

SAUL SMILANSKY

COMPATIBILISM: THE ARGUMENT FROM SHALLOWNESS

ABSTRACT. The compatibility question lies at the center of the free will problem. Compatibilists think that determinism is compatible with moral responsibility and the concomitant notions, while incompatibilists think that it is not. The topic of this paper is a particular form of charge against compatibilism: that it is *shallow*. This is not the typical sort of argument against compatibilism: most of the debate has attempted to discredit compatibilism completely. The Argument From Shallowness maintains that the compatibilists do have a case. However, this case is only partial, and shallow. This limited aim proves itself more powerful against compatibilists than previous all-or-nothing attempts. It connects to the valid instincts of compatibilists, making room for them, and hence is harder for compatibilists to ignore.

To the extent that the institution of blame works coherently, it does so because it attempts less than morality would like it to do . . . [it] takes the agent together with his character, and does not raise questions about his freedom to have chosen some other character.

Bernard Williams, *Ethics and the Limits of Philosophy*

The compatibility question lies at the center of the free will problem. Compatibilists think that determinism (or indeed the absence of libertarian free will irrespective of determinism) is compatible with moral responsibility and the concomitant notions, while incompatibilists think that determinism (or absence of libertarian free will) is not compatible with moral responsibility. The debate and the various options will be reviewed shortly. The topic of this paper, once those options are understood, is a particular form of charge against compatibilism: that it is *shallow*.

This is not the typical sort of argument against compatibilism: most of the debate has attempted to discredit compatibilism completely, to provide arguments to the effect that there is nothing right about compatibilism, and to suggest that it is an unmitigated failure. The libertarian Peter van Inwagen, for instance, has argued that, given that in a deterministic universe everything follows from



Philosophical Studies **115**: 257–282, 2003.

© 2003 Kluwer Academic Publishers. Printed in the Netherlands.

earlier states of the universe according to the laws of nature, and since one can control neither the initial states nor the laws, then, if determinism is the case one has no control and is not morally responsible (van Inwagen, 1983). Galen Strawson is, by contrast, a hard determinist, but he as well argued in a similarly absolute vein: since moral responsibility would require the possibility of self-creation, and that notion is ultimately incoherent, we do not have moral responsibility (Strawson, 1994). Compatibilists have by and large responded symmetrically, denying any merit to the all-or-nothing views of their opponents.

The charge that compatibilism is shallow is a strong one, and if correct the incompatibilists will have been to a significant degree validated. Compatibilists will certainly want to rebut this charge. Nevertheless, the Argument From Shallowness, as I shall call it, by contrast to standard arguments such as van Inwagen's and Galen Strawson's, is based upon a measure of respect for compatibilism. As I shall interpret the situation, this argument maintains that the compatibilists do have a case. However, this case is only partial, and shallow. This limited aim proves itself more powerful against compatibilists than previous all-or-nothing attempts. It connects to the valid instincts of compatibilists, making room for them, and hence is harder for compatibilists to ignore. Reflecting on the shallowness that emerges from reflection that is initially favourable towards compatibilism might even move some who are sympathetic to compatibilism away from it, and towards a more complex view.

Part 1 reviews the free will problem and some alternatives. Part 2 presents a view I call Fundamental Dualism, which, given the absence of libertarian free will, combines compatibilism with hard determinism. Part 3 shows, in turn, the moral, personal-existential, and pragmatic shallowness of compatibilism. Part 4 considers possible compatibilist counter-arguments.

1. THE FREE WILL PROBLEM AND SOME ALTERNATIVES

In compact form, the free will problem can be presented as the conjunction of two questions:

- a. The first question is whether there is libertarian free will, and it can be called the libertarian Coherence/Existence Question.

Libertarians think that there is libertarian free will; everyone else disagrees. This question is metaphysical, or ontological, or possibly logical.

- b. The second question is whether – if there is no libertarian free will – our situation is still satisfactory. It can be called the Compatibility Question, namely, are moral responsibility and related notions such as desert and justice compatible with determinism (or with the absence of libertarian free will irrespective of determinism)? Compatibilism and hard determinism are the opponents on the Compatibility Question. This question, as we will understand it here, is both ethical and personal-existential.

The traditional positions can be seen clearly from their answers to these questions:

First, *libertarian free will*. We are all, more or less, familiar with the idea of libertarian free will. For our purposes it can be characterised roughly as the ability to control one's action and actually do otherwise in exactly the same situation, with internal and external conditions held constant. People naturally assume that they have libertarian free will, and it has formed the basis of most of the ethical teaching of the Western religions and of major ethical systems such as Kant's. To help us intuitively hook on to it, it is the sort of freedom that determinism would preclude (although indeterminism would also not help and in fact the issue of determinism is not important). Libertarianism of course answers yes to the first question. Typically libertarians are incompatibilists, that is they think that if we did not have libertarian free will we would be in trouble, and there would not be, for example, moral responsibility. But luckily, we do have libertarian free will. In other words, the libertarian is *demanding* but *optimistic*.

Compatibilism is, roughly, the position that the forms of free will most people clearly have to some degree, such as the ability to deliberate and do as they wish, suffice to meet the requirements of morality and personal life insofar as they are affected by the issue of free will. In particular, the compatibilist rejects the idea that some sort of 'metaphysical' or 'libertarian' notion of free will, such as would be negated by a completely deterministic ontology, is necessary in order to have moral responsibility. Hence, the term 'compat-

ibilism': the compatibilist insists that free will, moral responsibility, and their concomitant notions are compatible with determinism (or with the absence of libertarian free will). For example, the compatibilist would claim that most people in the West choose a career with some measure of freedom, and are morally responsible for this choice, although it follows from their desires and beliefs. Lack of relevant freedom would result only from atypical causes eliminating or severely curtailing control (such as pathological compulsion or external coercion). It is important to stress that compatibilism is not utilitarian or consequentialist but maintains contact with the traditional paradigm requiring control for moral responsibility, and moral responsibility for blameworthiness and desert. Compatibilists maintain that the traditional paradigm can be sustained even in a deterministic world, and does not require libertarian free will. On the *compatibilist level* of deliberating, choosing, and acting, most people are basically free, such matters are within their control, and it is this that matters. The compatibilist is *non-demanding* and hence *optimistic*.

