Document Version
Peer reviewed version

Link to publication record in King's Research Portal

Citation for published version (APA):

Citing this paper
Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher’s definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher’s website for any subsequent corrections.

General rights
Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the Research Portal

Take down policy
If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
1 Introduction

In *Justice for Hedgehogs*, Dworkin covers a great breadth of moral and ethical theory, arguing against various forms of scepticism about value, and in favour of a unified view of morality and ethics, where the former comprises those values that govern our interaction with others, and the latter comprises those values associated with the more general project of living the good life.

Central to Dworkin’s project is his rejection of any attempt to illuminate ethical concerns from a distinctly “meta-ethical” perspective. This is a bold thesis, that stands opposed to much current thinking within ethics, and his conclusions here have an important bearing on some of the other moral and ethical questions he addresses in the book.

I want to examine the relation between his rejection of meta-ethically motivated, or “external” moral scepticism and his rejection of incompatibilist conceptions of moral responsibility. I hope to endorse Dworkin’s project as a whole, and to defend his rejection of external scepticism in the face of certain problems that arise in parts of his argument. But I also want to examine his use of parallel arguments in motivating his rejection of incompatibilist theories of moral responsibility, and to argue that it’s a mistake to see incompatibilism as a parallel form of “externally motivated”
challenge. I will argue that his rejection of incompatibilism is also driven, in part, by his commitment to certain further assumptions, which turn out to be problematic. If we were to reject these assumptions, I believe this would strengthen his overall viewpoint, as well as enabling it to be consistent with a broader range of perspectives within the free will dispute.

In the following section, I will examine Dworkin’s discussion of external scepticism. The section after that will be focused on his discussion of moral responsibility.

2 Moral Scepticism

Dworkin contrasts “internal” sources of scepticism, with “external” ones. External scepticism is described as “Archimedean”. It aims to undermine all first-order moral judgements on the basis of second-order, non-moral facts, that are taken to have a bearing on morality as a whole. E.g. anthropological or scientific facts, which are not themselves moral judgements. This contrasts with “internal” scepticism, which seeks to undermine some ethical claims, but does so on the basis of endorsing others.

Dworkin argues that it’s impossible for moral scepticism to be motivated by entirely non-moral considerations. He claims, firstly, that external scepticism is incoherent, and secondly, that no factors external to morality could possibly motivate the rejection of any first-order moral judgement.

Let’s begin by examining the first of these arguments. Dworkin maintains that any statement challenging morality must itself be a moral statement; that ‘philosophical challenges to the truth of moral judgements are themselves substantive moral theories’ [1], p. 34.

In support of this claim, he asks us to consider the differences between the following statements:

A. Abortion is forbidden.
B. Abortion is required.
C. Abortion is permissible.
D. Abortion is neither required, nor forbidden, nor permissible
Dworkin argues that the first three of these claims are all *moral* claims, and that these exhaust all of logical space. The final claim is utterly unintelligible. Since genuine external scepticism would require a commitment to this latter claim, and none of the previous ones, such scepticism would be nonsense. We could defend the claim that all actions fall into the same category as C here – that everything is permissible – but then we would automatically be in the realm of *internal* scepticism. External scepticism as distinct from this is an impossibility [1], pp. 40-44.

This seems too quick. Dworkin appears to take for granted that C *is* itself a moral claim, whereas it seems ambiguous at best. To say something is “permitted” could mean *either* that the correct system of morality neither requires nor forbids it, *or* simply that it’s *not* required or forbidden. It’s not at all obvious that the latter claim is a moral one. Permissibility is unlike obligation or prohibition, in that it entails no moral imperatives. It makes no *demands* on us, so it’s strange to suppose that it *must* be understood as a “moral” category, as opposed to just an absence of any moral demands. The case for the literal incoherence of external scepticism seems weak, then.

We can, however, undermine a view without showing it to be literally incoherent. Dworkin offers much more compelling arguments to the weaker conclusion that external sources of scepticism are really irrelevant to our moral convictions. Here, it is worth closely examining the *motivations* for external scepticism.

