

**Freedom of the Will: A Conditional Analysis**, by Ferenc Huoranszki. New York: Routledge, 2011, 208 pages.

DOI: 10.2478/disp-2013-0030

BIBLID [0873-626X (2013) 37; pp. 368-374]

Huoranszki's *Freedom of the Will* is a book length defence of classical compatibilism, a position which affirms, as one condition on an agent's freedom, that the agent possess the ability to do otherwise. The book is a rewarding read and contains useful commentary on a number of long standing debates surrounding moral responsibility. In Part 1 of the book Huoranszki investigates how a number of foundational issues in action theory relate to the issue of responsibility. High points include an argument to the effect that the abilities relevant to free will must be extrinsic (37-41) and examples purporting to demonstrate that intentional control is neither necessary nor sufficient for responsibility (44-47). Part 2 applies the account developed in Part 1 to issues such as rationality, autonomy, reasons, and self-determination. Particularly interesting here is the claim that sensitivity to reasons is a condition, not on free will, but on autonomy (100-108).

Huoranszki's primary concern is to present a conditional analysis of free will which he does in chapter 4. The standard way of formulating such accounts is as follows: *S can  $\phi$  iff S would  $\phi$  if S chose to  $\phi$* . This kind of conditional analysis renders free will compatible with determinism, but, as van Inwagen has taught us (Peter van Inwagen, *An essay on free will*, Oxford, 1983, 121), the conditional analysis of 'can' is really nothing other than the compatibilist's central premiss. To establish compatibilism over and against incompatibilism, therefore, it is not enough to present a conditional analysis of 'can' if at the same time there are compelling arguments for incompatibilism. Huoranszki agrees with this point (57), which is why, after presenting his broad framework in chapter 1, chapter 2 kicks off with an analysis and attempted refutation of the consequence argument. Huoranszki's discussion of the consequence argument is interesting because he does not spend much time on the usual intricacies concerning the various transfer principles, but instead aims to undermine the intuitive force of the argument.

Huoranszki draws a distinction between the concept of *determinism* as used in the argument and a concept of *determination* which uncontroversially threatens freedom (29). The former is *global*, in that it refers to states of the whole universe, and *abstract*, in that it does not refer to any particular laws. In addition, the consequence argument's assertion that the 'propositions expressing any physical state of the universe at one instant and propositions expressing the totality of laws of nature imply propositions about the physical states of the universe at all other instances' is itself a *consequence* of determinism, rather than a part of the thesis (29). Huoranszki's point seems to be that the determining going on in the consequence argument is all very theoretical and far removed from our ordinary lives. To further trivialise this determining he draws a parallel between the implications in the consequence argument and the obviously banal implications that hold between other kinds of propositions. A proposition about someone's being a bachelor *implies* a proposition about someone's being unmarried, and — so the thought seems to go — *in a similarly trivial manner* a proposition about the past *implies* a proposition about the future, if determinism is true.

What the incompatibilist gets right, Huoranszki says, is a hostility towards *determination*: the idea that something *local* has caused someone to do something (28). This is what 'determinism' means in the context of ordinary language, and such local determination — psychological, social, or genetic — would indeed be incompatible with freedom. Such notions have been shown false, Huoranszki thinks. The incompatibilist argument gains traction only by trading on the ambiguities in word determinism, and when 'we realise how abstract and global the notion of determinism involved in the consequence argument is, it is already less clear how that sort of determinism can deprive us of our free will' (29).

It is true, of course, that the consequence argument employs entailments between propositions, and also true that the content of those propositions is general and abstract in the way Huoranszki describes. But it is hard to see how that is anything but a virtue: the argument applies to every agent at every time, and it is not held hostage to any empirical discoveries. It is also hard to fathom why the implications highlighted by the consequence argument would be rendered irrelevant to issues of freedom because of the 'similarity'

they bear to the implications that hold between propositions about *bachelors* and *being married*. What underlies these implications is entirely different: in one it is facts about causation and natural laws, in another it is conventions of meaning. If the implications highlighted by the consequence argument were rendered trivial by this similarity, no argument in philosophy would be safe.

