

# LOCKE ON THE FREEDOM TO WILL

SAMUEL C. RICKLESS

## Introduction

In Book II, Chapter xxi of *An essay concerning human understanding*, Locke claims that a mind's will is its power 'to order the consideration of any *Idea*, or the forbearing to consider it; or to prefer the motion of any part of the body to its rest, and *vice versa* in any particular instance' (II. xxi. 5).<sup>1</sup> To exercise this power (that is, to will), Locke says, is to perform an act of volition (or: willing), volitions being actions of the mind directed to the performance of other actions (of the mind or body). Locke claims further that a man's being free (or at liberty) with respect to some action A is his having the power to perform or to forbear performing A according to his volition in the matter (8). Thus, X is free with respect to A *if and only if* (i) X is able to perform A if he wills to perform A and (ii) X is able not to perform A if he wills not to perform A. This analysis of freedom enables him to argue that prisoners are not free, because they are not able to perform the action of escaping towards which their volitions are directed.

Since volitions are *bona fide* actions, the question arises whether humans are free to perform them. Locke takes up this question in sections 23–25. He begins by distinguishing between two ways in which the question might be understood. In the first place, to ask whether a man is free to will might be to ask

<sup>1</sup> All the passages quoted below appear in Book II, Chapter xxi. Henceforth, I will refer to each passage by section number only. So, for example, instead of 'II. xxi. 23', I will simply write '23' or 'section 23'. I will also refer to the first five editions of the essay as E1 through E5.

whether he is free *'in respect of willing any Action in his Power once proposed to his Thoughts'* (23).<sup>2</sup> Let's call this question 'Q1' and Locke's original (E1) answer to it 'A1'. In the second place, to ask whether a man is free to will might be to ask whether he is free *'to will which of the two he pleases, Motion or Rest'* (25). Let's call this question 'Q2' and Locke's original (E1) answer to it 'A2'. Q1 is addressed in sections 23–24, Q2 in section 25.<sup>3</sup>

Having argued for A1 and A2 (about which more below), Locke spends the rest of II. xxi discussing the question, 'what determines the Will?', that is, 'What moves the mind, in every particular instance, to determine its general power of directing, to this or that particular Motion or Rest?' (29). In E1, Locke claims that the will is determined by 'the greater Good' (E1: 29). As is well known, Locke abandoned this view in subsequent editions. Starting with E2, Locke insists that the will is always determined by 'some uneasiness' (E2–5: 29), by which he means a 'pain of the body' or 'disquiet of the mind' (E2–5: 31). Locke changed his mind on this issue because he came to see that 'the greater good, though apprehended and acknowledged to be so, does not determine the will, until our desire, raised proportionably to it, makes us *uneasy* in the want of it' (E2–5: 35). That is, Locke saw that many people are subject to weakness of will: although they know perfectly well what is best for them, these people do the opposite instead.

<sup>2</sup> In E5, this question is worded somewhat differently. There, Locke asks whether a man is free *'in respect of willing, or the Act of Volition, when any Action in his power is once proposed to his Thoughts, as presently to be done'*. Whether Locke thought his rewording of the question significant is something I will be taking up below.

<sup>3</sup> I argue below that A1 is NO and A2 is YES. But there is some controversy about this, as we shall see, and I do not wish to beg any questions at the outset.

Now Locke thinks that the mind is full of different *'uneasinesses* always solliciting, and ready to determine the will'. Since the will is determined by some uneasiness or other, the question arises, by which of all these competing uneasinesses is the will ultimately determined? Locke's answer is that the will is determined by 'the most pressing' uneasiness, though this is 'for the most part, but not always'. As experience teaches us, the mind in most cases has the 'power to *suspend* the execution and satisfaction of any of its desires', that is, to suspend willing the performance of actions required to fulfil its desires. As a result, the mind is free to 'examine [the objects of these desires] on all sides, and weigh them with others'. The result of such an examination may be the 'raising' of a new uneasiness (perhaps a pain in the want of some absent good) which trumps the uneasiness that was originally seen to be 'most pressing' (see E2–5: 47).

Following Vere Chappell, let's call the thesis that the mind (in most cases) has the power to suspend the prosecution of its desires, 'the Doctrine of Suspension' (DS):<sup>4</sup>

(DS) The mind (in most cases) has the power to suspend willing the performance of actions required to fulfil its desires.

Now, much as recent commentators acknowledge the plausibility of this thesis, it has seemed to some that the addition of DS to the mix of Locke's views on freedom was pretty catastrophic. Chappell, for one, has recently argued for the following claims: (i) that DS conflicts both with A1 and with one of the premisses in Locke's main argument for A1; (ii) that Locke eventually (namely, by the time of E5) recognized the existence of this

<sup>4</sup> All references to Chappell are to his 'Locke on the freedom of the will', in *Locke's philosophy: Content and context*, ed. G. A. J. Rogers (Oxford 1994), pp. 101–21; reprinted in *Locke*, ed. V. Chappell (Oxford 1998).

conflict, and consequently repudiated A1 (even while he did not eliminate the passage of section 23 in which A1 is stated and defended); (iii) that Locke's argument for A2 relies on a premiss that he himself explicitly repudiated; (iv) that, although it is not at all clear that DS conflicts with A2, Locke eventually came to believe that these two theses were incompatible; and (v) that the perceived conflict between DS and A2 led Locke to abandon A2 explicitly in E5 (even while he did not eliminate the passage of section 25 in which A2 is stated and defended).