Hard determinism, despite its misleading name, is not only a position on determinism or on the existence of libertarian free will (although it of course that as well). It is the opponent of compatibilism, which is sometimes called 'soft-determinism', on the second question. In other words, hard determinism is a normative position according to which, given that there is no libertarian free will in the world, moral responsibility and desert are impossible. Libertarian free will is required, but does not exist. The hard determinist is the pessimist in our cast of characters: she agrees with the libertarian that compatibilist free will is insufficient and that we require libertarian free will; hence, both are incompatibilists. But, like the compatibilist, the hard determinist believes that libertarian free will does not exist. Consider again the example of the person who chose his career freely according to the compatibilist: the hard determinist would claim that on the *ultimate level* the career-choice was not up to the person, who could not in the end form the sources of his motivation. These sources, the hard determinist will emphasise, are the basis for his 'free' choice on the compatibilist level. In certain cases – such as the man choosing a criminal career – this absence of ultimate control is what matters, and elimi-

nates moral responsibility. The hard determinist is *demanding* and *pessimistic*.

2. FUNDAMENTAL DUALISM ON FREE WILL – AND THE LIMITED VALIDITY OF COMPATIBILISM

A complex recent alternative that I have proposed (Smilansky, 2000, Part I) provides the basis for understanding free will-related matters: Fundamental Dualism. This position incorporates two elements, corresponding to the two basic questions that, as we saw above, make up the free will problem:

- a. The rejection of libertarian free will.
- b. The attempt to *combine* the insights of both compatibilism and hard determinism.

In its answer to the first question, the Fundamental Dualism is at one with compatibilism and hard determinism. We shall assume the absence of libertarian free will in this discussion, for our concern lies with the answer to the second, Compatibility Question, which asks about the implications of this absence.¹

The second, and for us more significant element of the Fundamental Dualism is its dualism – the combination of compatibilism and hard determinism. This dual perspective can be explored concerning various matters: morality, desert, justice, self-respect, and so on. We shall see it emerging in the various spheres where the shallowness of compatibilism will be demonstrated. Initially, I shall explain the dualism by considering moral responsibility, blameworthiness and justice.

The free will problem is about control: it issues from the core normative intuition that we must take human agency, control and its absence very seriously, particularly when judging ourselves and others. To punish a person for an act that he did not commit is a paradigm of injustice: control is a condition for moral responsibility, moral responsibility is a condition for blameworthiness, and blameworthiness is a condition for just punishment. The absence of control precludes guilt and blameworthiness, and hence punishment would be manifestly unjust. Control is also crucial for our self-

evaluative and reactive attitudes, even beyond morality, as we shall see later on.

It turns out, however, that the pertinent forms of control are fundamentally *dualistic*: on the one hand, we need to take seriously distinctions in local compatibilist control, if we are to treat people as we ought. Questions about the existence of control, as well as about degrees of control, make sense and are morally and personally central. On the compatibilist level we take the person as a 'given', and ask about his or her control in pedestrian ways: did he willingly do X? Was he coerced? Was he under some uncontrollable psychological compulsion? Most people most of the time do have compatibilist control over their actions, even if there is no libertarian free will (say, if determinism applies to all human actions). The kleptomaniac or alcoholic are not in control of their pertinent actions in the way that, respectively, the common thief or occasional mild drinker are in control, irrespective of determinism. And this often matters. If we are to respect persons, we need to establish and maintain a social order and human relationships that broadly follow the compatibilist distinctions in terms of local control. More will be said on this below.

But we can ask the question about control also on the ultimate level. Given that there is no libertarian free will, asking about 'ultimate control' lands us with the hard determinist conclusion, where ultimately there can be no control. Any person whom we could agree was on the compatibilist level free (that is, could reflect on his options, decide to do what he wanted, was not coerced, etc.) would be seen in a new light: under the ultimate perspective, the sources of his character and motivation would also be queried. And if we have no libertarian free will, then ultimately we are just 'given', with our desires and beliefs, and any change in them is ultimately down to our earlier selves, which we ultimately cannot control. We are what we are, and from the ultimate perspective, with all our compatibilist choosing and doing, we operate as we were molded.

Compatibilist Justice is the sort of justice which exists when we follow compatibilist distinctions in terms of local control: in other words, when the pertinent results of the social order suffered and enjoyed by persons, reflect their free choices (on the compatibilist level). For instance, those who choose to keep to the law will be

safe, while those breaking it may be punished. Ultimate Injustice is the sort of injustice that, I claim, may follow when we do not take account of the absence of ultimate control, and follow Compatibilist Justice. Such injustice occurs, for example, when we punish the compatibilistically-guilty. We may well morally need to do so, overall, and as we noted doing so along compatibilist lines is just in a way that, say, punishment based on factors beyond people's control such as race would not be – because (compatibilist level) control and its absence is not being respected. But we must not hide from ourselves the (ultimate level) injustice that following Compatibilist Justice would involve.

In my view we have to take account of both valid perspectives on control, the compatibilist and the ultimate hard determinist, for they are each part of the complex truth on the free will problem. Hence a Fundamental Dualism encompassing both perspectives. In sum, many of the practices of a community based on such compatibilist distinctions, a Community of Responsibility, would be *in one way* unjust, owing to the absence of libertarian free will which implies that our actions are on the ultimate level not up to us, and that to hold us responsible for them is therefore morally arbitrary. Nevertheless, working according to such distinctions might be just *in another way*, because they correspond to a sense of being up to us, existing to some degree in specific cases – and which in cases such as kleptomania would not be applicable, and therefore it would be an injustice to treat in the same way. To *fail* to create a Community of Responsibility is also in one sense to fail to create a feasible non-arbitrary moral order, hence to fail in the proper respect for persons. There is a basis for working with compatibilist notions of fault and moral responsibility, based on the local compatibilist level control, even though we lack the sort of deep grounding in the 'ultimately guilty self' that libertarian free will was thought to provide. Moreover, we are required by the core intuitions pertaining to free will to work in this way. But doing so has often a 'hard determinist' moral price. We must recognise both the frequent need to be compatibilists and the need to confront that price. Both compatibilist control and ultimate-level arbitrariness and lack of control are morally relevant. The intuitive power of the requirement for control manifests itself in dual ways. *We see why the Community of Responsibility is deeply*

(*non-consequentially*) morally imperative, and why its results are deeply morally disturbing.

