …" is less *dangerously* ambiguous – less likely to slip from one of its senses to another in the course of a philosophical argument –, than "could have …". Nevertheless, the phrase "is able" (whatever its tense; whatever the infinitive it governs) *is* ambiguous. For example: Grisha Sokolov has been stranded on a desert island; is he able to play the piano? In one sense, yes, in another, no. Or: the loan officer at your bank knows that she would lose her job if she approved your application for a loan; is she able to approve it? Of course: she has only to sign this piece of paper – and yet she says to you, "I'm afraid I'm unable to approve the loan you've applied for". Is she mistaken? Lying? I have discussed ambiguities of these and various other kinds that attend the phrase "is able to" in Section 1.4 of *An Essay on Free Will* (van Inwagen 1983, 8–13). But the sense of "is able to" that figures in the argument of this paper may be specified by a simple device – by considering what is involved in being in a position to make a promise. Suppose that Alice asks Tim to give her a ride to work the following day (it's a serious matter: she'll lose her job if she counts on Tim for a ride and he fails to provide it). A necessary (and I think sufficient) condition for Tim's being in a position to promise to give Alice the requested ride is that he believes that he *is able to* give her a ride. And those italicized words have, in that context, the sense I mean "is able to" to have in the argument in the text. Suppose Winifred and Sokolov are both castaways on the same pianoless island; able though he is to play the piano (in one sense of "able"), he is not in a position to promise Winifred that he will play the piano that evening. And the loan officer is no doubt in a position to promise you to approve the loan ("no doubt": it might be that she is unsure whether it is psychologically possible for her to sign the piece of paper in those circumstances) – although of course it would be either foolish or dishonest of her to *make* such a promise.

And here is Thesis Two (note that Thesis Two refers to Thesis One):

> If the bad consequences of a decision are ever the fault of the person who made the decision, then Thesis One is true.

An alternative formulation of Thesis Two:

> If anyone can rightly be blamed for the bad consequences of some decision he or she has made, then Thesis One is true.

Consider, for example, the following story of a decision that has had bad consequences. One of Professor Lustig's students offered to have sexual intercourse with him if he would give her the highest marks on her moral philosophy examination. Lustig thought it over and decided to accept her offer – a decision that led to his losing his position, his family's being in serious financial need, and his wife's being driven nearly mad with rage and jealousy.

Now most people, on hearing this story, would say that all these bad things – Lustig's loss of his position, his family's severe need, his wife's near madness – were *Lustig's fault*. They would say that he was to *blame* for them. Let us suppose that most people would be right: these things were Lustig's fault and he was to blame for them.

Thesis Two implies that it follows from these bad consequences of Lustig's decision being his fault that Thesis One is true. It follows, that is, that it is false that it is *never* the case that when a human agent is trying to decide between two or more incompatible courses of action, that agent is able to perform each of them.

We next state the other two theses, determinism and indeterminism, that will figure in the statement of the problem.

- *Determinism* is the thesis that the past and the laws of nature determine a unique future.
- *Indeterminism* is the thesis that the past and the laws of nature do not determine a unique future.

We now proceed to a description of the dialectical situation that, as it were, generates the problem.

1.2   The Dialectical Situation

There are seemingly unanswerable arguments that (if they are indeed unanswerable) demonstrate that Thesis One is incompatible with determinism. I allude, of course, to the various versions of the Consequence Argument, as it is commonly called. And there are seemingly unanswerable arguments that (if they are indeed unanswerable) demonstrate that Thesis One is incompatible with *in*determinism (this part needs a little work, since indeterminism does not imply that a *given person's* actions are undetermined; the work can be done). I allude, of course, to the various versions of the *Mind* argument – named in honor of the august philosophical

journal in which so many variants on it have appeared. But if Thesis One is incompatible with both determinism and indeterminism, then Thesis One is false – necessarily false, in fact.

There are, moreover, seemingly unanswerable arguments that, if they are correct, demonstrate the truth of Thesis Two. But I can expect protests at this point. The protester speaks:

> What about Frankfurt's refutation of the Principle of Alternative Possibilities?[3] Has Frankfurt not shown that – or at any rate, can his arguments not easily be adapted to show that – a certain state of affairs can be someone's fault even if no one trying to decide whether to do A or to do B has ever been able to do A *and* able to do B?