If Chappell is right, Chapter xxi is a mess. First, having repudiated A1, Locke should have altered sections 23 and 24 accordingly. But, according to Chappell, this is something Locke failed to do. Rather, he kept DS and A1 in the E5 text of Chapter xxi, knowing full well that they were inconsistent. Second, since his argument for A2 rests on a premiss that he did not accept, Locke is guilty of intellectual laziness at best, blatant inconsistency at worst. And, third, having repudiated A2, Locke should have altered or dropped section 25 in E5. But this, too, is something Locke failed to do.

Now it is widely believed that Locke was a commonsensical, but noticeably less than systematic, philosopher; and Chappell's account of Chapter xxi reinforces this common presumption. Nevertheless, despite its attractiveness, the account is, I believe, mistaken. My aim in this paper is to establish that Locke did not err in the ways Chappell suggests. More particularly, I will argue for the following: (a) that, although DS and A1 are indeed inconsistent, Locke came to see that A1 was an overgeneralization that would have to be restricted; (b) that Locke effected the required restriction by means of explicit textual emendations in E2 and E5; (c) that the restricted version of A1 and all the premisses in Locke's main argument for it are perfectly consistent with DS; (d) that Locke's argument for A2 does not rely on a premiss which he explicitly (or implicitly) repudiated; and,

finally, (e) that, as Locke himself recognized, A2 and DS are, in fact, perfectly consistent.

The plan of the paper is as follows. I begin by considering sections 23 and 24, with a view to establishing (a), (b), and (c). I then consider section 25 and argue for (d) and (e). I conclude with a summary and a brief discussion of the philosophical significance of these results in the context of Locke's overall account of human freedom.

### The Unavoidability Thesis

In E1, section 23 appears as follows (I have divided the section into three parts, [A]–[C], for ease of reference):

[A] *Willing*, or Choosing being an Action, and Freedom consisting in a power of acting, or not acting, a Man in respect of willing any Action in his power once proposed to his Thoughts, cannot be free. [B] The reason whereof is very manifest: for it being unavoidable that the Action depending on his Will, should exist, or not exist; and its existence, or not existence, following perfectly the determination, and preference of his Will, he cannot avoid willing the existence, or not existence, of that Action; it is absolutely necessary that he will the one, or the other, *i. e. prefer* the one to the other: since one of them must necessarily follow; and that which does follow, follows by the choice and determination of his Mind, that is, by his *willing it*: for if he did not will it, it would not be. So that in respect of the act of willing, a Man is not free: Liberty consisting in a power to act, or not to act, which, in regard of Volition, a Man has not: it being necessary, and unavoidable (any Action in his power being once thought on) to prefer either its doing, or forbearance, upon which preference, the Action, or its forbearance certainly follows, and is truly voluntary. [C] So that to make a Man free in this sense, there must be another antecedent Will, to determine the acts of this Will, and another to determine that, and so *in infinitum*: for where-ever one stops, the Actions of the last Will cannot be free: Nor is any Being, as far as I can comprehend Beings above me, capable of such a freedom of Will, that it can

forbear to *Will, i. e.* to prefer the being, or not being of any thing in its power, which it has once considered as such.

In 23A, Locke provides us with A1, his answer to Q1. In 23B he produces an argument for A1, and, it seems, in 23C produces another. What, then, is A1? And what are Locke's arguments for it?

A1 is the thesis that a man considering an action is not free in respect of the act of willing. According to Chappell, and on this we agree, this claim is to be identified with what Chappell (106) calls 'the unavoidability thesis' (UT):

(UT) A man who thinks about doing something is unavoidably bound to will one way or the other regarding it.

Chappell then claims (again, rightly) that 23B contains the following constructive dilemma in support of UT:

- (1) Every action a man considers doing must either exist or not exist.
- (2) If the man considering such an action did not will it to exist, it would not exist.
- So (3) If such an action exists, the man wills it to exist. [From 2]
- (4) If the man considering such an action did not will it not to exist, it would not exist.
- So (5) If such an action does not exist, the man wills it not to exist. [From 4]
- So (6) Every such action must either be willed to exist or be willed not to exist by the man considering it. [From 1, 3, 5]
- (7) If a man wills an action to exist, he wills.
- (8) If a man wills an action not to exist, he wills.
- So (9) A man considering an action must will. [From 6, 7, 8]
- (10) If a man considering an action must will, then he is not free in respect of the act of willing: he cannot avoid willing.
- So (11) A man considering an action is not free in respect of the act of willing: he cannot avoid willing. [From 9, 10]

(Chappell does not tell us exactly how the sub-conclusions and conclusion of the argument are supposed to follow from previous premisses and sub-conclusions. I have inserted the required annotations to make the validity of the argument perspicuous.)

Now Chappell rightly claims (following Leibniz) that DS, which Locke introduced in E2, conflicts with one of the dilemma's premisses (namely, (4)) and with its conclusion as well. For, first, imagine that Fred is considering action A but suspends choosing whether or not to perform it. In such a case, A might fail to exist, even though Fred did not will it not to exist. So premiss (4) of the dilemma is false. Moreover, the very same case shows the conclusion to be false: for when Fred suspends choosing whether or not to perform A, he is not bound to will one way or the other regarding it.