Now, the immediate reaction of both compatibilists and hard determinists to such a dualistic account is likely to involve an attempt to discredit the other side's portion. 'Ultimate' injustice does not matter, the compatibilist might say, after all: you yourself tend to admit that we can distinguish between the guilty and innocent, and meet common intuitions about the way to treat various cases. Why care about 'ultimate fantasies' when people can have control of their lives, reform and even partially create themselves, and behave responsibly? The hard determinist is likely to attack me from the other side, saying that all talk about moral distinctions and about desert is groundless – do I not myself admit that any person, even the manifestly guilty in the compatibilist's eyes, is not ultimately responsible for being whoever he happens to be, and for the actions which result from this? What sort of control is it that is merely an unfolding of pre-set factors?

Both sets of arguments have some strength, which is why I think that any 'monistic' position is inadequate. However, once we make the conscious attempt to rid our minds of the assumption that either compatibilism or hard determinism must be exhaustively true (as there is no libertarian free will), we begin to see that there are aspects of the compatibilist case that the hard determinist cannot plausibly deny, and likewise with the hard determinist case. Since people tend to be immediately inclined in one way or the other, each reader will have to work on himself or herself *in order to see* the side he or she are blind to. One has to try to conquer one's blind side, and try to grasp how it is to inhabit it.

However deeply we might feel that all people are ultimately innocent, it is unconvincing to deny the difference between the control possessed by the common thief and that of the kleptomaniac, and to ignore the moral inadequacy of social institutions that would fail to take account of this difference. We have an intimate experience of control (or its lack). If a man believes that he is Napoleon then he is deluded, and his belief is *false*. But a woman's belief that her decision to see a movie and not a play is up to her is, even in a deterministic world, well founded on the compatibilist level. She did not ultimately create the sources of her motivation, and this hard

determinist insight is sometimes important, but her sense of local control is *not* illusory, although it is only part of the truth about her state. Irrespective of the absence of libertarian free will, the kleptomaniac is simply not in a condition for membership in a Community of Responsibility of which most people, having the required control, *can be*, and *would want to be* members.

The hard determinist is right to say that any punishment is in some sense unjust, but wrong when she denies that some punishments are more unjust than others because of the issue of compatibilist control. The eradication of free will-related distinctions does not make the hard determinist more humane and compassionate, but morally blind and a danger to the conditions for a civilised, sensitive moral environment. The same distinctions also apply to individual acts: there is a sense in which you may be properly blamed for not coming to class if you did not feel like coming and did not overcome this urge, but not if you were abducted (unless being abducted resulted from factors under your control). To blame a person who was abducted would be barbarous. The realisation of the absence of libertarian free will does not erase all these moral distinctions. We must take account of them and maintain the Community of Responsibility, in order to respect persons. An order not taking account of the compatibilist level distinctions is morally monstrous and inhuman.

Similarly, once we grant the compatibilist that his distinctions have *some* foundation and are partially morally required, there is no further reason to go the *whole* way with him, to claim that the absence of libertarian free will is of no great moral significance, and to deny the fact that without libertarian free will even a vicious and compatibilistically-free criminal who is being punished is in some important sense a victim of his circumstances. If we reflect upon the fact that many people are made to undergo acute misery while the fact that they have developed into criminals is ultimately beyond their control, it is hard to dismiss this matter in the way compatibilists are wont to do. Given the absence of libertarian free will, the appropriate notion of justice incorporates pity into the very fabric of justice. There is a sense in which Compatibilist Justice is very often, at best, 'justified injustice', and in which the proper compatibilist order can be seen as, in one way, morally outrageous.

The valid requirement to form, maintain, and enhance this moral order is hence tragic.

3. THE SHALLOWNESS OF COMPATIBILISM

Let us review where we are. We are assuming that there is no libertarian free will. Following Smilansky (2000), the implications were seen to be inherently dualistic: neither the compatibilist interpretation nor the hard determinist one cover all that needs to be said. Fundamental Dualism hence means that both traditional approaches are inadequate. Of hard determinists who reject all possibility of making sense of moral responsibility and of a moral order based on it we say that they are over-reacting. Here the compatibilists will be quick to agree. This side of the debate is not the topic of this paper. Of compatibilists who rest content with moral responsibility and the concomitant notions even in the face of the absence of libertarian free will, we say that they are morally, ‘personally’ and even pragmatically shallow. This does not follow from some error in compatibilist presentations of their views; on the contrary, compatibilists today present sophisticated positions and have made genuine philosophical progress. Nevertheless, the inherent limitations of the compatibilist perspective mean that it is shallow. Now we shall proceed to defend this charge.

3.1. *The Ethical Shallowness of Compatibilism*

Since in our presentation of the dualistic picture we have focused upon moral responsibility and justice, the materials for seeing the moral shallowness of compatibilism are close at hand. Let us focus on an individual criminal who is justly being harmed, in terms of Compatibilist Justice. Even if this criminal significantly shaped his own identity he could not, in a non-libertarian account, have created the original ‘he’ that formed his later self (an original ‘he’ that could not have created his later self differently). If he suffers on account of whatever he is, he is a *victim* of injustice, simply by being. Even if people can be morally responsible in compatibilist terms they lack ultimate responsibility: this lack is often morally significant, and in cases such as the one we have considered having people pay dearly for their compatibilistically-responsible actions

is unjust. Not to acknowledge this prevailing injustice would be morally unperceptive, complacent, and unfair.

