

Van Inwagen on introspected freedom

Jean-Baptiste Guillon

© Springer Science+Business Media Dordrecht 2013

Abstract Any philosopher who defends Free Will should have an answer to the epistemological question: “how do we know that we have such a capacity?” A traditional answer to this question is that we have some form of introspective access to our own Free Will. In recent times though, many philosophers have considered any such introspectionist theory as so obviously wrong that it hardly needs discussion, especially when Free Will is understood in libertarian terms. One of the rare objections to appear as an explicit argument was proposed by van Inwagen in his *Essay on Free Will*. In this paper, I address van Inwagen’s anti-introspection argument; I argue that it is both inconsistent with his overall treatment of the Existence Question (namely, with his defence of the existence of Free Will from reflections about morality), and inconclusive in itself (at least for anyone not ready to endorse general scepticism about perception). In passing, I give a clarification of the notion of Introspection, in the case of Freedom, that also sets a more favourable stage for the evaluation of further objections.

Keywords Free Will · Libertarianism · Introspection · Scepticism · Epistemic closure

A Libertarian, in the traditional sense, is someone who accepts not only that Free Will is incompatible with determinism, but also that Free Will exists. There is a large agreement among libertarians as to the proper method to ascertain the first half of this view: incompatibilism is to be argued for on conceptual grounds, by philosophical reasoning. But there is much less agreement on how the second

J.-B. Guillon (✉)
Université de Nantes, 1, avenue de la Tranquillité, 78000 Versailles, France
e-mail: guillon@phare.normalesup.org

half—the “Free Will Thesis”, to use van Inwagen’s terminology—could and should be justified. How could a libertarian claim to *know* that he is free?

An answer that seems to naturally come to mind is that we just *feel* that we’re free. In more philosophical terms, a libertarian could claim that we have some kind of introspective access to, or direct experience of, our own freedom. And some prominent libertarians, like Reid,¹ Campbell² or O’Connor,³ have indeed defended varieties of such a view.⁴

But there are other answers on the market: some libertarians have claimed to know about freedom from reflections about morality (van Inwagen 1983), from the crucial role that this belief plays in our system of beliefs (Donagan 1987), or from some kind of self-referential or transcendental argument (Jordan 1969). Other authors, who could be called libertarians in the weak sense that they endorse a libertarian concept of Free Will, even though they remain agnostic about whether we do have Free Will, consider that the answer could only come from future empirical research (Balaguer 2010; cf. Kane 1996). Most of these libertarian authors join with compatibilists and hard-incompatibilists in pronouncing a sweeping condemnation of any introspectionist strategy for libertarianism. Such pronouncements are typically not very long and not always argumentative in form, because of the wide consensus expected among readers. In this paper, I want to address one explicit anti-introspection argument, proposed by Peter van Inwagen in the *Essay on Free Will*.

¹ “In certain motions of my body and directions of my thought, I know, not only that there must be a cause that has power to produce these effects, but that I am that cause; and I am conscious of what I do in order to the production of them. From the consciousness of our own activity, seems to be derived, not only the clearest, but the only conception we can form of activity, or the exertion of active power” (Reid 1788, p. 36). And of course, Reid’s notion of agent-causation (or “active power”) is an incompatibilist one, since according to him: “power to produce any effect implies power not to produce it” (Reid 1788, p. 35).

² “Can we *trust* the evidence of inner experience? [...] The misguided, and as a rule quite uncritical, belittlement, of the evidence offered by inner experience has, I am convinced, been responsible for more bad argument by the opponents of Free Will than has any other single factor. [...] Are there objections to a freedom of this [libertarian] kind so cogent that we are bound to distrust the evidence of ‘inner experience’? [...] The arguments which seem to carry most weight with Determinists are, to say the least of it, very far from compulsive” (Campbell 1957, pp. 168–174).

³ “It does not seem to me (at least ordinarily) that I am caused to act by the reasons which favor doing so; it seems to be the case, rather, that I produce my decision in view of those reasons, and could have, in an unconditional sense, decided differently. [...] then it is natural for the agency theorist to maintain that they involve the perception of the agent-causal relation” (O’Connor 2000, p. 30).