But Chappell claims further, and it is here that we diverge, that these two inconsistencies (the one between DS and premiss (4), and the one between DS and UT) existed in the text of E2 'and continued to do so thereafter'. 'For', Chappell writes, 'Locke made no changes in the doctrine of suspension in subsequent editions of the *Essay*, nor did he excise or amend the (false) premiss of his argument in section 23' (107). I fully agree that Locke made no changes in DS, and it is also clear that he did not 'excise' premiss (4) from his argument in 23B. But, as I will now argue, he explicitly restricted the scope of both premiss (4) and UT, first in E2 and later in E5.

To understand how Locke came to give both (4) and UT a more restricted reading, we need to look at some key passages in section 24. In that section, Locke provides us with an example designed to illustrate UT. In E1-3, the case is described as follows:

[A] Man that is walking, to whom it is proposed to give off walking, is not at liberty, whether he will *will*, or no: He must necessarily prefer one, or t'other of them; walking or not walking.

In E4–5, the language is scarcely different:

[A] Man that is walking, to whom it is proposed to give off walking, is not at liberty, whether he *will* determine himself to walk, or give off walking, or no: He must necessarily prefer one, or t'other of them; walking or not walking.

This example contains two important features. First, the action being considered is the action of stopping ('giving off') another. The man who is walking is already engaged in a process, and the question that he is pondering is *whether to stop or to continue the process*: he is not considering whether to *initiate* a process. Second, it is clear that the walking man is not being asked to consider giving off walking *sometime in the future*: he is being asked to consider giving off walking *right this minute*.

Now consider what happens when (4) is restricted in such a way that the relevant actions include only *stoppings of processes* that are to be considered only *at the time they are to be performed*. (Keep in mind here that, for any process X, to say that stopping X does not exist is just to say that X continues.) The restricted version of (4) (call it '(4<sub>R</sub>)') reads as follows:

- (4<sub>R</sub>) If a man considering presently stopping a process in which he is presently engaged did not will it to presently continue, it would not presently continue.

Interestingly enough, (4<sub>R</sub>) is both plausible and consistent with DS. To see this, consider the walking man again. First, it seems perfectly plain that, if this man did not will his walking to

presently continue, it would presently stop.<sup>5</sup> Second, it is also plain that DS concerns actions that come under consideration *before* the time they are to be performed. As Locke says in the section in which DS is first introduced, the '*suspension* of any desire' occurs 'before the *will* be determined to action, and the action (which follows that determination) done' (E2–5: 47).

Now consider what happens to the argument of 23B if all of its premisses are restricted in the way of (4<sub>R</sub>). Here is how the amended argument reads:

- (1<sub>R</sub>) Every process a man considers stopping when he is presently engaged in it must either stop or continue at that time.  
(2<sub>R</sub>) If a man considering presently stopping a process in which he is presently engaged did not will it to presently stop, it would not presently stop.  
So (3<sub>R</sub>) If such a process presently stops, the man wills it to presently stop. [From 2<sub>R</sub>]

<sup>5</sup> Some might worry that this is not, in fact, 'perfectly plain'. For it might be argued that if our pedestrian is walking voluntarily, then his walking is caused and sustained by a prior volition. Thus, upon the proposal to 'give off' walking, the pedestrian might well make no decision about the matter, in which case the prior and sustaining volition to walk will continue to cause him to walk. On this account, then, it would be false that: if the man did not will his walking to presently continue, it would presently stop.

Let me briefly explain why I do not find this line of argument convincing. To begin, we are asked to suppose that the man's walking is not only *caused*, but also *sustained*, by a prior volition. How are we to make sense of this? Recall that, for Locke, volitions are mental commands or directives (5). As such, they are either occurrent events (like promulgations or decrees) or standing states. Now, on the one hand, if volitions are events, then it seems impossible for a volition to be a cause of walking that is both *prior* and *sustaining*. For a cause must exist at the time at which it *sustains* its effects, but a *prior* event no longer exists at the time at which it is meant to sustain its effects. On the other hand, if volitions are standing states, then, upon the proposal to 'give off' walking, the man must indeed make a decision about the matter: if he continues to walk, this is because he continues to *will* to walk; and if he stops walking, this is because he wills not to walk. So, either the putatively problematic supposition makes no sense, or it yields the very result I characterized as 'perfectly plain'.

- (4<sub>R</sub>) If a man considering presently stopping a process in which he is presently engaged did not will it to presently continue, it would not presently continue.
- So (5<sub>R</sub>) If such a process presently continues, the man wills it to presently continue. [From 4<sub>R</sub>]
- So (6<sub>R</sub>) A man considering presently stopping a process in which he is presently engaged must either will the process to presently stop or will the process to presently continue. [From 1<sub>R</sub>, 3<sub>R</sub>, 5<sub>R</sub>]
- (7<sub>R</sub>) If a man wills a process to presently stop, he wills.
- (8<sub>R</sub>) If a man wills a process to presently continue, he wills.
- So (9<sub>R</sub>) A man considering presently stopping a process in which he is presently engaged must will. [From 6<sub>R</sub>, 7<sub>R</sub>, 8<sub>R</sub>]
- (10<sub>R</sub>) If a man considering presently stopping a process in which he is presently engaged must will, then he is not free in respect of the act of willing.
- So (11<sub>R</sub>) A man considering presently stopping a process in which he is presently engaged is not free in respect of the act of willing. [From 9<sub>R</sub>, 10<sub>R</sub>]

The amended argument is valid, and its premisses do not conflict with DS. Its conclusion (namely, (11<sub>R</sub>)) is, in fact, a restricted version of UT (call it 'UT<sub>R</sub>')

(UT<sub>R</sub>) A man who thinks about *presently stopping a process in which he is presently engaged* is unavoidably bound to will one way or the other regarding it.