Consider the following quotation from a compatibilist:

The incoherence of the libertarian conception of moral responsibility arises from the fact that it requires not only authorship of the action, but also, in a sense, authorship of one's self, or of one's character. As was shown, this requirement is unintelligible because it leads to an infinite regress. The way out of this regress is *simply to drop* the second-order authorship requirement, which is what has been done here. (Vuoso, 1987, p. 1681) (my emphasis)

The difficulty, surely, is that there is an *ethical basis* for the libertarian requirement, and, even if it cannot be fulfilled, the idea of 'simply dropping it' masks how *problematic* the result may be in terms of fairness and justice. The fact remains that if there is no libertarian free will a person being punished *may suffer justly* in compatibilist terms for what is ultimately her luck, for what follows from being what she is – ultimately without her control, a state which she had no real opportunity to alter, hence not her responsibility and fault.

Consider a more sophisticated example. Jay Wallace maintains the traditional paradigmatic terminology of moral responsibility, desert, fairness and justice. Compatibilism captures what needs to be said because it corresponds to proper compatibilist distinctions, which in the end turn out to require less than incompatibilist stories made us believe. According to Wallace, "it is reasonable to hold agents morally accountable when they possess the power of reflective self-control; and when such accountable agents violate the obligations to which we hold them, they deserve to be blamed for what they have done" (p. 226).

I grant the obvious difference in terms of fairness that would occur were we to treat alike cases that are very different compatibilistically, say, were we to blame people who lacked any capacity for reflection or self-control. I also admit, pace the incompatibilists, that there is an important sense of desert and of blameworthiness that can form a basis for the compatibilist practices that should be implemented. However, the compatibilist cannot form a sustainable barrier, either normatively or metaphysically, that will *block* the incompatibilist's *further* inquiries, about all of the central notions: opportunity, blameworthiness, desert, fairness and justice. It is

unfair to blame a person for something not ultimately under her control, and, given the absence of libertarian free will, ultimately nothing *can* be under our control. Ultimately, no one can *deserve* such blame, and thus be truly blame-*worthy*. Our decisions, even as ideal compatibilist agents, reflect the way we were formed, and we have had no opportunity to have been formed differently. If in the end it is only our bad luck, then in a deep sense it is not morally our fault – *anyone* in ‘our’ place would (tautologically) have done the same, and so everyone’s not doing this, and the fact of *our* being such people as do it, is ultimately just a matter of luck. Matters of luck, by their very character, are the opposite of the moral – how can we ultimately hold someone accountable for what is, after all, a matter of luck? How can it be fair, when all that compatibilists have wanted to say is heard, that the person about to be e.g. punished ‘pay’ for this?

Without libertarian free will, no matter how sophisticated the compatibilist formulation of control is, and whether it focuses on character, reflection, ability to follow reasons or anything else available at the compatibilist level, in the end *no one can have ultimate control over that for which one is being judged*. While ‘forms of life’ based on the compatibilist distinctions about control are possible and morally required, they are also superficial and deeply problematic in ethical and personal terms. When ultimate control is seen to be impossible, we must take notice.

Note that my own claim is not that Ultimate Injustice is all that there is of free will-related justice: I acknowledge that compatibilism captures much of fairness and justice, and indeed that compatibilist distinctions are to have a dominant role in establishing social practices. The compatibilist cannot dismiss the dualist by saying, as she might attempt to do in response to the hard determinist, that the hard determinist is denying the manifest moral differences between cases and the very conditions for a civilized moral order. The dualist, after all, *acknowledges* all this. What the dualist resists is the claim that this is all that matters, and that the ultimate arbitrariness of it all is, somehow, of no moral import. The proper description of such a case is indeed dualistic: given that we need to order social life within certain constraints, we are obliged to follow compatibilist distinctions in terms of control and its absence,

if we are to respect persons. But those who pay the price, by ultimately acting as they have been molded, are in the end victims as well. Their treatment is hence, on a deep view, manifestly morally disturbing. This needs to be acknowledged.

Let us take a step back and reflect on the compatibilist task. There are two ways in which this task can be understood: first, as a project of defending the compatibilist distinctions, of making some sense of what the compatibilists value. Hence, focusing on this task, then, once they manage to show the unreasonable reductionism of the typical incompatibilists, who group together the compatibilistically free and unfree, compatibilists understandably think that they have triumphed. Any claim that something is amiss is dismissed as the denial of the obvious, namely, that compatibilist distinctions in term of control *are* manifestly salient. However, as my dualistic argumentation showed, it is not sufficient for compatibilists to show the limitations of those who deny their case: there is a second compatibilist task, that of showing that the compatibilist captures *everything* important that is at stake. And while on the first task the compatibilists are successful, this does not seem to be the case with the second one. The dualistic case I presented allows, as it were, the compatibilist to enter the house, but denies her claim to take full control of it. Valid compatibilist insights need to share the pertinent philosophical accommodation with hard determinist insights.

3.2. *The Personal-Existential Shallowness of Compatibilism*

The free will issue combines a wealth of moral and amoral notions: not only moral responsibility, blameworthiness, desert and justice, but also (self-)respect, love, gratitude, appreciation, resentment, and so on. I cannot consider here in detail how matters appear beyond morality and justice, but it is important to see that even if compatibilism is viable on some shallow level with respect to social arrangements, it is dubious with respect to much of our free will-related emotional life. Take, for example, self-respect and respect for others. A similar course of argument to that which I have made concerning blameworthiness and justice can be made concerning (self-)respect.

Compatibilism is a shallow position because the basis for human self-respect and respect for others, in central senses of respect, is

dependent upon what one freely does. The Fundamental Dualism means that we can make some sense of compatibilist distinctions, e.g., one can be respected for her hard work but not for her height or skin colour. The social order certainly should express and follow the moral importance of such compatibilist distinctions, for reasons such as we saw when we considered moral responsibility and justice. Alongside the compatibilist level there nevertheless exists an ultimate hard determinist one, whereby ultimately all the basis for (self-)respect is a matter of how one was constituted, of one's luck. And this matters; as Michael Walzer put it: "The reflexive forms of recognition, self-esteem and self-respect, our most important possessions . . . must seem meaningless to individuals all of whose qualities are nothing but the luck of the draw" (Walzer, 1983, p. 261). Compatibilist assurances remain unconvincing and shallow here.