⁴ Other recent defenders of some variety of introspectionism include Richard Swinburne and Keith Lehrer. For the former, see Swinburne (2011, p. 82): “It is a basic principle of rationality that things are probably the way they seem to be (in the sense that we are inclined to believe that they are) in the absence of counter-evidence. [...] When we make a decision, it seems that we choose and are not caused to choose as we do.” For Keith Lehrer, see (Lehrer 1960, p. 156): “The introspective datum that we deliberate about future actions provides adequate evidence for believing that some of our actions are in our power, that is, that we can choose to perform the action and also that we can choose not to perform the action, and in this sense we have adequate evidence that we have free will.” It is a disputable matter whether Lehrer’s conception is ultimately as introspectionist as he himself thinks it is, but this is not the place to discuss his particular view.

I shall argue that van Inwagen's argument against Introspectionism suffers from two defects. First, it is inconsistent with van Inwagen's own strategy in answering the Existence Question, for *if* the anti-introspection argument works, so should a parallel argument devised against our knowing Free Will from reflections about morality. But van Inwagen wants to maintain that this parallel objection against knowledge from morality *doesn't* work. So the anti-introspection argument shouldn't work either (Sects. 2–4). Second, I shall argue that the argument in itself is unsuccessful, at least for anyone not prepared to swallow general scepticism about perception (Sect. 5).

1 Van Inwagen's anti-introspection argument

Van Inwagen's objection to Introspectionism goes as follows:

Some philosophers believe that it is possible to find out whether we have Free Will by *introspection*. But this seems just obviously wrong, since, if it were right, we could find out by introspection whether we were fitted out with Martian devices like those we imagined in our discussion of the Paradigm Case Argument in Chapter IV. And, of course, we cannot do this. (van Inwagen 1983, p. 204)

The form of the argument is obvious enough:

- (1) If we could know by introspection that we are free, then we could know by introspection (plus a priori reasoning) that we are not fitted out with Martian devices.⁵
- (2) We cannot know by introspection (plus a priori reasoning) that we are not fitted out with Martian devices.

From (1) and (2) follows by modus tollens:

- (C) We cannot know by introspection that we are free.

The acceptability of the premises depends on two crucial features of the Martian Devices here mentioned (hereafter MD). First, MD are used by Martians to manipulate every single one of our choices, and therefore they are logically inconsistent with our having Free Will (even in most compatibilist construals of Free Will).⁶ In other words:

- (a) we are fitted out with MD \Rightarrow we are not free

⁵ I assume that what van Inwagen has in mind with the phrase "find out by introspection" is the possibility to *deduce* that "we are not fitted out with Martian devices" from the (hypothetical) introspective knowledge of "we are free" and the (presumably a priori) knowledge that freedom excludes Martian manipulation. The resulting knowledge, therefore, wouldn't be strictly introspective, or "by introspection" full stop, but rather "by introspection plus a priori reasoning".

⁶ See van Inwagen (1983, p. 109): "whenever that person must make a decision, the device *causes* him to decide one way or the other according to the requirements of a table of instructions."

or conversely:

(a') we are free \Rightarrow we are not fitted out with MD

This feature, together with a principle of epistemic closure, gives support to premise (1). There are of course various principles in the literature which attempt to capture the core of the closure intuition, and not all of them are very cogent. What we need here, though, is only the weak and very plausible version which hinges on the notion of Competent Deduction⁷:

(CD-Closure) If x knows that p , and x is in a position to acquire the belief that q by competently deducing q from p (due to a priori considerations), then x is in a position to thereby come to know that q .⁸

The second crucial feature of MD, is that they generate in us whatever “feeling” or “experience” of Free Will an introspectionist may purport to have.⁹ In other words:

(b) the experience of freedom we would have in the MD scenario is subjectively indiscernible from the experience of freedom we would have if we were indeed free.

And feature (b) is brought in support to premise (2).

2 Analogy between van Inwagen’s argument and the sceptical argument

There is an obvious analogy between this argument and the famous family of sceptical arguments based on the same principle of closure. For comparison, just consider:

(1*) If we could know by perception that we have hands, then we could know by perception (plus a priori reasoning) that we are not brains in vats.

(2*) We cannot know by perception (plus a priori reasoning) that we are not brains in vats.

(C*) We cannot know by perception that we have hands.¹⁰

⁷ It should also be noted that what we need is a principle of closure for *knowledge* rather than closure for *justification* alone, or some other epistemic notion.