Notice now that UT<sub>R</sub>, like (4<sub>R</sub>), is perfectly consistent with DS. For, again, DS concerns actions that are to be performed *after* they are considered. So, for example, although Locke's walking man can suspend his willing to attend a play the following evening, he cannot suspend willing to stop his walking *now*, but is bound to will one way or the other regarding it.

I would argue that Locke recognized as early as E2 that he was better off with *some* restriction of the argument in 23B. My evidence for this is a short phrase added at the end of section 24 in E2. The relevant passage appears below (with the added phrase underlined):

The Mind has not a power to forbear *willing*; it cannot avoid some determination concerning [the actions in its power], let the Consideration be as short, the Thought as quick, as it will, it either leaves the Man in the state he was before thinking, or changes it; continues the Action, or puts an end to it.

The added phrase makes it plain that Locke takes the relevant principles (namely, (4) and UT) to cover *stoppings of processes*, rather than actions in general. Locke intimates that it is only with regard to continuing an action or stopping it that the mind is not able to forbear willing.

That this restriction is not adequate in the end is evident from the following consideration. It is possible for a man to consider stopping a process in which he will *later* be engaged. For example, I might now know that I will be driving to the airport tomorrow, and might now further consider whether or not to stop for breakfast before reaching the airport. With regard to the act of stopping for breakfast tomorrow, am I now unavoidably bound to will one way or the other? It seems that I am not. For, by DS, I now have the power to suspend willing to stop for breakfast tomorrow in order to consider the advantages and disadvantages of doing so.

There is strong textual evidence that Locke came to see that the restriction of his argument to the stoppings of processes in E2 was vulnerable to this objection. For, in E5, Locke decides to restrict the scope of his argument in 23B even further. In between the example of the walking man and the passage just cited, he adds the following (the added material is underlined):

And so it is in regard of all other Actions in our power so proposed, which are the far greater number. For considering the vast number of voluntary Actions, that succeed one another every moment that we are awake, in the course of our Lives, there are but few of them that are thought on or proposed to the Will, 'till the time they are to be done: And in all such actions, as I have shewn, the Mind in respect of willing has not a power to act, or not to act, wherein consists Liberty: The Mind in that case has not a power to forbear willing ...

Locke is now explicitly claiming to have 'shewn' that it is only regarding voluntary actions that are proposed to the will *at the time they are to be done* that the mind is not free in respect of willing. These, says Locke, far outnumber those actions that are proposed to the will *at a time that is earlier than the time they are to be done*.<sup>6</sup>

To make matters even clearer, Locke adds an important passage to section 56 of E5:

Liberty 'tis plain consists in a Power to do, or not to do; to do, or forbear doing as we *will*. This cannot be deny'd. But this seeming to comprehend only the actions of a Man consecutive to volition, it is farther enquired, whether he be at Liberty to *will*, or no? and to this it has been answered, that in most cases a Man is not at Liberty to forbear the act of volition; he must exert an act of his *will*, whereby the action proposed, is made to exist, or not to exist. But yet there is a case wherein a Man is at Liberty in respect of *willing*, and that is the chusing of a remote Good as an end to be pursued. Here a Man may suspend the act of his choice from being determined for or against the thing proposed, till he has examined, whether it be really of a nature in it self and consequences to make him happy, or no.

<sup>6</sup> Notice also that Locke's restatement of Q1 in E5 includes the addition of the phrase '*as presently to be done*' (see footnote 2).

Chappell cites this passage as evidence that Locke 'takes the doctrine of suspension to conflict with his unavoidability thesis' and abandons the latter as a result (118). Although I agree with Chappell that DS conflicts with UT (as originally stated) and that Locke abandons UT (so stated) in E5, I believe that this passage was written to make a different point, namely that DS does not conflict with UT<sub>R</sub>. Again, Locke emphasizes that 'in most cases' (those cases, I suggest, in which the action is to be done at the time it is considered), 'a Man is not at Liberty to forbear the act of volition'. Here he adds that DS allows for cases in which 'a Man is at Liberty in respect of *willing*', but insists that these are cases to which UT<sub>R</sub> does not apply. For the action to be considered in these relatively rare cases is the pursuit of a remote good, that is, a good to be attained sometime in the future.

Let me now summarize the picture that I have been painting. In E1, Locke proposed an argument for UT in 23B, and illustrated UT by means of his 'walking man' example in section 24. The argument of 23B depends on an assumption (namely, (4)) that conflicts with DS. Locke, who had not accepted DS in E1, realized later that DS was true, and rewrote sections 28ff. of Chapter xxi accordingly (in E2). Looking back on sections 23–24, Locke realized that UT and DS were incompatible, but also saw that his 'walking man' example supported a more restricted version of UT. So he added a phrase to section 24 to remind us that we should read the argument of 23B as covering only the stoppings of processes. But Locke eventually saw that this addition was inadequate. So, in E5, he added a longer passage to section 24 explicitly indicating that the conclusion of his argument was UT<sub>R</sub>, not UT. The final result of Locke's efforts is a perfectly consistent theory composed of UT<sub>R</sub> and DS, along with an argument for the former that does not conflict with the latter.