Once more, we need not follow the full extremity of such pronouncements. Within our lives, compatibilist distinctions can go some way towards making sense of local feelings of (self-)respect. Even one's pedestrian accomplishments through efforts in daily life – say, succeeding in losing weight through dieting – can generate self-confidence and self-respect. Nevertheless, reflection will show us that such feelings will not go very far, in the face of the free will problem. Once the ultimate perspective is taken, such local triumphs are seen within their inevitable nexus, and even if the sense of success survives, it would not go very deep. For, in the end, all is merely an unfolding of the given, and one's increasing self-respect is put in perspective. What one considers to be one's own accomplishment, the source of one's pride, sense of value, and appreciation by others, begin to seem not deeply one's achievement, but rather merely something that came into being through one as a result of factors ultimately beyond one's control.

Consider friendship: we can remain on a 'pre-free will' level in friendship, where joy in togetherness suffices (in this sense one can be a friend with a dog). Emotional attachment similarly does not require any presupposition of free will – parents can be attached to their new-born baby. We can, however, see a compatibilist level that assumes free will. Normally your friend's motives and the amount of effort she puts into the relationship clearly matter to you and to

the value of the friendship. But, once again, the hard determinist perspective looms: under the ultimate level description our view of our friends will be affected. Your friend's reactions to you are seen to follow from her 'nature', which is ultimately beyond her control. That is the sort of person she has been formed into being. Deep gratitude now seems out of place. A similar development to that witnessed with regard to respect occurs, namely, the deepest sort of appreciation withers.

3.3. *The Pragmatic Shallowness of Compatibilism*

We by now understand what the moral and personal-existential shallowness consists of: avoiding to take seriously what must be taken seriously, and not pursuing the same control-valuing principle that the compatibilist herself affirms up to the worrisome level, where that very principle comes up with the ultimate negative result. But what is pragmatic shallowness? By this I understand the corresponding lack of seriousness as to the grave dangers that recognizing the implications of the absence of libertarian free will can have. In a word, the deep practical *complacency* of compatibilism in the face of what is perhaps the most serious conceptual, ethical, and personal-existential challenge of modernity. The naturalistic-deterministic picture challenges our most central values and self-images. Compatibilists, who think that their half-way substitutes, for all of their limited validity, suffice, are simply naïve.

A number of brief illustrations should make the point. We often want a person to blame himself, feel guilty, and even see that he deserves to be punished. There is no viable picture of the moral life that can completely bypass this need. Such a person is not as likely to do all this if he internalised the ultimate perspective, according to which, in the actual world, nothing else could in fact have occurred, and he could not strictly have done anything else except what he did do. But, as we saw, it was not very difficult to enter the ultimate perspective, at least in one's thoughts. I have considered elsewhere the mechanisms through which we are saved from such catastrophic realizations (Smilansky, 2000, Chs. 7, 8), but the present point is to see that compatibilists hardly seem to notice the very difficulty. Here, it seems to me, the common person's incompatibilist intuitions, for all of their vagueness and crudeness,

have captured something that has escaped philosophical compatibilists. A realization of the absence of the sort of tacit libertarian free will that is typically assumed by most people (at least in the West) is likely to be detrimental to the acceptance of responsibility and accompanying emotions.

There are perhaps two levels at which we can consider this. One is the level at which people are in fact eager to look for excuses. How, for instance, would we be able to develop a sense of responsibility in children if there was the culturally available possibility of using the ultimate perspective as a way of escaping responsibility for one's past actions or omissions? The partial validity of compatibilism is likely to be pragmatically defeated when confronted with the eagerness to put forward the fruits of a deeper look, as an excuse. On another level, even good will and eagerness to accept responsibility would not suffice. One can surrender the right to make use of the 'ultimate level excuse' for normative reasons, and yet perhaps not be able to hold oneself truly responsible (e.g. to engage in remorse), if one has no grain of belief in something like libertarian free will. One can, after all, accept responsibility for matters that were not up to one in any sense, such as for the actions of others, for normative reasons. But here we are dealing with a different matter: not with the acceptance of responsibility in the shallow sense of 'willingness to pay', but rather with feeling *compunction*. Compunction seems conceptually problematic and psychologically (non-pathologically) dubious when it concerns matters that, it is understood, ultimately one could not help doing. But such genuine feelings of responsibility (and not mere acceptance of it) are crucial for being responsible selves.

Consider now briefly the issue of self-respect and respect for others that we touched upon before. Clearly, if people really thought of themselves or of their parents as determined outcomes of what existed a century ago (perhaps with some small random indeterminism thrown in), this would make a substantial difference to their attitudes of (self-)respect and pride. The appreciation of achievement or lack of it cannot emerge unscathed from such reflection: in retrospect, we might tend to say of an achiever that 'Well, *he had it in him*'. When applied to ourselves or to others, such deprecatory thoughts can be extremely damaging to our sense of achievement,

worth, and self-respect. For if *any virtue that one has exhibited, if all that one has achieved*, was ‘in the cards’, just an unfolding of one’s predetermined self, one’s view of oneself (or important others) cannot, surely, remain the same.²

It seems that the shallowness of the substantive compatibilist diagnosis of the situation in moral and personal-existential terms translates into blindness as to the pragmatic dangers.

4. COMPATIBILIST COUNTER-ARGUMENTS

Compatibilists can of course be expected to dispute my claim that their view is shallow. Let us attempt to consider some of the ways in which such arguments might proceed.

One argument dismisses any incompatibilist worries by claiming that we cannot make sense of what incompatibilists want. This is the familiar ‘Libertarian free will isn’t worth wanting anyway’ line (see e.g. Dennett, 1984). I cannot take up this issue in detail (cf. Smilansky, 2000, pp. 48–50), but do not think that it is worrisome. All we need is to point out the *limitations* of the compatibilist view about free will-related justice, for instance. Given that we shall continue to put people in prisons, the absence of libertarian free will means that this practice will have much shallower grounding, hence be much more unjust, than it would have been were justification based on the ‘ultimately guilty self’ possible. The severe shallowness of desert and value, or the grave injustice, that exist in a world without libertarian free will, are – for all the ‘worth wanting’ rhetoric – ethically and existentially meaningful.