⁸ Competent Deduction Closure was famously defended by Hawthorne (2004, 2005). I have chosen to use a version of CD-Closure slightly different from Hawthorne’s exact wording for two main reasons. First, for simplicity, I have omitted some cautious qualifications that don’t play any particular role in our context. Second, in order to fit the “modal” wording of van Inwagen’s argument (“if we *could* know... then we *could* find out...”), I have translated in dispositional terms (“ x is in a position to competently deduce...”) what Hawthorne states in conditional terms (“if x competently deduces...”).

⁹ See van Inwagen (1983, p. 109): “someone might object that [MD] is not in fact consistent with our observations, since we can normally “feel” our decisions “flowing” naturally from our desires and our beliefs; but if [we were fitted out with MD], we should “feel” ourselves being interfered with. But to meet this objection we need only suppose that the Martian device causes us to have desires and beliefs appropriate to the decisions it will cause us to make”.

¹⁰ The detailed formulation of the argument was chosen in order to elicit the analogy with van Inwagen’s argument. One may wonder whether this formulation is really equivalent to more usual formulations, in particular with formulations that do not mention the *modes* of knowledge, and require only that we should

The premises are supported by the following two features of the brains-in-vats scenario:

- (a*) we have hands \Rightarrow we are not brains in vats
- (b*) the experience as of having hands we would have in the brains-in-vats scenario is subjectively indiscernible from the experience as of having hands we would have if we did have hands.

More precisely, premise (2*) is supported by feature (b*). And premise (1*) is supported by feature (a*) together with the Closure Principle for Competent Deduction: anyone who gives (a*) a thought is able to come to know it to be true a priori, and to competently produce the corresponding modus ponens.

This analogy quite straightforwardly suggests a way of defending the Introspectionist stance by the following analogical argument:

Despite its plausibility, the anti-introspection argument here mentioned cannot be right, because it is just analogous to the anti-perception sceptical argument. And of course, the latter must be wrong some way or other. So whatever defeater we choose to rebut scepticism will also defeat the anti-introspection argument.

But analogy isn't identity. And though such a defence by analogy may have some intuitive pull, still there *might* be a defeater of the sceptical argument that wouldn't also be a defeater of the anti-introspection argument.

Nevertheless, I believe the comparison with the sceptical argument can be quite devastating for van Inwagen's discussion, for two reasons I first present in a nutshell.

Footnote 10 continued

have knowledge (of *whatever kind*) that we are not brains in vats (hereafter \sim BIV) (see Pritchard 2005, p. 28; DeRose and Warfield 1999, p. 2). The main worry here is that our formulation may be a weak one if it ignores the possibility of our having some kind of independent or antecedent knowledge of \sim BIV (knowledge in no way perceptually based, presumably a priori). I shall discuss this possibility in Sect. 5. For now, it need only be noticed that this possibility isn't *ignored* but *negated* by one of the premises—namely (1*)—so that it remains a possible way to refuse the argument. Here is how the dialectics should be understood. I take it that a developed account of the closure-based reasoning—like that of Avnur (2012, p. 2)—has three premises rather than two, namely: (4) We don't have antecedent (a priori) knowledge of \sim BIV; (5) we cannot legitimately come to know \sim BIV by deduction from everyday perceptual premises (because this would constitute "Easy Knowledge"); (6) if we had everyday perceptual knowledge, Competent Deduction Closure would allow us to come to know any trivial consequence of such knowledge which we wouldn't antecedently know (in particular \sim BIV, according to (4)); therefore, we don't have everyday perceptual knowledge (or at least can't retain it upon reflection). The most classical two premise versions of the closure-based argument conjoin premises (4) and (5) to assert that we cannot have *any kind* of knowledge of \sim BIV (neither antecedent nor resultant). Our version, rather, conjoins premises (4) and (6) as premise (1*): if we know by perception that we have hands, then given our lack of antecedent knowledge of \sim BIV, Competent Deduction would allow us to acquire easy knowledge of \sim BIV. But, second premise ((premise (5) or (2*)), easy knowledge is illegitimate. This way of presenting the sceptical reasoning obviously lays the stress on Easy Knowledge—and for that reason, it comes closest to Cohen (2002)'s discussions. But it should be noticed for our discussion that this presentation differs from more traditional ones *only in emphasis* or stress: any reply to the traditional formulations of the argument are available to our formulation as well. In particular, the "antecedent knowledge" reply, whatever its plausibility may be, is available by refusing the conditional premise (1*).