## Volitional Determinism

The careful reader will no doubt have noticed that I haven't yet discussed 23C, in which Locke appears to offer an infinite regress argument in defence of UT. As it happens, Locke kept this passage in E1 through E3, amended it slightly in E4, and then dropped it completely in E5. So, if 23C does indeed contain an argument for UT, the story that I have been telling turns out to be no more complimentary to Locke than Chappell's. For it would appear that Locke retained at least one argument for UT, even while abandoning UT itself, in E2 through E4. Is there a way around this difficulty?

There is. For it might be claimed that the argument of 23C isn't an argument for UT after all. In the rest of this section, I will consider two ways of defending this claim. The first, proposed by Chappell, will be found wanting, the second, not.

Here, to begin, is the relevant passage as it appears in E4:

Besides, to make a Man free after this manner, by making the Action of *Willing* to depend on his *Will*, there must be another antecedent *Will*, to determine the acts of this *Will*, and another to determine that, and so *in infinitum*: For where-ever one stops, the Actions of the last *Will* cannot be free.

Chappell finds three implicit premisses in this passage. The first (call it 'Voluntariness') is that every act with respect to which an agent is free is voluntary. The second (which Chappell dubs the 'Heteronomy Principle') is that no will determines itself, that is, no will determines the acts of willing that belong to it. The third (which Chappell dubs the 'Inheritance Principle') is that an agent is free with respect to an act, only if he also is free with respect to the act of willing which produces that act. Chappell then reconstructs the argument of 23C as follows (I have simplified slightly to avoid redundancy):

- (1) Acts of willing are acts.
  - (2) Every act with respect to which an agent is free is voluntary. [Voluntariness]
  - (3) An act is voluntary iff there is an act of willing which produces it.
  - (4) An act of willing is voluntary iff there is a will which determines it.
  - (5) Every act of willing belongs to a will.
  - (6) The act of willing which produces a voluntary act of willing belongs to the will which determines that act.
  - (7) No will determines the acts of willing which belong to it. [Heteronomy]
  - (8) An agent is free with respect to an act, only if he also is free with respect to the act of willing which produces that act. [Inheritance]
- So (9) For any voluntary act of willing, there is an infinite series of wills running back from the will to which it belongs. [From 1–8]
- (10) An infinite series of wills running back from the will to which a specified act of willing belongs is an absurdity.
  - (11) There are no absurdities.
- So (12) No agent is free with respect to an act of willing. [From 1, 2, 9, 10, 11]<sup>7</sup>

Chappell claims that the conclusion of the argument of 23C is not UT, but rather a thesis he calls 'Volitional Determinism' (VD):

(VD) No agent is free with respect to an act of willing.

<sup>7</sup> Chappell claims that his reconstruction of the argument of 23C is valid (110). It isn't, in fact, for (9) does not follow from (1)–(8). Intuitively, the problem is that there is nothing to stop the series of wills generated by (1)–(8) from looping back on itself. The possibility of such a loop makes it impossible to infer from (1)–(8) that the series of wills running back from the will to which the first act of willing belongs must be *infinite*. So an additional premiss is needed to preclude the possibility of such loops. Further discussion of this matter would take us too far afield, since the matter of the argument's validity is tangential to the issues discussed in this paper.

This is something I will question shortly. But, before I do, let me emphasize from the start that Chappell is right to distinguish between UT and VD. For, first, VD does, but UT does not, attribute lack of freedom to each individual volition. And, second, VD is, but UT is not, universal in scope, ‘since not all of the actions that agents perform, or even their voluntary actions, are “proposed to their thoughts” before being done’ (107).

Against the argument of 23C, Chappell levels two criticisms. First, he argues that Locke explicitly rejects the Inheritance Principle. For, in section 33 of E1 (which is reproduced almost verbatim in section 50 of E2–5), Locke claims that a prisoner who is granted the freedom to leave his prison may yet choose to remain, where this choice is ‘absolutely determined’ by ‘the desire of some convenience to be had there’. If the fact that the prisoner’s choice is ‘determined’ entails that it is not free, then we have a counterexample to the Inheritance Principle. For the prisoner is free to remain in his prison even while he is not free to *will* to remain there. If Chappell is right about this, then Locke is in trouble. For, according to Chappell, the argument of 23C relies (even if only implicitly) on the Inheritance Principle, and so depends on a principle that Locke himself rejects.

Second, Chappell claims that Locke eventually came to believe (by the time of E5) that the conclusion of the argument of 23C (namely, VD) conflicts with DS, and consequently that VD is false. Chappell’s evidence for this is the passage from the E5 version of section 56 quoted above (p. 54). If Chappell is right, then Locke has blundered again. For this means that, in E5, Locke used section 56 to reject a conclusion for which he argued in section 23.<sup>8</sup>

<sup>8</sup> In Locke’s defence, it might be noted that Locke excised 23C from the text of E5, thereby eliminating the purported contradiction. But, as Chappell rightly points out, an argument very similar to the argument of 23C remains in section 25 of E5.