I might say in reply that I explained many years ago why Frankfurt's arguments do not show any such thing[4], and that if people have not been listening, it's *not my fault*. But that reply, like the protest it is a reply to, is really beside the point. I am stating a problem, not discussing possible solutions to that problem. And I am not stating this problem with an eye towards presenting a solution of my own to it. Frankfurt's arguments are best looked at as a proposed solution to the problem I am in the process of stating. Similar remarks apply to supposed refutations of the Consequence Argument and the *Mind* Argument. (This may be a difficult exercise, but do your best not to attend to the question whether the statements I make in the course of laying out the problem I mean to discuss are true or false; what you should be attending to is the fact that, in making these statements, I never use certain words and phrases, to wit: "free", "freely", "freedom", "could have done otherwise", and "moral responsibility".) All right – that point having been made, let us continue with the statement of the problem.

We had got as far as saying that there are seemingly unanswerable arguments both for the falsity of Thesis One and the truth of Thesis Two. But if Thesis One is false and Thesis Two is true, then nothing is ever anyone's fault. And it is evident that it is simply *false* that nothing is ever anyone's fault. It must, therefore, be that at least one of the following four propositions is true:

- The seemingly unanswerable arguments for the incompatibility of Thesis One and determinism are in fact answerable; these arguments are fallacious.
- The seemingly unanswerable arguments for the incompatibility of Thesis One and indeterminism are in fact answerable; these arguments are fallacious.
- The seemingly unanswerable arguments for Thesis Two are in fact answerable; these arguments are fallacious.
- It is *not* evident that it is simply false that nothing is ever anyone's fault; and not only is it not evident, it's not even *true*: The apparent self-evidence of that thesis is illusory.

---

[3] The supposed refutation was first presented in his celebrated and widely reprinted essay "The Principle of Alternate Possibilities" (Frankfurt 1969). The essay has generated a vast body of discussion and commentary.

[4] See, e.g., van Inwagen 1978, 1999.

1.3   The Statement of the Problem

My statement of the problem is in the form of three interrelated questions:

> Which of these four propositions are true? For each of the first three: if it is
> true, what are the fallacies in the arguments to which it alludes? If the fourth
> proposition is true, what is the nature of the illusion that has made it seem
> self-evident to me and many other philosophers (and, indeed to the great mass
> of humanity) that many of the bad things that have happened in the course of
> human history are someone's fault?

For reasons that I hope I have made clear, I decline to call this problem "the problem
of free will". It will, however, be convenient to have a name for it. I will call it "the
Culpability Problem" – with the understanding that this name is a mere tag whose
purpose is to facilitate reference – a proper name, if you will. The fact that *culpa* is the
Latin word for "fault" or "blame" should be regarded as a mere *aide-mémoire*. (I
introduce a name for the problem only at this late point in my argument with the spe-
cific intention of underscoring the fact that I ascribe no significance to the name I have
chosen. And if anyone does find some reason to dislike this name, if anyone regards it
as tendentious or in any other way objectionable, I'll simply call it something else –
"the Three Questions Problem" or "Peter's Problem" or "Arthur"[5].) Note that the only
philosophical technical term that occurs in the statement of the Culpability Problem is
"determinism" – "indeterminism" being merely the contradictory of "determinism".

I have given up on the Culpability Problem. It's too hard for me. But my purpose
here today is not to solve it or even to examine proposed solutions to it. It's rather to
show that much philosophical work whose announced subject is "the problem of
free will" is simply irrelevant to the Culpability Problem – not addressed to that
problem at all. (I don't deny that this work may be valuable for other reasons. After
all, it's no objection to the discipline of social psychology that its investigations and
theories are of no relevance to the problems of astrophysics.)

I do not mean to imply that *all* work that is specifically addressed to "the problem
of free will" is irrelevant to the Culpability Problem. My own work is a case in
point. The subject-matter of, e.g., "The Incompatibility of Free Will and
Determinism" (van Inwagen 1975) and *An Essay on Free Will* (van Inwagen 1983)
is the Culpability Problem, although not under that name. The work on "free will"
that is irrelevant to the Culpability Problem is – at any rate most of it is – work that
involves what I shall call *verbal essentialism*.