A different although related thought might be that, from the ultimate hard determinist perspective, desert itself ceases to make sense: it is not that no one deserves, say, to be treated worse than others because there is no free will, but that the very notion of desert disappears. Desert-based justice, it might seem natural to hold, cannot exist from the hard determinist perspective. Hence there is no basis for understanding the very notions, primarily that of injustice, which ground my claim that compatibilism is shallow. This, however, is a mistake. As I have shown elsewhere in greater detail (Smilansky, 1996a, 1996b, 2000, pp. 43–44), the conceptual structure here is as follows:

1. We assume background conditions delineating those to whom the categories of desert and justice apply (persons of a certain capacity).
2. We assume a baseline of desert or entitlement (e.g., everyone should be treated as innocent).
3. The only way in which people can come to deserve not to get the baseline is through being responsible for not getting the baseline. Hence, if people suffer from the lack of the baseline without being so responsible, they do not deserve to suffer.
4. In terms of free will-based desert, such suffering would be unjust, and might generate second order desert or entitlement that one be compensated for not getting the baseline.

Insofar as one takes the ultimate hard determinist line, one denies that one possesses the sort of capacity for responsibility that, through one's actions, might justify one's not getting the baseline. Hence any departure from the baseline is unjust.³ Not recognizing this is morally shallow. Whether one is an absolutist hard determinist who denies that the compatibilist has any sort of case, or, like myself, a dualist who seeks to combine the true but partial insights of both compatibilism and hard determinism, there is no conceptual difficulty in conceiving of non-compatibilist injustice.

John Martin Fischer has offered another sort of argument that seeks to pull the rug from under my position (Fischer, 2001).⁴ He focuses upon my claim (when wearing my incompatibilist hat) that since everything one does is an unfolding of the given, over which one had no control, then ultimately one cannot be said to have had control over one's action. Fischer presents a number of cases where luck is paramount, such as not being abused by a close relative when one was young, something that we assume would have precluded one's development into an adequate moral agent (Fischer, 2001, p. 28). Since we do not think that this sort of counter-casual luck precludes an evaluation of the person as he is now, why care about the ultimate sort of luck that I focus upon?

The answer is that the presence of luck in itself is not the issue: after all, we are all fortunate that the earth was not hit yesterday by a huge meteorite, but that seems scarcely relevant to the free will issue. Luck is undoubtedly present in our lives, but the central question is how its presence is manifested. The problem which, I argue,

makes compatibilist control shallow (although real) is that the reach of such control in the actual sequence of events is limited: it is not so much that certain things are beyond our control (as in Fischer's sort of case), but that what *is* within our control, in compatibilist terms, is ultimately infected by arbitrariness, with what this implies about desert, moral worth, or self-respect. In other words, the difficulty is that compatibilist control is set by the way we are constituted so that every choice we make and every action we undertake is an unfolding of the arbitrarily given – the luck is not located in some corner but, when we look deeply, we see that it goes *all the way through*. Therefore, our central moral notions invite the further searching inquiry into the elements of the compatibilist control scenario, until everything ends up being ultimately a matter of luck.

Fischer might counter by saying that, after all, the counterfactuals hold at every point, even on a non-deterministic model ('if I had been significantly abused as a child . . .'); hence the alleged difficulty is not peculiar to compatibilism. But this does not damage my previous reply: compatibilism is attractive as establishing an 'Island of Control' countering random arbitrariness; a space where, as it were, arbitrariness has no entry. The sense of compatibilist control is not without some basis, but nevertheless it becomes incorporated in the deeper picture, where all is arbitrary. Rather than being merely a factor (whether counterfactual or not) on the margins of action *luck makes its way through compatibilist agency*, ultimately pervading all.

Fischer attempts to turn the table on the sceptics by pushing them to acknowledge that they themselves do not mind the presence of a large measure of luck. But this wrongly locates the sceptical worry, which is whether, if we look deeply, anything but luck *remains*. For certain purposes, compatibilist senses of local control might still be important, as we have noted, but the attempt simply to dismiss the sceptical worry about the moral implications of the ultimate arbitrariness of it all fails.

Other philosophers sympathetic to the compatibilist case have dealt with the difficulties in a different way, as it were 'changing the subject'. Farthest away from the traditional free will debate are utilitarians. In the broad sense utilitarians often wish to preserve talk about moral responsibility and blameworthiness (e.g. Sidgwick,

1963; Smart, 1961), for doing so is pragmatically useful. As Sidgwick says:

[T]he determinist can give to the terms 'ill-desert' and 'responsibility' a signification which is not only clear and definite, but, from an utilitarian point of view, the only suitable meaning. In this view, if I affirm that A is responsible for a harmful act, I mean that it is right to punish him for it; primarily, in order that the fear of punishment may prevent him and others from committing similar acts in future. (Sidgwick, 1963, pp. 71–72)

Nevertheless, it would be misleading to understand utilitarians as compatibilists. It is more correct to interpret utilitarianism as a radically revisionist position that gives no inherent weight to free will considerations; as diverging, in other words, from the very concern with free will from the start. All that matters is whether it would be useful to punish someone, and it is wrong to refer to such consequentialist concerns by using the term desert. If one is a utilitarian then he or she may well not care very much about the free will problem in the first place. However, for someone who begins with the basic values and concerns brought into focus by this perennial problem, the utilitarian option will seem almost outside the scope of morality.

Contractualists have recently made interesting attempts to accommodate much of the conceptual structure related to the traditional free will-based paradigms, but to ground it in an alternative yet non-utilitarian way (see Scanlon, 1984; Scanlon, 1998; Lenman, 2000; Lenman, 2002; cf. Smilansky, 2003a). This is an interesting attempt, for by bypassing the paradigmatic discussion contractualists may rebut the charge of shallowness I leveled against more traditional compatibilists. According to Scanlon, for instance, there is no robust pre-institutional sense of desert (1998, p. 274). Rather, we respect choice and consider the compatibilist distinctions because this is mandated by the contractualist concern for what reasonable persons would agree to (e.g. Scanlon, 1984, p. 166).

However, on reflection this will not do. Why, for example, must the innocent not be punished? Consider three broad forms of reply:

- a. Punishing the innocent is wrong because it transgresses the set of moral rules that will deliver maximal overall human well-being (rule-utilitarian).