I will argue shortly that Chappell’s reconstruction of the argument of 23C is erroneous, and that the argument, rightly reconstructed, is not vulnerable to these two criticisms. But it is worth considering whether Chappell’s criticisms are well taken, even assuming that his reconstruction is accurate. On this score, I believe they miss the mark.

First, it is not clear that Locke would have accepted the case of the freed prisoner who chooses to stay in his prison as a counterexample to the Inheritance Principle. For, as Chappell himself recognizes, Locke was a *compatibilist*. That is, Locke believed that it is possible for human actions to be both determined and free. But if compatibilism is true, then the fact that a freed prisoner’s act of willing to remain in prison is ‘determined’ by some desire of his does not entail that he is not free with respect to that act.

Second, I argued above that section 56 does not indicate that Locke thought DS to be incompatible with UT. For similar reasons, I do not believe that the passage indicates that Locke thought DS to be incompatible with VD. Rather, the point of the passage is to remind the reader that DS is, in fact, perfectly compatible with UT<sub>R</sub>.<sup>9</sup>

So, even if Chappell’s reconstruction is correct, the argument of 23C is not vulnerable to the criticisms Chappell levels at

<sup>9</sup> Chappell also claims that Locke rejected VD in a letter to Philippus van Limborch (*The correspondence of John Locke*, ed. E. S. de Beer, Clarendon Edition, 9 vols. (Oxford 1976—), vol. 7, no. 3192: 28 September 1702). There Locke writes that, ‘generally ... a man is free in every action, as well of willing as of understanding, if he was able to have abstained from that action of willing or understanding; if not, not. More particularly, as regards the will: there are some cases in which a man is unable not to will, and in all those acts of willing a man is not free because he is unable not to act. In the rest, where he was able to will or not to will, he is free’. But, unlike Chappell, I read this passage as an affirmation of a restricted version of UT, not as an affirmation of VD. Notice in particular that Locke’s reference to ‘some cases’ in this passage recalls his addition of ‘in such a case’ and ‘in that case’ to sections 23–24 of E5, sections in which UT (or a restricted version thereof) is discussed.

it. For section 33 of E1 (=section 50 of E2-5) does not indicate that Locke rejected the Inheritance Principle, and section 56 of E5 does not indicate that Locke rejected VD.

However, as I will now argue, Chappell's reconstruction of 23C is incorrect. Moreover, having understood how the argument is supposed to go, we will see that Chappell's two criticisms are simply inapplicable. For, first, the argument of 23C does not rely on the Inheritance Principle, and, second, its conclusion is the very opposite of Volitional Determinism!

Recall the text of 23C (as it appears in E4):

Besides, to make a Man free after this manner, by making the Action of *Willing* to depend on his *Will*, there must be another antecedent *Will*, to determine the acts of this *Will*, and another to determine that, and so *in infinitum*: For where-ever one stops, the Actions of the last *Will* cannot be free.

To understand this passage, let us begin by imagining a man (call him 'Fred') who is free with respect to his acts of volition, and whose actions of willing are made to depend on his will. Now, by Voluntariness, X is free with respect to act Y only if X wills to perform Y. Moreover, if X wills to perform Y, then X performs an act of volition to perform Y.

So consider any volition of Fred's, say 'V<sub>0</sub>'. It follows that Fred performs an act of volition to perform V<sub>0</sub> (call it 'V<sub>1</sub>'). But now, if Y depends on X's will, then any volition to perform Y that X performs is distinct from Y. Given that V<sub>0</sub> depends on Fred's will and that V<sub>1</sub> is an act of volition to perform V<sub>0</sub> that Fred performs, it follows that V<sub>1</sub> is distinct from V<sub>0</sub>. Further, according to Locke, every volition consists in the exercise of a will to which the volition belongs. So there is a will (call it 'W<sub>0</sub>') to which V<sub>0</sub> belongs, and there is a will (call it 'W<sub>1</sub>') to which V<sub>1</sub> belongs. But, in Locke's parlance, *if* X is a volition to perform act Y *and* X is distinct from Y *and* X belongs to will Z,

*then* Z determines Y. So, since V<sub>1</sub> is a volition to perform V<sub>0</sub> *and* V<sub>1</sub> is distinct from V<sub>0</sub> *and* V<sub>1</sub> belongs to W<sub>1</sub>, it follows that W<sub>1</sub> determines V<sub>0</sub>. But according to Heteronomy, no will determines the acts of willing that belong to it. So, since V<sub>0</sub> belongs to W<sub>0</sub>, it follows that W<sub>0</sub> does not determine V<sub>0</sub>. Consequently, by Leibniz's Law, W<sub>1</sub> is distinct from W<sub>0</sub>. It follows that W<sub>1</sub> is an 'antecedent *Will*' that 'determines the acts of' W<sub>0</sub>. Since we are assuming that Fred is free with respect to *all* his acts of volition, repetition of this reasoning shows that there is an 'antecedent *Will*' distinct from W<sub>1</sub> that 'determines the acts of' W<sub>1</sub>, and so on, *ad infinitum*.<sup>10</sup> But an infinite chain of wills, each one determining the acts of its successor, is impossible. Contradiction. It follows by *reductio* that, *if* Fred's acts of volition are made to depend on his will, *then* Fred is not free with respect to his acts of volition.

Notice first that although this *reductio* depends on the Heteronomy Principle, it does *not* depend on the Inheritance Principle. I argued above that Locke did not reject this principle. But even if he had (as Chappell claims), this would not have represented an embarrassment. For Locke does not appeal to this principle in 23C.