Dworkin draws on Hume’s principle that you cannot derive an “ought” from an “is”. But he argues for a particularly strong version of this principle. He goes further than Hume in arguing for the impossibility of any negative entailment too – that is, he argues that we cannot undermine an “ought” claim with reference to any “is” claims [1], pp. 44-46. This part is significantly more controversial. If we take seriously this strong version of Hume’s principle, then there could not, even in theory, be any phenomena external to morality that could motivate a rejection of it.

Why, then, should we accept this principle? It will *certainly* be true that for any non-moral fact we care to state in favour of our scepticism (be it anthropological, evolutionary, or whatever), that no such fact will by itself imply the falsity of any moral statement. But that does *not* imply that there is *no* possible way in which our moral convictions might be undermined with reference to non-moral facts, so at first this reasoning looks rather puzzling.

To illuminate this worry, it will help to examine a parallel argument: No facts about
my psychology, or about the physical world, would imply the falsity of the claim that there is a non-physical ghost in the room with me. It’s implausible to suppose there could be any direct entailment from purely physical claims to supernatural ones. But this does not imply that such facts could never provide a good basis for me to reject my ghostly thesis. If my reason for believing there is a ghost present is to do with my believing I can see a ghost, then there are many sorts of natural fact that could undermine my belief. E.g. that the lighting conditions are set up so as to make probable certain sorts of perceptual illusion, that I am under the influence of hallucinogenic drugs, prone to paranoid delusions, etc. We might wonder, then, why there could not be parallel concerns that would justify scepticism in the case of morality.

In morality, typically cited motivations for external scepticism include the diversity of moral opinion, the fact that our actual moral convictions are better explained by accidental social and evolutionary factors than by direct appeal to truths about those convictions, and the fact that there do not appear to be moral entities – value particles or “morons” (as Dworkin calls them) that could explain the truth of our moral claims.

The reasoning here appears to be parallel to the ghost example. The sceptic points out that our beliefs are best explained not by moral truths themselves, but by extraneous, non-moral causes, and this is intended to undermine our reasons for having such beliefs in the first place. But are these phenomena really parallel, as this argument supposes? There seem to be important differences between beliefs about values, and beliefs about facts (even supernatural facts), and these differences are brought out nicely in Dworkin’s discussion.

In the case of my belief in the ghost, the justification for that belief rested crucially on the strength of evidence in favour of it. If my evidence for believing there is a ghost is that I can see one, then it’s important that the actual presence of a ghost is causing my sighting of the ghost, and my subsequent belief in it. If my sighting is much better explained by appeal to a bunch of non-supernatural factors, this would completely undermine my justification for believing there is a ghost present.

Dworkin argues that morality is different to this. Even if we found out that our moral convictions were caused in a wholly irrelevant way (e.g. that a medical scan was known to have caused us to have the belief that positive discrimination is morally okay), this would be no reason on its own to abandon that belief. Rather, we would have to subject it to rational scrutiny, and if it stood up to that test, then the cause of the belief would be irrelevant [1], p. 77-82.
Similarly, if we were to find moral particles, or “morons” it would be hard to imagine how these could be relevant to the truth of our moral convictions, the way ghosts would be relevant to the truth of our supernatural ones. Suppose we discovered that “morons” were the real cause of our beliefs about morality, we might still want to question the validity of beliefs so caused. The only way we could do that would be by subjecting those beliefs to rational scrutiny, whilst protected from the causal influence of “morons” [1], 72-76.

Our moral convictions cannot be undermined merely by the fact that they came about as a result of anthropological accident, then. Even if our moral opinions are best explained by accidental features of our society and upbringing, rational argument is the sole basis on which we could seriously undermine those convictions. In Dworkin’s view, morality is an “interpretive project”. Unlike factual convictions, moral convictions are justified by argument rather than evidence.