A piece of philosophical writing exhibits what I call verbal essentialism if there
is some philosophical term of art (either a word or phrase invented by philoso-
phers – like "actualism" and "nowness" – or a dictionary-entry word that is used in
by philosophers in a special technical sense, like "proposition" and "validity") such
that the thesis presented in that text could not be stated *without using that word or
phrase*. I have, for example, recently accused Karen Bennett of falling prey to verbal

---

[5] Reporter: "What do you call that haircut?"; George Harrison: "Arthur."

essentialism – you will have guessed that I regard verbal essentialism as a Bad Thing, something one can properly be said to fall prey to – in her paper "Proxy 'Actualism'" (Bennett 2006). In that paper, Bennett presented certain criticisms of Alvin Plantinga's so-called actualism (in his philosophy of modality). In the introductory paragraph of my critique of Bennett's criticisms of Plantinga, I wrote,

> My conclusion will be that [Bennett's] criticisms fail, owing to the fact that they depend on the historical accident that the customary designation for Plantinga's position is "actualism" – that if this position had been given a name that did not contain "actual" or any word formed from "actual" the criticisms of the position that are presented in "Proxy 'Actualism'" could not even be stated. (van Inwagen 2012, 69)

The works I mean to call attention are like that, mutatis mutandis: you could not rewrite them in such a way as to eliminate the phrase "free will" and the words "freedom" and "freely" from them – there would simply be nothing left; a translation of these works into "free"-free language is impossible. (That would not be the case with, say, *An Essay on Free Will* – which is essentially an essay on the Culpability Problem; nothing in its substantive content hangs on my choice of "free will" as a term of art.)

It is time to turn to examples. I begin with Daniel Dennett's latest thoughts on free will, contained in his recent book *Intuition Pumps and Other Tools for Thinking* (Dennett 2013, chap. 8). But lest you think that my accusation of verbal essentialism is directed only at those philosophers who, like Dennett, take what might be called a "deflationary" position on "free will", let me assure you that I will also consider examples of philosophers (conveniently cited by Dennett) who speak favorably of things like "ultimate moral responsibility" and "absolute free will" and agents who are "perfectly free to do otherwise". Those philosophers and Dennett, in my view, occupy two sides of the same coin – the same verbally essentialist coin.

One of Dennett's targets in that chapter is those scientists who contend that science has shown that free will is an illusion:

> [These scientists] have typically been making a rookie mistake: confusing the manifest image with what we might call the folk ideology of the manifest image. [Dennett regards free will, like color and solidity, as a denizen of Wilfrid Sellars's "manifest image".] The folk ideology of color is, let's face it, bonkers; color just isn't what most people think it is, but that doesn't mean that the manifest world doesn't really have any colors; it means that colors – real colors – are quite different from what most folks think they are […] Similarly, free will isn't what some of the folk ideology of the manifest image proclaims it to be, a sort of magical isolation from causation.[6] […] I wholeheartedly agree with the scientific chorus that that sort of free will is an illusion, but that doesn't mean that free will is an illusion in any morally important sense. It's as real as colors […] (Dennett 2013, 355)

> [These scientists] are making the mistake people make when they say that nothing is ever [colored][7], not really. They are using an unreconstructed popular concept of free will, when they should be adjusting it first, the way they do with color […] (Dennett 2013, 356)

---

[6] A very misleading phrase. I shall explain why in a moment.

[7] Dennett actually has "solid" and not "colored" here. I have substituted "colored" for "solid" in order that my scattered quotation should express a unified thought. Dennett had earlier used solidity and color as parallel examples of things that are *real* and yet not much like what we thought they were before science revealed their true nature.

Now I am not convinced that science has shown us that color just isn't what ordinary folk who know nothing of the physics of electromagnetic radiation or the psychology and physiology of color perception think it is. (Let's call them "the folk".) This is not because I think that what science has revealed about color shows us that color *is* what the folk think it is. It's rather that I'm not sure whether the folk think color is *anything*, bonkers or not – whether they have *any* particular views on what color is. I'm not sure whether there *is* a 'folk ideology' of color. If there is such an ideology – I would ask –, what is its propositional content? Dennett mentions the indisputable facts that electromagnetic radiation in the 390–700 nm range is not made of little colored things and that atoms have no colors, but I can't tell from the context of that remark whether he meant it to imply that the denials of these two propositions were contained in the folk ideology of color. If he did, he would certainly be wrong, since this folk ideology of color, if it ever existed, must have antedated our knowledge of atoms in the modern chemical sense of the word and of electromagnetic radiation.