One of the most thought provoking portions of the book is Huoranszki's discussion of the nature of abilities and their connection to responsibility. It is common in the philosophy of action to think of responsibility deriving from some set of basic actions which are often thought to be one's physical bodily movements. On this view, responsibility 'flows' from an agent's bodily movements to their complex actions. We are directly responsible for basic bodily actions (moving my finger) and we are derivatively responsible for complex actions and their consequences (turning on the light). Huoranszki thinks this view is mistaken. Although it might be correct to say of any complex action that we do it *by* doing some basic physical action, this *by*-locution is not the *by*-locution which connects cases of direct and derivative responsibility. Responsibility, in other words, does not originate in basic actions.

Consider the following example, which Huoranszki uses to argue for this view: suppose, intending to insult someone, I say something rude to them. My bodily movement here is a set of tongue and mouth movements, but it is highly plausible that what I'm directly responsible for is saying something rude. This is because I have no conceptual representation of the mouth movements qua mouth movements and so I *could not* make the mouth movements directly. The only way for me to make those mouth movements is by *saying those words*. What this shows is that responsibility for complex actions such as insults cannot be derivative, being built up from the responsibility from basic actions which compose them, because the basic actions which compose them cannot be performed independently of the complex action (39).

Huoranszki takes this example, and a number like it, to show that the actions we hold each other directly responsible for are specified with reference to extrinsic results, and from this he concludes that the abilities relevant to free will must also be specified extrinsically. This idea forms a theme which runs through the whole book

(32, 36-44, 62, 84-89). Despite being clear about the broad outline, however, Huoranszki's account does not fill in as many of the details as one might like. Here are some of the questions that need to be asked of the thesis:

- (1) What is the scope of the thesis?
- (2) What is the sense of extrinsic in play?
- (3) Do the examples provided support it?

Consider question (1). At the close of chapter 2 Huoranszki says that 'those actions for which we're responsible ... are *almost never* intrinsically identified' (32), but a few pages later we are told that 'the types of action for which agents are responsible *must* be extrinsically identified' (36). This might be a minor issue, but it is not entirely inconsequential: allowing exceptions would preclude the set of abilities relevant to free will being of a unified metaphysical kind, and this might preclude certain kinds of explanation for that class of ability. (A simple example suggests that we do need to allow exceptions: e.g., one can be directly responsible for the basic bodily movement of *stepping* onto a patch of grass in the vicinity of a 'Do not step on the grass' sign.)

Question (2) is more important. What is it to be extrinsically specified? Huoranszki says that abilities are 'extrinsic in the sense that the ascription of such abilities is sensitive to conditions that lie beyond the agent's body' (62), and elsewhere he is clear that abilities can be lost even when the agent undergoes 'no internal change' (85). This suggests that 'extrinsic' is to be understood as 'external' as opposed to 'relational.' And external circumstances feature centrally in Huoranszki's account of abilities: he eschews any ability/opportunity distinction as a useful way of thinking about the free will problem (31-2), saying instead that we must pay attention to when opportunities affect the possession of abilities.

But *which* extrinsic conditions affect whether an agent possesses an ability? Huoranszki says that abilities need to be *maximally specific* (24-5, 84). We are not told what this means. The implication is that a specific ability, in contrast to a general ability, will be sensitive to (more?) features of the environment. If we take *maximal* at face value we might think such a specific ability will be sensitive to *all* the ex-

trinsic conditions. But his discussion of Frankfurt style-cases shows that this is not what he means: the presence of the intervener *does not* remove the agent's ability to perform the action in question (82-9). The intervener — an actual, extrinsic factor — is not to be taken into account when assessing the agent's ability. (Interveners are sometimes described as *counterfactual*. This is unfortunate and leads to many misunderstandings. It is only the intervener's intervention which might properly be described as counterfactual). Compare the above with Huoranszki's judgement of Locke's man unknowingly put in a locked room case. Here Huoranszki says that the man *does not* stay in the room of his own free will (32); the locked door removes the man's ability to leave. This is problematic for Huoranszki for these cases appear analogous. In both we have an extrinsic factor which blocks the agent doing something. Locke's man *will* remain in the room and in the Frankfurt-style case the result the intervener wants *will* occur. In neither case is there anything the agent can do to avoid the end result. Not only, therefore, do we lack a principled way to judge which extrinsic circumstances to use when assessing an agent's ability, we seem to have a set of analogous cases for which Huoranszki has given differing judgements.