Whilst external scepticism is a coherent viewpoint, then, it looks like such a view would always be unmotivated. Any arguments in support of it would set up an inappropriate form of justification for morality – one that rests on an evidence-finding idea of truth, which is arguably unsuited to the task of verifying interpretive truths. A classic culprit would be Mackie’s argument that moral values would be “queer entities”, and unlikely to exist [3]. Such arguments seem to lose all force once we see moral truths as verified by argument, as opposed to being verifiable only by supposing there exist moral “entities” that bear the right sort of causal relation to our beliefs about them. If, then, we seek to undermine any moral statements, we must do so on grounds that are internal to morality.

3 The Free Will Challenge

So far, so good. But it’s not easy to fit all challenges to aspects of morality neatly into either the “internal” or “external” category. This becomes especially clear when we turn to the third part of the book, in which Dworkin discusses moral responsibility, and the challenge from the “pessimist” perspective. For the most part, Dworkin treats this as another external challenge, and it is this treatment that I want to question.

The view taken to be most problematic here is that belonging to the incompatibilist about moral responsibility and determinism. That is, any theorist who thinks that if determinism is true, then no agent could really be judgementally responsible; that no
one could ever really *deserve* blame or praise for their behaviour.

Dworkin contrasts a “causal conception” of control with a “capacity conception” of control. According to the “causal conception”, a person is ‘not in control when his decision is determined by external forces in the way determinism holds that all behaviour is’ [1] p. 228. According to the “capacity conception”, an agent is in control ‘when he is conscious of facing and making a decision, when no one else is making that decision through and for him, and when he has the capacities to form true beliefs about the world and to match his decisions to his normative personality – his settled desires, ambitions, and convictions.’ [1] p. 228.

Typically, incompatibilists maintain that causal control is necessary for responsibility, whilst compatibilists maintain that only capacity control is required. If judgemental responsibility requires causal control, then it’s easy to see why determinism poses a threat. If it merely requires capacity control, then it looks as if determinism poses no such threat.

The “pessimist” typically holds the causal view of control, which motivates an incompatibilist perspective on responsibility and determinism, and then further holds that determinism is likely to be true.

In aiming to refute incompatibilism, and particularly the “pessimist” perspective, Dworkin defends the following claims:

1. That we cannot practically take the pessimist’s picture seriously.
2. That the view can be motivated only by external factors, and in this sense is parallel to the sorts of external scepticism already discussed.
3. That there could be no internal support for the position.

Let us deal with each of these claims in turn.

3.1 Taking the Pessimist’s Challenge Seriously

Dworkin begins by arguing that we cannot possibly take the pessimist’s view seriously, either from a first-person or a third-person perspective.

From the first-person perspective, he maintains that you ‘cannot be convinced, even
intellectually, that you are not responsible for your actions, because you cannot make any reflective decision without judging which decision it would be better to make’ [1] p. 223. Again, he says ‘you cannot wrestle apart the thought “What should I do?” from the thought “Which decision would it be better for me to make?”’ [1] p. 223.

From here, he argues that attempting to apply the pessimist’s conclusion in the third-person is equally fruitless: ‘If no one ever has judgemental responsibility, then officials who treat accused criminals as responsible for their actions are not responsible for their own actions, and it is therefore wrong to accuse them of acting unfairly’ [1] p. 225.

Similarly, he maintains that ‘if determinism extinguishes our judgemental responsibility, it must extinguish our intellectual responsibility as well’. So, if pessimistic determinism is true, ‘no one could think he had made a wise decision in believing it’ [1] p. 225.

It is important to note the pattern of reasoning at work here. Since we cannot live without making value judgements, or so the argument goes, we cannot take seriously any view that undermines our judgemental responsibility for those value claims, because this would equally undermine the value claims themselves. If we are not judgementally responsible, then nothing could count as good or bad, fair or unfair, wise or unwise.

These arguments clearly rely heavily on a certain kind of inference: That if we are not judgementally responsible – if no one ever really deserves praise or blame, as the pessimist maintains – then there can be no valid values at all. Lets call this the “no-desert-no-values inference”.