But let's suppose that Dennett is right: There *is* a folk ideology of color and scientific investigation shows that it's entirely wrong – *but* we should not infer from the wrongness of the folk ideology that color is an illusion; color is real, but it's not what everyone used to think it was and the folk still think it is.

But what then does it mean to say that color is real and why does Dennett think it is real in that sense? Dennett does not define "real" and he does not exactly argue for the reality (in any sense) of color; what he says in defense of the reality of color is more along the lines of an appeal to common sense:

> [Color] is not an illusion in the sense that matters: nobody thinks Sony is lying when it says that its color televisions really show the world of color or that Sherwin-Williams should be sued for fraud for selling us many different colors in the form of paint. (Dennett 2013, 69)

I would myself offer something a little more theoretical in support of the reality of color – although what I would offer is certainly consistent with the point Dennett means those two examples to illustrate. I would say something along the following lines. Most of the statements we make in ordinary life that contain color-words ("The car that left the scene of the accident was a dark green Lexus"; "Titanium dioxide is the most common white pigment"; "The predominant color of Picasso's *La Vie* (1903, Cleveland Museum of Art) is blue") are *true*; And if the sentences "Color is an illusion" or "Color is not real" mean anything at all (they certainly don't wear their meanings on their sleeves), they must mean something that implies that most of those statements are false. But that's by the way.

Dennett's thesis, then, is that free will is like color (as he represents color). It is real enough, but it cannot be what the folk ideology of free will says it is: There is no place for *that* in the scientific image (that's the image of the world that Sellars – and, following him, Dennett – opposes to the manifest image). The "unreconstructed popular concept of free will" is inconsistent with what science has discovered about the nature of the beings (us) to whom it is supposed to apply.

I said a moment ago (in effect) that I was not sure whether there was any such thing as the *unreconstructed popular* concept of color – although I was willing to

grant for the sake of argument that there was such a thing. However that may be, there is certainly such a thing as the concept of color. That is to say, there are such things as the meaning of the word "colored" and the meanings of the words "green" and "brown" and "mauve". (Let nominalists understand that statement as they will: There has to be some sense in which it is true.) Here is why the case of the concept of color (even assuming that Dennett is right when he says that there is an unreconstructed popular concept of color and that nothing in the real world corresponds to it) is not parallel to the case of the concept of free will: There is no such thing as the concept of free will. And, of course, if there is no such thing as the concept of free will, there is no such thing as the unreconstructed popular concept of free will.

But I suppose I am getting ahead of myself. Before I say anything more about the non-existence of the unreconstructed popular concept of free will, I should tell you what Dennett thinks the unreconstructed popular concept of free will is, and I should tell you what the scientific-image-friendly concept with which he means to replace it is. As to the former, he says (towards the end of his chapter on free will):

> People care deeply about having free will, but they also seem to have misguided ideas about what free will is or could be (like their misguided ideas about color […]). Our decisions are not little miracles in the brain that violate the physics and chemistry that account for the rest of our bodies' processes, even if many folk think this must be what happens if our decisions are to be truly free. We can't conclude from this, however, that then we don't have free will, because free will in this bonkers sense is not the only concept of free will. (Dennett 2013, 406)

Now this passage contains a statement that is so egregiously confused that I must depart for a moment from the line of argument I have been pursuing and correct it. Whatever "many folk" may think, the philosophers who call themselves "incompatibilists" and who also believe that we human beings are sometimes able to act otherwise than we do do *not* think that "our decisions are little miracles in the brain that violate the physics and chemistry that account for the rest of our bodies' processes". According to incompatibilism, if an agent decides to do A rather than B, then – if the agent was *able* to choose to do B – *neither* a decision to do A *nor* a decision to do B would have violated the laws of physics and chemistry. In Dennett's defense, however, it should be noted that he elsewhere mentions an unnamed philosopher who "has frankly announced that a free choice is a 'little miracle'". As President-for-Life of the World Society of Incompatibilists, it is my unpleasant duty to inform that philosopher that if he or she is a member of the Society, he or she is hereby excommunicated for having made this heretical statement. (The heresy in question is sometimes called "contra-causal freedom".)