Second, as I have reconstructed it, the *reductio* of 23C is not an argument for VD. For VD says that we are not free with respect to our volitions. However, the conclusion of the argument of 23C is a *conditional* claim: it states that a man is not free with respect to his volitions *if* those volitions are made to depend on his will. I argued above that Locke did not believe that DS was incompatible with VD. But even if he had (as Chappell claims), this, too, would not have represented an embarrassment. For 23C does not contain an argument for VD.

<sup>10</sup> Assuming, of course, that the chain of wills does not loop back on itself (see footnote 7).

Third, similar reasoning establishes that the *reductio* of 23C is not an argument for UT. And this is a desirable result. For, as Locke himself saw, DS *does* conflict with UT. The fact that 23C should not be read as an argument for UT removes the last piece of evidence for the claim that Locke retained UT in E2–4. It follows that there is no reason to believe that there is any edition in which these conflicting principles appear together.

The argument of 23C, then, is designed to show that you cannot make a man free with respect to his volitions by making these volitions depend on his will. In particular, you cannot consistently imagine that Fred is free with respect to his volitions if you suppose that his willings to will are distinct from the volitions to which they are directed. More particularly still, you cannot consistently imagine that Fred is free with respect to  $V_0$  (and his other volitions) if you suppose that  $V_1$  and  $V_0$  are distinct; and, conversely, if you suppose that  $V_1$  and  $V_0$  are identical, then you *can* consistently imagine that Fred is free with respect to  $V_0$  (and his other volitions).

But this leaves us with one final problem. For there is a distinct echo of 23C at the end of section 25, an echo that remains throughout E1–4. Now Chappell claims that section 25 contains an explicit statement of VD, and that the echo of 23C (like 23C itself) contains an argument for VD. If Chappell is right, then I am in the unenviable position of having to acknowledge that Locke accepted and argued for VD in E1 and that his commitment to VD never wavered thereafter.

Let us then look at section 25 to see whether it includes a statement and defence of VD. The relevant passage of E5 reads as follows:<sup>11</sup>

<sup>11</sup> With regard to section 25, the differences between E5 and any one of the other editions are insignificant, aside from the addition of the phrase 'in most cases' in the first sentence of the section in E5. This phrase is meant as a further reminder that sections 23–24 contain a defence of  $UT_R$ , rather than a defence of UT (see above).

Since then it is plain, that in most cases a Man is not at liberty, whether he will *Will*, or no; the next thing demanded is, *Whether a Man be at liberty to will which of the two he pleases, Motion or Rest*. This Question carries the absurdity of it so manifestly in itself, that one might thereby sufficiently be convinced, that Liberty concerns not the Will. For to ask, whether a Man be at liberty to will either Motion, or Rest; Speaking, or Silence; which he pleases, is to ask, whether a Man can *will*, what he *wills*; or be pleased with what he is pleased with. A Question which, I think, needs no answer: and they, who can make a Question of it, must suppose one Will to determine the Acts of another, and another to determinate that; and so on *in infinitum*.

The question at issue in this passage is whether we are free with respect to our individual acts of willing, that is, whether VD is true or false. It is plain that, for Locke, the question admits of some 'absurdity', that it 'needs no answer'. What does he mean by this? Chappell claims that,

[Locke] suggests that the absurdity lies not in the question itself, but in an affirmative answer to it ... It seems plain that Locke is urging a substantive, albeit negative answer to the question at issue, and that he is basing it on an argument of the *reductio ad absurdum* form. This negative answer is in fact the thesis of volitional determinism. (108)

This, I believe, is exactly the reverse of the truth. Locke states plainly that the relevant 'absurdity' is contained in the *question*. There is no suggestion in the text that the 'absurdity' is contained in any *answer*, let alone the *affirmative* answer, to it. Moreover, from the fact that a question contains an absurdity, it does not follow that the answer to it must be negative. If anything, Locke clearly indicates that his answer to the question is *positive*. For he says that the question 'whether a Man be at liberty to will either Motion or Rest ... which he pleases' reduces to the question 'whether a Man can *will*, what he *wills*; or be pleased with what he is pleased with'. But the answer to the latter question is: 'Of course!' It is *obvious* that a man can

will what he wills, and it is likewise *obvious* that a man can be pleased with what pleases him, for it is plain that whatever is actual is possible. Thus, far from claiming the truth of VD, Locke is insisting on its falsity.

'But how', you might ask, 'could an absurd question be given an affirmative answer?' The answer is that *what makes a question absurd is precisely that the answer to it is obvious*, in this case *obviously true*, though in other cases perhaps *obviously false*. This usage of the word 'absurd' as applied to the word 'question' may seem odd, but is not all that removed from contemporary linguistic practice. If I ask you, in a serious tone of voice, whether bachelors are really unmarried, you might well reply: 'What an absurd question! Of course they are'.

You might still ask why Locke would have thought that the question 'whether a Man be at liberty to will' reduces to the 'absurd' question 'whether a Man can will, what he wills'. The answer depends on Locke's definition of freedom. Locke says that a man is free with respect to action A if and only if (i) he can do A if he wills to do A and (ii) he can fail to do A if he wills not to do A. I suggest that Locke thought that willing to perform a volition is tantamount to performing that volition. Likewise, or so I conjecture, Locke would have added that willing not to perform a volition is tantamount to not performing it. Assuming that action A is a volition, it follows that a man is free with respect to A iff (i) he can do A if he does A and (ii) he can fail to do A if he fails to do A. Since Locke is happy to 'comprehend the forbearance' of any action 'under the word *Action*' (E5: 28), he can conclude that the question whether one is free with respect to one's volitions reduces to the question whether one can perform the volitions that one performs.