It’s notable that those theorists best known as defendants of the pessimist’s viewpoint (e.g. Galen Strawson [5] and Derk Pereboom [4]) reject the no-desert-no-values inference. And not without good reason. If we examine the driving motivations behind most of these value claims, we find that they are justified largely on the basis of concerns for which judgemental responsibility is irrelevant.

First, lets consider ethical and intellectual values. We decide which actions are good based on how they will effect our lives, and whether they will further our projects and values. My justification for reading a good novel, or doing some travelling, certainly does not rest on concerns about whether I would count as judgementally responsible for doing these things. I eat because I’m hungry, not because I’ll be responsible
for eating, so it’s hard to see how facts about moral desert could possibly have any bearing on my reasons for judging these actions as good or bad. Similarly, I form beliefs on the evidence because those are likely to be true. Not on the basis that I will deserve praise or blame for forming such beliefs.

The same holds true with moral values. When deciding what the morally correct thing to do is, we tend to draw solely on concerns about the way our actions affect others. E.g. whether they cause harm, show due respect to others, could reasonably be consented to, etc. The strength of these reasons would seem to be utterly unaffected by whether we would genuinely deserve punishment if we failed to act on them. Whilst my deserving punishment if I mug old ladies might provide further reason to avoid such actions (and the likelihood of my actually incurring such punishment, further reason still), it is certainly not the prime reason. The prime reasons rest on concerns about the negative effects such an action would have on its potential victim, and that will be the same whether or not I am genuinely judgementally responsible.

The question, then, is why Dworkin thinks we must accept this inference? Strangely, we are offered very little argument on this point. He notes that some theorists reject this inference, citing Galen Strawson as an example, and he dismisses this strategy. But where we might expect some defence of an “ought implies can” principle (the usual route to this sort of conclusion), he merely states that morality is ‘an integrated web of standards’ and that judgemental responsibility is ‘the weft of all moral fabric’ [1] p. 224. But this merely restates precisely what is at issue. Unless we already have a good argument for the no-desert-no-values inference, we have no reason to believe this.

Once we see that the pessimist’s position need not imply the rejection of all values, it seems to pose no serious threat to morality as a whole, or to Dworkin’s overall project, which requires just that there be no significant disparity between our treatment of moral concerns, and our treatment of more broadly ethical ones. Since neither our ethical nor our moral values are motivated by concerns about our own judgemental responsibility, there is no reason why accepting this inference would require us to treat these two areas differently. We should reject it for both.

Without this principle, there seems to be only one way in which the incompatibilist’s view might be taken to be inconsistent with Dworkin’s project as a whole, or with his analysis of moral truth: We might suppose that this is yet another form of external challenge – a way of rejecting portions of ethics from an external standpoint. And in fact, Dworkin does appear to treat the pessimist’s viewpoint as precisely a challenge
of this sort. However, it seems wrong to suppose that the challenge here really is an external one, so let’s examine this claim next.

3.2 External Motivations

If the incompatibilist’s challenge were externally motivated, then presumably it could be dealt with in a similar vein to the sceptical challenges raised in the first part of the book. In fact, Dworkin appears to think of this merely as a more broadly ethical analogue of the sceptical doubts of Part One. He contrasts the set of metaphysical and factual questions that we must deal with in order to determine when an agent is free, and the set of moral questions we need to deal with in order to determine when it’s appropriate to exercise judgemental responsibility, and argues that this first set of factors can have no bearing on the second set. He says that because ‘Hume’s principle applies just as firmly in the ethical context it does in the moral context of Part One, no conclusion about responsibility can follow directly from any answers we give to questions in the first set’ [1] p. 222.

But is it correct to see this argument as an analogous form of external scepticism? Even if we thoroughly accepted the arguments of Part One, it would be insane to maintain that no factual matters ever had any bearing on any moral conclusions we might reach. All that has been said is that we have no reason to suppose that the validity of our moral principles depends crucially on questions about the physical world (whether there are “morons”’, etc.) And presumably the same will be true of ethical principles.