- b. Punishing the innocent is wrong because it transgresses the set of moral rules that would be constructed by free, informed people interested in the regulation of their lives (contractualist).
- c. Punishing the innocent is wrong because the innocent do not deserve to be punished (free will-based desert).

Utilitarianism is notoriously dubious on the issue of the punishment of the innocent, even in its diluted rule-utilitarian forms. The reason is simple: as we saw, utilitarianism grants no inherent value to the blamelessness of the innocent, and their treatment is only a component of the master consideration, maximizing the good. Should we devise our justice system so that ten guilty will go free in order that one innocent person will not find himself in prison? Perhaps not, if the guilty are dangerous enough; and we could always clinch the argument against the innocent by upping the numbers. The innocence of the innocent not being an inherent consideration for the utilitarian, let alone a trumping one, it will surely often lose out. We do not need to posit fanciful ‘Sheriff and Lynching Mob’ type stories, considerations of efficient social planning may well suffice to dilute common legal practices designed to prevent the punishment of the innocent (see, e.g., Smilansky, 1990). I do not think that contractualism is quite as problematic as utilitarianism here. The only well developed contractualist system we have on these issues is Scanlon’s (1984, 1998), and the role of social utility considerations in it is not clear. Nevertheless, I think that it is striking to see the resemblance between the contractualist and the utilitarian formulations. Neither, it seems natural to say, is really concerned with true morality or justice, in the free will context. Only the desert-based account is unwilling to accept social arrangements whereby some innocent individual will pay the price of the social good, and makes innocence in itself an inherently superior moral concern.

The contractualist perspective is some distance away from common concerns with free will, morality and justice. From the perspective of genuine concern with morality and justice in the free will context, contractualist revisionism looks not very different than utilitarian revisionism. It may have a role to play in our ethical life, but in free will-related matters we cannot give up the traditional paradigm, working with it as far as we can along compatibilist

lines – and then acknowledging the limitations of the compatibilist’s resources.

Finally, we need to address the ‘compatibilist revisionist’ element in P.F. Strawson’s well-known “Freedom and Resentment” (Strawson, 1981). My discussion here summarizes points I have made previously in greater detail (Smilansky, 2000, Ch. 9, 2001), where I also referred to the valuable literature that P.F. Strawson’s seminal paper has generated. Strawson sets out to suggest ways in which the sceptic’s view about moral responsibility can be countered. His approach primarily centres around our ‘reactive attitudes’ to people’s good or ill will towards us (e.g. gratitude, resentment, forgiveness, indignation). Considering the predominance of human reactive attitudes and their centrality in human life, indeed, in being human, any intellectual considerations, such as the truth of determinism, cannot seriously be posed as a threat. And, even if we can imagine having a choice whether to engage in interpersonal relations, founded as they are on reactive attitudes, rational choice would be based on the expected gains and losses to human life, and the outcome would be clear (p. 70).

This view does not adequately recognize, in my opinion, that it is the reality of compatibilist distinctions in control that influence our knowledge when a reaction is *appropriate*. Our reactive attitudes are not independent and self-validating. Even with P.F. Strawson’s paradigmatic attitude, resentment, a belief *transcending and underlying* the attitude itself seems necessary. As Joel Feinberg says: “it is clear, I think, that resentment without an ostensible desert basis is not resentment” (Feinberg, 1970, p. 71). Moreover, the needs of our reactive lives cannot be the main consideration. Justice, for instance, involves matters other than safeguarding the reactive attitudes. If we doubt whether free action exists in a significant sense, this must be crucial to the view we take of, for example, the justification of blame and punishment.

Like the consequentialist ‘effects of blame’, safe-guarding the reactive attitudes is just not the *kind* of reason bound up with free will-related moral life, and will not be recognised by most people as appropriate. Here Strawson’s position confronts the same sort of difficulties that confronted utilitarianism and contractualism. Dependency on grounds is inherent in the notions under considera-

tion, such as resentment, blame, and punishment. Reactive attitudes follow the existence of such grounds (e.g. gratitude has to be deserved), and cannot in themselves replace it.

If Strawson would retreat to a merely pragmatic ‘no need to worry in practice’ interpretation, then already no rebuttal of my claims for shallowness has occurred. But even on this pragmatic level I think that Strawson would be much too optimistic. Societies with very different conceptions of justice have existed, and even if they cannot be reinstated, significant doubts as to the justness of our own institutions, resulting in an uncaring cynicism, cannot be ruled out. Focusing too much on the emotional side of life might place the central role of free will-related beliefs in establishing practices such as just punishment, beyond our view. Our reactive lives in themselves cannot truly legitimise such practices, and cannot be thought to do so by reflective persons in the face of doubts about the relevant beliefs. We cannot easily envisage a justice system functioning as ours does today when what are commonly considered to be its pillars (‘real’ blameworthiness based on libertarian free will) are threatened by deep and widespread doubt. A similar case for practical concern can be made with respect to our sense of self-respect, and other matters (see Smilansky, 1997, 2000, sec. 6.4).

Those three very different revisionist ‘quasi-compatibilist’ attempts, the utilitarian, contractualist, and ‘reactive-naturalist’, are not successful. Morally, ethically, and pragmatically they fail to address the initial concerns that generate in the first place the urgency of the free will issue. They are supportive of my case, however, by their very existence: much of their perceived force derives, surely, from vague worries that within the traditional free will paradigm compatibilism will not succeed. Such worries, I have argued, are in one way exaggerated: traditional compatibilism can make some sense of our moral and personal worlds. Hence we need to be dualists on the Compatibility Question rather than going for any extreme position. Nevertheless, in another, deeper way, the revisionists are right, for compatibilism is indeed unsatisfactory. It is the best we can do, and much preferable to the utilitarian or some other revision, but it is nevertheless shallow.⁵

We must at once take seriously the compatibilist distinctions and insights as a foundation for our lives, and acknowledge their

shallowness. This shallowness, in the light of the free will problem, is part of the human condition.⁶