First, because van Inwagen uses just such an argument by analogy in order to defend his own conception of “knowledge from morality” (against a parallel objection). But if some analogy has any argumentative value, so should a *closer* analogy. And the introspectionist case is, if anything, *more* analogous to the perception case than is “knowledge from morality”. Therefore, *if* van Inwagen’s analogical defence of “knowledge from morality” works, a fortiori should the analogical defence of introspective knowledge work as well. Or conversely, *if* van Inwagen’s objection against Introspection worked, *so* would the parallel objection against his version of knowledge from morality (Sect. 3, 4).

Second, I shall argue that, if we properly construe the case for introspective knowledge of Free Will, the anti-introspection argument isn’t really an *analogon* of sceptical anti-perception arguments but rather an *instance* of them. Therefore, if one doesn’t want to be committed to scepticism, one has to consider that van Inwagen’s argument against Introspection fails (Sect. 5).

Let us examine those two points more carefully.

3 Problem of consistency for van Inwagen’s general strategy

My first point is that van Inwagen’s *acceptance* of the anti-introspection argument is inconsistent with his own *rebuttal* of a parallel argument against knowledge from morality. Here is the objection he addresses (and rebuts) against knowledge from morality. The objector speaks:

If all your arguments are correct, then our (alleged) knowledge of the existence of moral responsibility, coupled with certain arguments a priori, can constitute a good reason for believing that determinism is false. But these things are not the *sorts* of things that can be a good reason for believing in indeterminism. [...] Therefore, your arguments represent just one more attempt by a philosopher to settle by intellectual intuition and pure reason a question that should be left to empirical science. (van Inwagen 1983, p. 210)

Again, it is not difficult to get at the structure of the argument:

(1**) If we could know from moral responsibility that we are free, then we could know from moral responsibility (plus a priori reasoning) that indeterminism is true.

(2**) We cannot know from moral responsibility (plus a priori reasoning) that indeterminism is true.

(C**) Therefore, we cannot know from moral responsibility that we are free.¹¹

(1**) and (2**), like their counterparts in the anti-introspection argument and the sceptical argument, are backed by two assumptions:

¹¹ Here again, some adaptation (from “reasons to believe” to “knowledge”) was necessary for the analogy with the other arguments to be clearer. My adaptation requires only to assume that “knowledge from reasons” entails “having good reasons to believe”.

- (a**) we are free \Rightarrow indeterminism is true¹²
 (b**) the belief in moral responsibility isn't sensitive to indeterminism
 (only scientific investigations, if anything, could be sensitive to such a state
 of affairs)

Van Inwagen acknowledges that this argument has some force, but his “considered opinion” is that it fails. His reason for rejecting this argument is the *analogy* with the kind of sceptical argument we mentioned a few paragraphs above. It is important to emphasize here that van Inwagen doesn't consider the argument to be an *instance* of such sceptical arguments: he presents the two kinds of arguments as “very much alike”, and sufficiently similar to suggest “two types of response [...] that are exactly *parallel* to the two types of response available to the anti-sceptic” (van Inwagen 1983, p. 212, emphasis mine).

The first type of response is to refuse the closure principle on which (1*) and (1**) rely: maybe we can know that we have hands (by perception) and that “having hands” entails “not being a brain in a vat” (a priori), but that doesn't provide us with a way to come to know that we are not brains in vats. And similarly, maybe we can know that we are free (from morality), and that “being free” entails “being undetermined” (a priori), but that doesn't provide us with a way to come to know that we are undetermined.

The second type of response (the one favoured by van Inwagen) is, so to speak, to make the modus ponens instead of the sceptic's modus tollens: since we *can* know by perception that we have hands, maybe we can, after all, know that we are not brains in vats. And since we *can* know from our belief in morality that we are free, maybe we can, after all, come to know that we are undetermined from that same belief (see van Inwagen 1983, pp. 211–212).

Again, van Inwagen's point is *not* that the two arguments (the sceptical argument on the one hand, and the argument against knowledge from morality on the other), being of the same *kind*, necessarily stand or fall together. His point is rather that he fails to see how the disparities between the two cases could be relevant to justify a different treatment between the two.

I am not quite sure what to say about this defence by analogy. What seems to me quite clear, though, is that the anti-introspection argument presented by van Inwagen is, if anything, *more* analogous to the sceptical argument than is the argument against knowledge from morality. It is more analogous in two respects.