There are, however, very important factual questions to settle in order to determine whether any particular instance of an action counts as an instance of moral wrong. E.g. if we accept the principle that harming people is wrong, then we need to determine whether a particular action is harmful in order to determine whether performing that action counts as wrong, according to our principle. If everyone should become suddenly immune to harm tomorrow, this would make lots of varieties of action okay tomorrow, which would not be okay today. But the principle that harming people is wrong would be unaffected by such a change. The principle itself gives us only a conditional conclusion – that if something is an instance of harm, then it’s wrong. Whether the antecedent condition is actually met is irrelevant to the truth of the principle itself.

The conjecture that “morons” are required for morality is importantly different to
this. It violates Dworkin’s strong version of Hume’s principle, since it entails that the moral principles themselves are dependent on matters of metaphysics. A person who believes there is no morality because there are no “morons” is externally sceptical. A person who believes that an unlikely physical phenomena has made us all, in fact, immune to harm is not externally sceptical. It is not his principles, but his judgements about the relevance of those principles to actual scenarios which have been revised. I maintain that the free will challenge is not a form of external scepticism, since it’s more like the “harm” of this example, than the “morons” of the previous one.

Let’s suppose we endorse the principle that we ought not to blame anyone (ourselves included) who is unable to do otherwise. Such a principle would in no way depend for it’s truth on whether or not determinism is true. The truth of determinism would affect only our reasons for believing that there are agents who actually meet our antecedent conditions for genuine desert of blame. It would not affect the moral principle itself, so it would not violate Hume’s rule at all, so this is not an external challenge.

In fact, I believe incompatibilism is motivated by concerns that play a central role in our usual ideas about fairness and desert, and hence that there is a strong internal motivation for incompatibilism. Dworkin himself argues that there could be no such internal support, so let us turn to this point next.

### 3.3 Internal Motivations

I believe that Dworkin’s argument to the conclusion that there is no internal support for incompatibilism rests on a misunderstanding about the causal control requirement, and about its moral basis.

Dworkin seems to think that the only possible motivation for the causal control requirement draws on the differences between those agents who are usually regarded as responsible (normal healthy adults, say), and those who ordinarily are not (children, animals, the mentally ill, etc.), and would rest on the claim that the only way to explain those differences is with reference to facts about whether the behaviour in question is determined [1], pp. 237-240. But this approach seems to miss the real nature of the incompatibilist’s account of freedom, as well as the real source of motivation for it.

Clearly, the causal control view does not capture the morally important differences between adults and children, or between the sane and insane. It would be bizarre to
think that adults were somehow immune from determinism in a way children were not. However, it seems a mistake to think that the incompatibilist’s view would only have internal support if every difference in responsibility were explained by a corresponding difference in causal control.

Causal control is generally thought to be a necessary condition for moral responsibility, not a sufficient one. There is no reason at all to suppose that wherever an agent lacks responsibility, this must be because their actions are determined by factors beyond their control; to say that determinism undermines moral responsibility is not to say that there are no other factors that undermine it – such as ignorance, unsound reasoning, inability to make mature judgements, etc. The incompatibilist generally maintains that causal control and capacity control are each necessary, and only jointly sufficient for freedom. It is perfectly consistent with this view to say that being delusional undermines responsibility just on the basis that delusional people do not adequately understand their situations. Nothing in the incompatibilist’s position commits her to saying that determinism is the only factor that could possibly be responsibility-undermining.

These cannot, then, be the motivations for incompatibilism. Incompatibilists generally fall into one of two camps – those who think that it is unfair to blame an agent for actions that were unavoidable (leeway-incompatibilism) and those who think it’s unfair to blame an agent for events that do not originate with the agent themselves, but instead have their origins in factors outside of that agent’s control (source-incompatibilism). Both principles have their roots in our ideas about fairness and desert.