Consider now question (3): does the example support the thesis? What seems immediately clear is that the speech example is a strong counter-example against the basic/non-basic action distinction as drawn by Danto. This is because the agent cannot conceptualise the mouth movements but they can conceptualise insult. But this does not show that the agent's *possession of the ability* depends on extrinsic facts about the agent's environment, it shows only that the agent's representation needs to be about some external state of affairs.

Let us move now to core of the account, the conditional analysis of free will, Huoranszki's version of which runs as follows (66):

S's will is free in the sense of having the ability to perform an actually unperformed action A at t iff S would have done A, if

- (i) S had chosen so, and
- (ii) had not changed with respect to her ability to perform A at t, and

- (iii) had not changed with respect to her ability to make a choice about whether or not to perform A at t.

The antecedent of the conditional contains three conditions, in contrast to the usual one. As Huoranszki notes (55), however, Moore's own account of free will included two conditions over and above the simple conditional that it is sometimes thought to consist in. Huoranszki doesn't think Moore's further conditions were adequate, but takes himself to be improving on the model supplied by Moore. How does Huoranszki's account fare? The objection deemed decisive in burying the simple conditional analysis was articulated with great clarity by Lehrer (Keith Lehrer, 'Cans without ifs', *Analysis* 29 (1), 1968: 32), and the key thought is this: the truth of the conditional is not sufficient for the truth of the ability ascription. Lehrer's example has become prodigious in the literature: suppose I am presented with a bowl of red candy, and while I might like candy in general, and am not paralysed, I have a pathological aversion to taking one of these candy because they remind me of drops of blood. The following two things are true:

I cannot take a red candy.

I would take one, if I choose to.

This suffices to show that 'I can' cannot mean 'I would ..., if ...'. Huoranszki's clause (iii), which affirms not just that the agent is unchanged in their ability state but that they do indeed possess the ability, is introduced explicitly to address this problem. The clause works by conceding Lehrer's point that 'would if I choose' only stands a chance of being part of the correct analysis of 'can' if 'I can choose' is also included.

Crucially, clause (iii) includes reference to the agent's *ability to choose*. Another ability. And one which cannot be analysed using the above account on pain of circularity. Not that Huoranszki advocates such a thing. He maintains that the ability to choose is an entirely different kind of ability because *making a choice is not an action* and the control we have over our choices is different to the control we have over our actions (47, 51).

This is what allows Huoranszki to resist any charge of circularity. In bottoming out in a (putatively) non-actional component, his account is akin to that of Davidson's, who cited a belief/desire pair as the antecedent of the conditional. But whereas beliefs and desires are uncontroversially non-actional, choices are far from being so. The key, Huoranszki thinks, is to understand *choice* as referring not to a stretch of deliberation but to the end result, the 'coming to a practical conclusion' (51). This is non-actional because agents do not control the results of choices in the same way they control the actions they perform. In support of this Huoranszki invokes Locke's famous point, namely, that once the possibility of an action has occurred to us, our choice concerning it cannot be free: we can choose whether or not to do it, but we cannot choose whether or not to choose about it. These claims are contentious, but concede them for argument's sake. Huoranszki's account still has a major problem because alongside the non-actional aspect of choice he is keen to countenance the actional aspect: he speaks of deliberation as an *act* (51), and he refers to the *activity of choice making* as both intentional and voluntary (52-3). Whatever we call this actional component of choice, Huoranszki's account needs to apply to it. This will either make the account circular or, if it is argued that the conditions of responsibility for this kind of action are for some reason different (which itself risks being ad hoc), the account will be incomplete.

Despite the above problems, Huoranszki's account is a strong defence of compatibilism. The high level of detail in many sections of the book will repay careful study.

Simon Kittle  
Department of Philosophy  
University of Sheffield  
45 Victoria Street  
Sheffield  
S3 7QB  
United Kingdom  
simon@kittle.co.uk