First, the alternative scenarios considered in the anti-introspection argument (Martian devices) and in the sceptical argument (brains in vats) are both philosophical

¹² This incompatibilist assumption may be either accepted wholeheartedly, or assumed for the sake of argument by van Inwagen's objector, according to whether she is a hard incompatibilist or a compatibilist. We could of course bypass the compatibility problem by setting the whole argument in terms of “libertarian freedom” or “being free in the libertarian sense”. But I share van Inwagen's doubts about such phrases (see van Inwagen 2008), and anyway, given the dialectical situation of this paper, it is clearer to keep with van Inwagen's setting of the problem.

inventions lacking any kind of positive support, scenarios just made up for the sake of philosophical argument. Whereas the alternative scenario in the argument against knowledge from morality is arguably a live option for which we might have some kind of inductive support from science. It should therefore be “easier” to justifiably rule out the first kind of alternative than it is to rule out the second kind.

Second, the kind of knowledge considered in the anti-introspection and the sceptical arguments (introspection on the one hand, perception on the other) are arguably *basic* (or non-inferential) kinds of knowledge, whereas knowing that we are free *from* (our belief in) morality is presumably inferential knowledge. And the defence of basic beliefs *as such* has traditionally played an important role in the refutation of scepticism.

In these two respects (which both seem potentially relevant to rebutting the sceptical argument), van Inwagen’s anti-introspection argument is more analogous to the sceptical argument than is the argument against knowledge from morality.

Now, even though the logic of analogical reasoning is rather hard to ascertain, it seems reasonable to accept the following “a fortiori principle”:

- (AF) if it is legitimate to reason by analogy from A is F to B is F because of the analogy between A and B,
 then if C is *more* analogous to A than B is, then it is legitimate to reason by analogy from A is F to C is F because of the (closer) analogy between A and C.¹³

If we apply (AF) to our discussion, the result is that if it is legitimate to reason by analogy from “the sceptical argument must somehow be inconclusive” to “the anti-knowledge-from-morality argument must somehow be inconclusive” because of the analogy between these two arguments, then given that the anti-introspection argument is *more* analogous to the sceptical argument, then a fortiori is it legitimate to reason by analogy from “the sceptical argument must somehow be inconclusive” to “the anti-introspection argument must somehow be inconclusive”. In short, *if* the analogy with the sceptical argument does save knowledge from morality *then* a fortiori does a closer analogy save introspective knowledge. And conversely: *if* the argument against introspective knowledge *does* work in spite of its analogy with the sceptical argument, then a more remote analogy cannot undermine the argument against knowledge from morality.

So if I am right up to that point, there are two conclusions one might draw from our discussion: one way to go would be to accept van Inwagen’s argument against introspection and reject his defence of knowledge from morality. Another way to go would be to reject the anti-introspection argument and say that, for all this argument can show, it remains possible that we should know by introspection that we are free.

¹³ This seems to be a natural specification of Bartha’s very first principle of analogy: “(CS1) The more similarities (between the two domains), the stronger the analogy” (Bartha 2009, p. 19).

4 Objection: the nature of introspection

There is an important objection against the a fortiori line of reasoning presented in the previous section; this objection relies on the peculiar nature of introspection.

Isn't it a *grammatical error*, or a *logical falsehood*, to say that we could know *by introspection* about what Martians are doing? Since introspection is by definition knowledge *of our experiences* or internal states, it is also impossible *by definition* that we should know by introspection about something that is *not* an experience. Now, in the other two arguments, there may be some degree of initial *implausibility* to the fact that we should know that we are not brains in vats by perception, or know that we are undetermined from morality, but this is certainly not *impossible by definition*. So there seems to be an important reason why the analogical argument could save knowledge from morality while it couldn't possibly save introspective knowledge of free will.

The easiest way to understand this objection in formal terms is probably to view it as exploiting the inductive nature of analogical reasoning: it is well known that reasoning by analogy isn't *valid* reasoning, but only reasoning giving *some inductive support* to the conclusion. So suppose we accept the conclusion that the analogy with scepticism gives the introspectionist stance *at least as much inductive defence* than it gives to the knowledge from morality stance. Still, whether this amount of inductive defence is enough to *save* the introspectionist stance depends on the rest of the evidence for and against that stance, in other words, it depends on its prior plausibility. So given the inductive nature of analogical reasoning it might be that the analogical defence of knowledge from morality is weak but successful (given a favourable prior plausibility), while the analogical defence of introspection is stronger but unsuccessful (given a much less favourable prior plausibility). In other words: the introspectionist stance might be, for other reasons, too implausible to be saved.