The idea that a person could genuinely deserve blame for behaviour that’s unavoidably fixed by events outside of that agent’s control, seems to sit uneasily with our ordinary notion of fairness. Such “moral luck” strikes many of us as paradoxical and disturbing. It seems unfair to inflict punishment purely for the sake of retribution when the victim of that punishment could not reasonably be expected to avoid the wrongdoing.

Of course, there may be other grounds on which to justify punishment: As a deterrent, say, or as a means of keeping the public safe, or as a way of encouraging reflection on a wrongdoing where this is likely to aid reform, or likely to bring about a willingness to make amends, or because of the practical benefits of having a system of legal liability, etc. But it’s important to note that none of those justifications are relevant here. To say an agent has genuine judgemental responsibility is to argue that
the agent deserves to suffer retribution, whether or not such suffering has any other beneficial consequences. When we focus just on this justification, a lack of causal control starts to look rather more obviously relevant. Even theorists who reject this condition generally acknowledge its prima facie intuitive appeal.

Dworkin, however, denies that a lack of causal control could ever appear relevant to the question of desert. He argues that agents always appear responsible when they meet the conditions of the capacity view alone: When there is the correct sort of mesh between an agent’s normative personality – their beliefs and values – and that agent’s decisions and behaviour. He argues that the only reason agents may appear to lack responsibility when subject to manipulation or hypnosis, say, is because such agents do not make genuine decisions, and do not act on the basis of their own normative personality [1], p. 239.

But this does not appear to be true in all conceivable cases of manipulation. It may be that the agent’s normative personality itself results from manipulation. In this case, the agent’s choices will result from causes which are extraneous to the agent, but the mesh between their choices and their normative personality would remain intact. It’s certainly not obvious that this sort of manipulation would be irrelevant. The point of Pereboom’s “Four Case” Argument [2], pp. 93-101, is to show how difficult it is to draw a neat line between the sorts of influences that are present in this kind of example, which seems to many to be intuitively responsibility-undermining, and the sorts of influence that are present for ordinary agents given the truth of determinism.

Whilst Dworkin argues that we would think it irrelevant if we were to find out that only some of our choices lacked causal control [1], p. 233, it would be surprising if this intuition were broadly shared. This is, after all precisely the point that’s at issue within this dispute, and an immense breadth of literature on the free will dispute shows us that opinion on this point is split, at best.

True, a great many philosophers do come to reject the incompatibilist intuition, finding the numerous difficulties associated with the incompatibilist’s account to be too costly in the end. But this is generally a conclusion arrived at after much careful reflection. History shows us quite how resilient the incompatibilist intuition has remained in the face of a very strong tradition of compatibilist arguments against the view. It’s difficult to shake this intuition precisely because it’s deeply entrenched in our moral outlook, and coheres with other core elements of our moral and ethical thinking, particularly our ideas about fairness, desert, and luck.
The causal control condition is not, then, an “interpretive orphan”, as Dworkin supposes. It only appears to be an interpretive orphan, when we are presented with its allegedly absurd consequences. But these are generated mainly by the introduction of the no-desert-no-values inference, and the treatment of causal control as sufficient for responsibility, as opposed to merely necessary for it. It appears to be these assumptions which are the real interpretive orphans, and which lack support within the rest of our moral framework. It is these assumptions, then, which I suggest we should reject.

4 Conclusion

The question we must ask now is whether any of this should seriously trouble Dworkin. I maintain that it should not. Once we see that these are really the offending claims, we also see that the incompatibilist’s perspective does not pose anything like the threat to Dworkin’s overall project that he envisages, and in aiming to refute it, he fights an unnecessary battle.

If instead, Dworkin were to reject these problem assumptions, along with the idea that incompatibilism constitutes an external challenge rather than an internal one, then the incompatibilist picture becomes relatively unproblematic. Whilst nothing I have said must commit him to accepting the incompatibilist’s perspective, I have tried to show that nothing in his system requires him to rule it out either. But we could certainly broaden the appeal of his picture by leaving this an open question.

References