Now Locke acknowledges that his opponents might wish to 'make a Question' of the question to which, as I have been arguing, he provides a positive answer. What Locke means by this

strange turn of phrase is that his opponents might refuse to accept the question's 'absurdity'; that is, they might take the answer to it to be less than obvious. The only way for them to do this is to deny one of the assumptions needed to effect the reduction just discussed, presumably the claim that willing to perform a volition is tantamount to performing that volition. But, according to Locke, those who reject this assumption 'must suppose one Will to determine the Acts of another, and another to determinate that; and so on *in infinitum*'. Locke's support for this is the same as the reasoning of 23C (see above), and I suspect that this redundancy is the main reason why Locke removed 23C from E5.<sup>12</sup>

The upshot of this discussion is that neither section 23C nor section 25 contains a defence of VD. Section 23C contains an infinite regress argument to the conclusion that human beings are not free with respect to their volitions if their acts of willing to perform volitions are distinct from those volitions themselves. Section 25 contains an argument to the conclusion that human beings are free with respect to their volitions, and this precisely because their acts of willing to perform volitions are *not* distinct from those volitions. The two sections are perfectly complementary, coexisting happily until E5, at which point Locke eliminated 23C at least in part because it makes a point already made in section 25. Locke's premisses and his conclusions are all compatible with DS, and are also compatible with UT<sub>R</sub>. So it is not surprising that Locke neither amended nor removed section 25 from the text of the *Essay* after the change of mind that resulted in his acceptance of DS.

<sup>12</sup> In saying this, I do not mean to discount the possibility (raised by Chappell) that Locke removed 23C from E5 in part because he realized that the intended conclusion of the argument of 23C is other than UT (or UT<sub>R</sub>).

## Conclusion

Let me now summarize the main results of this paper. When he first wrote II. xxi of the *Essay*, Locke accepted (and argued for) UT, but did not accept DS. Locke then changed his mind about DS, and, in E2, rewrote the second half of II. xxi accordingly. At the same time, he realized that UT was an overgeneralization and amended section 24 to indicate that he now wished to be read as arguing for a restricted version of UT. More explicit additions and emendations to the effect that his intended conclusion was UT<sub>R</sub> appeared later in sections 23–25 of E5, when Locke realized that the changes he had incorporated into E2 were inadequate. All of this is contrary to Chappell's version of events, according to which Locke never gave up UT, even though he recognized that UT was not compatible with DS. Also mistaken is Chappell's account of Locke's relationship to VD, according to which Locke accepted and argued for VD even after he came to believe (wrongly) that VD and DS were incompatible. For Locke never accepted and never argued for VD. More, since he explicitly rejected VD (in section 25), he was never worried about the possibly objectionable consequences of simultaneously accepting both VD and DS.

What, then, is the philosophical significance of these results in the context of Locke's overall theory of human action? In particular, what is the philosophical importance of Locke's abandoning UT in favour of UT<sub>R</sub> and of his rejection of VD? First and foremost, of course, is Locke's respect for the truth. Having raised questions Q1 and Q2, Locke did not wish to be read as arguing for false answers to them. But it is also important to recognize the place of Locke's considered answers to these questions in the chapter as a whole. Locke begins his discussion of freedom by arguing that it is nonsense to speak of the will's being free, since the will is not an agent, freedom is a

power, and only agents are capable of having powers (14). But Locke insists that although human wills can't be free, humans themselves are often free to act as they will. The purpose of section 23, in this context, is to provide an important qualification: there are circumstances in which humans cannot avoid willing one way or the other with respect to a given action. And this result in turn serves to underscore the importance of the circumstances in which humans *can* avoid willing. For these circumstances are those in which humans have the power to suspend the prosecution of their desires, and it is in this power that Locke finds 'the source of all liberty' and what is improperly called 'free will' (47). Moreover, as Locke emphasizes in section 25, the (qualified) negative result of section 23 does not preclude humans from being free with respect to their individual volitions. Finally, and perhaps most importantly, Locke was able to make a reasonable case for these claims without falling into inconsistency. If Locke deserves our praise, it is mostly for finding a way to expound a consistent and reasonable defence of a viable compatibilist libertarianism.<sup>13</sup>

<sup>13</sup> This paper grew out of ideas I presented in a Spring 1997 graduate seminar on Locke's metaphysics and epistemology at Florida State University. I wish to thank my students in that seminar (Justin Barnard, Brad Hadaway, Greg Smith, and Chip Summey) for helping me to sharpen my thoughts in various ways. In November 1999, I presented a version of the paper at a meeting of the Southeastern Seminar in Early Modern Philosophy at Duke University. I would like to thank all those who attended that seminar for their incisive comments, especially Don Garrett, Sukjae Lee, Paul Lodge, Miriam McCormick, and Tad Schmaltz. I am particularly grateful to my wife and colleague, Dana Nelkin, for her invaluable assistance and encouragement, and to an anonymous referee for some useful suggestions.