NOTES

¹ The reason why libertarian free will is impossible is, in a nutshell, that the conditions required by an ethically satisfying sense of libertarian free will, which would give us anything beyond sophisticated formulations of compatibilism, are self-contradictory, and hence cannot be met. This is irrespective of determinism or causality. Attributing moral worth to a person for her action requires that her action follow from what she is, morally (her reasons-based choices or the like). The action cannot, for instance, be produced by a random occurrence and count morally, for then it is the random occurrence that makes the crucial moral difference. We might think that two different things can follow equally from a person, but which one actually does, say, a decision to steal or not to steal, again cannot be random but needs to follow from what she is, morally. But what a person is, morally, cannot be under her control. We might think that such control is possible if she creates herself, but then it is the early self that creates a later self, leading to vicious infinite regress. The libertarian project was worthwhile attempting: it was supposed to allow a deep moral connection between a given act and the person, and yet not fall into being merely an unfolding of the arbitrarily given, whether determined or random. But it is not possible to find any way in which this can be. For a more detailed presentation of this argument, see Strawson (1994); Smilansky (2000, Ch. 4).

² In one way we need to fear most the over-reaction of the simple-minded, who may not see the compatibilist level at all, and fall from holding the false libertarian beliefs into 'pragmatic' consequentialist temptations, or an unprincipled nihilism. But since the doubts are genuine, often it will be those who are morally and intellectually serious, and emotionally deep, who will be at *greater* risk of being affected. The very absence of libertarian free will is shattering to those who realise it and have ethical and personal depth. The grounding for matters such as self-respect, deep ethical appreciation, justifying social practices on the basis of desert, true internal acceptance of responsibility and remorse, and even our view of our loved ones, is so *diluted*. Even when we overcome the dissonance with the ultimate perspective and hold true to valid compatibilist practices, once we begin to realise the absence of libertarian free will our picture of human reality, of what is going on and of what is possible, becomes disheartening. Both in conceptual and in psychological terms, the diet that the compatibilists want us to live by is poor.

³ See for example G.A. Cohen's choice-based egalitarian position (Cohen, 1988): when there is no justification for inequality based upon people's free choices, the fall-back position is strict equality. This is not to say that such a position is convincing. See, e.g., my argument for the existence here of a 'Paradox of the Baseline' (Smilansky, 2003b).

⁴ This argument was developed by Fischer (2001) as a response to my position in Smilansky (2000).

⁵ One possible and curious application of an awareness of this shallowness pertains to the puzzle of modesty. See Smilansky (2003c).

⁶ I am grateful to Iddo Landau, Daniel Statman and to an anonymous referee for *Philosophical Studies*, for helpful comments on drafts of this paper.

REFERENCES

- Cohen, G.A. (1989): 'The Currency of Egalitarian Justice', *Ethics* 99, 906–944.
- Dennett, D. (1984): *Elbow Room*, Oxford: Clarendon Press.
- Feinberg, J. (1970): 'Justice and Personal Desert', in *Doing and Deserving*, Princeton: Princeton University Press.
- Fischer, J.M. (2001): 'A Matter of Luck or Judgment', *Times Literary Supplement*, No. 5143 (26 October).
- Lenman, J. (2000): 'Contracting Responsibility', in T. van den Beld (ed.), *Moral Responsibility and Ontology*, Dordrecht: Kluwer Academic Publishers.
- Lenman, J. (2002): 'On the Alleged Shallowness of Compatibilism: A Critical Study of Saul Smilansky: *Free Will and Illusion*', *Iyyun* 51, 63–79.
- Scanlon, T.M. (1984): 'Contractualism and Utilitarianism', in A. Sen and B. Williams (eds.), *Utilitarianism and Beyond*, Cambridge: Cambridge University Press.
- Scanlon, T.M. (1998): *What We Owe to Each Other*, Cambridge: Belknap Press of Harvard University Press.
- Sidgwick, H. (1963): *The Methods of Ethics*, 7th edn., London: Macmillan.
- Smart, J.J.C. (1961): 'Free Will, Praise and Blame', *Mind* 70, 291–306.
- Smilansky, S. (1990): 'Utilitarianism and the "Punishment" of the Innocent: The General Problem', *Analysis* 50, 256–261.
- Smilansky, S. (1996a): 'Responsibility and Desert: Defending the Connection', *Mind* 105, 157–163.
- Smilansky, S. (1996b): 'The Connection Between Responsibility and Desert: The Crucial Distinction', *Mind* 105, 385–386.
- Smilansky, S. (1997): 'Can a Determinist Respect Herself?', in C.H. Manekin and M. Kellner (eds.), *Freedom and Moral Responsibility: General and Jewish Perspectives*, College Park: University of Maryland Press.
- Smilansky, S. (2000): *Free Will and Illusion*, Oxford: Oxford University Press.
- Smilansky, S. (2001): 'Free Will: From Nature to Illusion', *Proceedings of the Aristotelian Society* 101, 71–95.
- Smilansky, S. (2003a): 'On Free Will and Ultimate Injustice', *Iyyun* 52, 41–55.
- Smilansky, S. (2003b): 'Choice-Egalitarianism and the Paradox of the Baseline', *Analysis* 63, 146–151.
- Smilansky, S. (2003c): 'Free Will and the Mystery of Modesty', *American Philosophical Quarterly* 40, 105–117.

- Strawson, G. (1994): 'The Impossibility of Moral Responsibility', *Philosophical Studies* 75, 5–24.
- Strawson, P.F. (1981): 'Freedom and Resentment', in Gary Watson (ed.), *Free Will*, Oxford: Oxford University Press.
- van Inwagen, P. (1983): *An Essay On Free Will*, Oxford: Clarendon Press.
- Vuoso, G. (1987): 'Background, Responsibility and Excuse', *Yale Law Journal* 96, 1661–1686.
- Wallace, R.J. (1994): *Responsibility and the Moral Sentiments*, Cambridge: Harvard University Press.
- Walzer, M. (1983): *Spheres of Justice*, Oxford: Martin Robertson.
- Williams, B. (1985): *Ethics and the Limits of Philosophy*, London: Fontana.

Department of Philosophy
University of Haifa
Mt. Carmel, Haifa 31905
Israel
E-mail: Smilsaul@research.haifa.ac.il