Having got this out of the way, let us return to the questions we were considering – "What does Dennett think the folk conception of free will is?" and "What is the scientifically more adequate conception with which he proposes to replace it?" He says this about the latter conception, early on in the chapter:

> The intuition pumps in this [chapter] are designed to wean you from [the folk ideology of] free will and get you to see a better concept, the concept of real free will, practical free will, the phenomenon in the manifest image that matters. (Dennett 2013, 356)

Unfortunately, Dennett never gets round to spelling out the precise content of either concept (maybe he cannot be blamed for that in the case of the folk concept, supposing it to exist; maybe it *has* no precise content). He does tell us a quite a bit about both concepts, however. For example, he quotes statements by various philosophers that are intended to illustrate the folk concept – such as this passage from Jerry Fodor's review of his, Dennett's, book *Freedom Evolves*:

> One wants to be what tradition has it that Eve was when she bit the apple. Perfectly free to do otherwise. So perfectly free, in fact, that even God couldn't tell which way she'd jump. (Fodor 2003, 18)

And this passage, from Galen Strawson's review of the same book,

> [Dennett] doesn't establish the kind of absolute free will and moral responsibility that most people want to believe in. (Strawson [2003] 2008, 333)

Dennett supposes that this "being perfectly free to do otherwise" this "absolute free will" are not philosophers' inventions but are components of the folk ideology of free will; he supposes that Fodor and Strawson have correctly identified "what one wants to be" and "what most people want to believe in". He seems to come down rather hard on Fodor's words "even God couldn't tell which way she'd jump" – that is, he seems to suppose that absolute unpredictability, unpredictability even in principle, of human behavior is the essential core of the folk-ideological concept of absolute free will. And I think that he sees Strawson's well-known contention that to enjoy absolute free will one would have to be the sole and ultimate cause of one's actions as also being essential to the folk ideology. He wonders why anyone would want to be an in-principle-unpredictable agent and the ultimate cause of one's actions, and why anyone would be attracted to the belief that we were agents of this sort. And well he might. I would too – at least if I were sure I understood what Fodor and Strawson were talking about. In my view, however, Fodor and Strawson are simply reproducing some ideas invented by philosophers and not reporting what the man on the Clapham omnibus (as he was once called in Oxford) wants to be or wants to believe in the existence of. Or, better, not *ideas* invented by philosophers but words and phrases invented by philosophers – in J. L. Austin's immortal phrase, "a certain special, happy style of blinkering philosophical English".

In any case, he wants to replace the folk-ideological concept of free will with something else: a better candidate for the office "free will", something that is consistent with our present scientific knowledge, something that isn't bonkers, something that it would actually make sense to want to have oneself and to want to believe that one's fellow human agents had.

My only problem with this project is that there's no such office. Whatever the replacement he may have in mind may be, there's nothing for it to replace.

Otiose or not, it isn't entirely clear what the proposed replacement is, although it certainly has these features:

- It involves a certain amount of unpredictability in practice, but not unpredictability in principle. (Like it or not, life occasionally requires us to compete with other inhabitants of the world, and organisms that are deficient

in unpredictability tend not to pass their genes along to their descendants –
think of a gazelle that *always* swerves left when it's being pursued by a lion.
But since we're unlikely ever to be in competition with the Laplacian
Reckoner, unpredictability in principle would enjoy no advantage over unpre-
dictability in practice – unpredictability by the organisms with which we are
actually in competition.)
- It involves certain kinds of "freedom from" – from coercion, from physical
  bondage, from illusion and hallucination …
- It is compatible with determinism.

Well, I'm happy to give Dennett the words "free will". Let him spell out the details
of the concept he intends this phrase to denote as he wants.

Let's suppose we have Dennett's concept before us. Whatever precisely it may
be, I am, as I have said, happy to let him call it 'free will' – or for that matter, the
one 'the one possible non-bonkers concept of free will' if that's what he wants to
call it. And then let him and Fodor and Strawson – those fellows on the other side of
the verbally essentialist coin – fight over the words 'free will'. Let them fight over
who's got free will *right* or over what free will *really is* (or really would be if it
existed). Let Fodor and Strawson accuse Dennett of purveying "*ein elender Behelf*"
(in Kant's words) or of leading his readers into "a quagmire of evasion" (in James's).
Let Dennett reply that a belief in the phenomenon Fodor and Strawson want the
words 'free will' to designate would be as bonkers as a belief in levitation (a com-
parison he uses at one point).

I think such a debate would be about an entirely meaningless issue. It is certainly
about a meaningless issue if, as I suppose, there is no concept that goes by the name
"free will". But suppose I am wrong about that. It is at least *not evident* that the
debate would be a meaningful one even in that case. Suppose that most people *do*
believe or want to believe that what they do is in principle unpredictable and *do*
believe or want to believe that they are the sole causes of what they do; why would
what Dennett offers as a substitute for what they want or believe in or want to
believe in *be* a substitute – even a disappointing substitute – for *those* things? For
suppose that what Fodor and Strawson say people want is indeed non-existent –
still, not just any existent thing counts as substitute for just any non-existent thing,
if I may so express myself. ("I want to find the fountain of youth." "Oh, you're like
Ponce De León. You want to find the unreconstructed, folk-ideological fountain of
youth. That's bonkers. There's no such thing. But, fortunately, the fountain of youth
exists; it just isn't what we thought it was. It isn't a fountain that, when one drinks
from it, one is magically restored to youth and then never ages. It's a regimen of diet
and exercise that can extend the vigor of one's youth by as much as fifteen years.")
The only thing that ties Dennett's substitute to the Fodor-Strawson original –
whether that original is an actual folk ideology or a mere tangle of essentially mean-
ingless words and phrases invented by philosophers and known only to philosophers
and their students – is that he calls it "free will". Now I'll concede that he has as
much right to call some concept of his devising "free will" as Fodor has to say that
a person whose behavior is not predictable even in principle is "perfectly free to do

otherwise" or as Strawson has to call the ability to create one's character *ex nihilo* "absolute free will". Since none of these terms – "free will", "perfectly free to do otherwise", "absolute free will" – means anything in particular, they are available to be put to any use a writer wants to put them to.)

But perhaps I am wrong. Perhaps a debate about "what free will really is" or "what free will should be" or "the proper meaning of 'free will'" would make sense and would even be a valuable debate. I'm sure of one thing, however: whatever the value of such debates might be, none of them would have any relevance to the great, central philosophical problem that I have called the Culpability Problem.

References

Bennett, Karen. 2006. Proxy "Actualism". *Philosophical Studies* 129: 263–294.

Dennett, Daniel. 2013. *Intuition pumps and other tools for thinking*. New York: W. W. Norton & Company.

Fodor, Jerry. 2003. Why would mother nature bother? Review of Dennett's "Freedom evolves". *London Review of Books* 25 (5): 17–18.

Frankfurt, Harry. 1969. The principle of alternate possibilities. *The Journal of Philosophy* 66: 829–839.

Pothast, Ulrich, ed. 1978. *Seminar: Freies Handeln und Determinismus*. Frankfurt am Main: Suhrkamp.

Santayana, George. 1963. *Persons and places: The background of my life*. New York: Charles Scribner's Sons.

Strawson, Galen. 2008. Freedom evolves. In *Real materialism and other essays*, ed. Daniel C. Dennett, 333–335. Oxford: Clarendon Press. Originally published in *New York Times Book Review* 108, 2003.

van Inwagen, Peter. 1975. The incompatibility of free will and determinism. *Philosophical Studies* 27: 185–199.

———. 1978. Ability and responsibility. *The Philosophical Review* 87: 201–224.

———. 1983. *An essay on free will*. Oxford: Clarendon Press.

———. 1999. Moral responsibility, determinism, and the ability to do otherwise. *The Journal of Ethics* 3: 341–350.

———. 2008. How to think about the problem of free will. *The Journal of Ethics* 12: 327–341.

———. 2012. "Who sees not that all the dispute is about a word?": Some thoughts on Bennett's "Proxy 'Actualism'". *Hungarian Philosophical Review* 3: 69–81.

Chapter 2
Rational Peer Disagreement upon Sufficient Evidence: Leaving the Track to Truth?

Frieder Bögner, Thomas Meyer, Konstantin Schnieder and Markus Seidel

2.1 Introduction

Anyone who has or will read this volume inevitably, yet implicitly, runs across the problem this chapter will explicitly deal with. As in many other spheres of life, scientists tend to disagree about a wide range of subjects, questions, theories etc. One could easily read this whole book as an exercise and a manifestation of revealed disagreement, and one could easily point to further instances of disagreement encountered throughout everyday life. Due to epistemologists' generally increasing interest in the social dimensions of human knowledge, the fact of the ubiquitous presence of disagreement has also found enhanced philosophical attention by asking what the rational response to situations of revealed peer disagreement (RPD) should be. These are situations where at least two peers regarding a specific topic disagree about one and the same proposition while both are aware of the others diverging belief. In the eyes of some epistemologists such situations create the problem that there seems to be a prima facie reason for both to question one's own belief because a recognized peer holds the opposite belief. Situations of RPD seem to be higher-order defeating reasons of one's beliefs. In this philosophical debate Peter van Inwagen was one of the first researchers to explicitly confront this delicate matter in religion, politics, philosophy and science. Both in the private and public

F. Bögner (✉)
Department of Philosophy, Osnabrück University, Osnabrück, Germany
e-mail: frieder.boegner@uni-osnabrueck.de

T. Meyer
Department of Philosophy, WWU Münster, Münster, Germany
e-mail: thomas.meyer@uni-muenster.de

K. Schnieder
Center for Bioethics (CfB), WWU Münster, Münster, Germany
e-mail: konstantin.schnieder@uni-muenster.de

M. Seidel
Center for Philosophy of Science, WWU Münster, Münster, Germany
e-mail: markus.seidel@wwu.de

sphere, different people tend to hold widely diverging beliefs about, e.g., whether God exists, whether certain technical innovations are a good or bad thing, or whether Ronald Reagan was a good president. And similar disagreement can also be found among scientists from various disciplines, ranging from arts and humanities to the (natural) sciences, who may dispute how to interpret certain fossils, whether to accept the Copenhagen interpretation of quantum mechanics, and many other questions.

Yet why and in what sense is disagreement an interesting topic for epistemologists and not simply a rather brute fact of human psychology? Answering this question will be one of the central goals of the following section of this article. There we will set out a number of conditions that capture schematically the problem of RPD as it is suggested both by van Inwagen's descriptions and examples and as it presents a serious puzzle for epistemology (Sect. 2.2). Building on this fine-grained picture we will subsequently discuss four attempts to solve the puzzle of RPD and reconstruct why van Inwagen seems to find neither of these options a promising candidate (Sect. 2.3). We reflect on where this path of negative argumentation leaves us and – following a suggestion by van Inwagen himself – ask whether there might have been any hidden assumptions along the path that maneuvered us into this unsatisfying situation (Sect. 2.4). After briefly discussing a hidden metaphilosophical presupposition of van Inwagen, we then focus on a further concealed premise in his argumentation by showing how it poses an internal dilemma for his position – thus explaining van Inwagen's difficulties in solving the problem of RPD. Complementing this internal critique and closing our argumentation, we then ask whether van Inwagen might be willing to reject this assumption and, consequently, allow for rational peer disagreement upon sufficient evidence. Since this proposal comes with a cost – though it does not necessarily detract us from the track to truth – we will close by highlighting some advantages and disadvantages of our moderate proposal (Sect. 2.5).

2.2 Revealed Peer Disagreement: The Puzzle According to van Inwagen

2.2.1 The Situation in General

As we have already mentioned, disagreement seems to be a fact found throughout everyday life as well as in multiple areas of scientific research. In his discussion of the topic, Peter van Inwagen echoes this pervasiveness of disagreement across various domains in that he takes his paradigm cases of disagreement from diverse fields of both scientific and public debate (cf. van Inwagen 1996, 147–149). While disagreement is indisputably a feature of other scientific disciplines as well, one sort of disagreement has come to occupy a central stage in van Inwagen's articles on this topic and we will here follow this lead. More precisely, revealed peer disagreement