In our case, the other reason would be provided by a logical analysis of "introspective knowledge": this analysis is seen as giving a *null* prior plausibility to the notion of introspecting freedom (this notion being considered as a "logical falsehood"). But I shall argue that this logical analysis is misconceived. Or more precisely: this line of reply relies on an ambiguity in the notion of "knowledge by introspection", so that the analysis here given is appropriate for some uses of that phrase but not for the uses in the context of "knowledge of Free Will by Introspection". (The following discussion, being an analysis of "knowledge by introspection", will also have upshots much beyond the present section.)

"Knowledge by Introspection" may designate either of three kinds of knowledge in which experience plays a decisive role. Let us take, as an example, the experience as of hearing a roaring lion. From this experience, we may gain:

- (a) knowledge *of* experience: when I have the experience as of hearing a roaring lion, I can form the belief (and usually I thereby come to know) that *I have an experience as of a roaring lion*. The content of the belief (and the knowledge) in that case is not the *content* of the experience but, so to speak, the experience itself (or rather, its occurring).

- (b) knowledge *from* experience: when I have acquired a knowledge *of* experience, of the kind just mentioned, I can also draw some conclusions from this first belief. I might for example draw the conclusion that *some experiences have a rich content*. The content of the belief is now *inferred*... but it is not inferred from the *content* of the experience, it is inferred from the belief *that* this experience occurs.
- (c) knowledge *by* experience: of course, when I have an experience as of a roaring lion, I can also form the belief (and in favourable circumstances come to know) that *a lion is roaring*. That is: I can come to have a belief the content of which is the very same content as that of the experience.¹⁴

The third kind of knowledge (knowledge *by* experience) is what we typically call *perception*, while introspection in the narrow sense (the sense involved in the “logical falsehood reply”) designates the first kind of knowledge (knowledge *of* experience). Now, if anyone were claiming that we can know about Free Will by introspection in this narrow sense, she would be asserting the rather weird view that “being free” requires no more than “feeling free” just because Free Will *is* that experience.¹⁵ Whatever the philosophical defensibility of such an extreme view, it surely cannot be a libertarian conception of Free Will.¹⁶

So, when libertarians like Reid or O’Connor assert that our experience of freedom can give us knowledge that we are free, they certainly don’t understand it as introspection in the narrow sense. It seems obvious, on the contrary, that what they have in mind is a knowledge the content of which is the very content of the experience that grounds it... in other words a kind of perception.¹⁷ Since this kind of “perception” isn’t perception of states of affairs *external* to our body, but rather

¹⁴ Of course, some epistemologists, most notably Descartes, have tried to reduce what I call “knowledge by experience” to a certain form of “knowledge from experience”—for instance, I could be said to acquire knowledge *of* my experience as of a roaring lion, and then *infer* from my having such an experience plus a belief in my creator’s trustworthiness that (probably) there is a roaring lion. Whether or not such a reduction succeeds, my point here is only the logical point that knowledge by experience doesn’t have the same kind of content as knowledge of experience. In the former case—and not in the latter—experience and belief share the same (object-oriented) content. The mere distinction between the two is all that is really needed here.

¹⁵ This kind of view was discussed by Galen Strawson, as an attempt to spell out our common phenomenology of freedom. Strawson interprets this phenomenology as an experience that presents itself as somehow logically sufficient to ascertain freedom: “Consider [...] *a*, an ordinary intelligent self-conscious purposive agent. [...] It is (it seems) quite clear to him that he is able to choose freely. But then surely he *is ipso facto* able to choose freely?” (Strawson 1986, p. 60) The crucial phrase here is “*ipso facto*”. Strawson develops this point with a theory of what he calls “experiential facts”: “What is experienced? Freedom. This really is experience of freedom. It is not just experience *as of* freedom. For freedom is what is experienced” (Strawson 1986, p. 68). This is not the place to discuss such a view: I doubt that there is any wholehearted defender of it; and if there is, he doesn’t need any special argument to accept an introspectionist view on free will anyway.

¹⁶ Nor could it be any kind of traditional compatibilism. Indeed, all traditional conceptions of Free Will are, according to Strawson’s terminology (Strawson 1986, p. 12), “Objectivist theories”—that is: theories in which “*x* believes/experiences himself to be free” doesn’t play any constitutive role in defining “*x* is free”.

¹⁷ Indeed, O’Connor explicitly uses the word “perception” for the kind of knowledge he has in mind. See footnote 3.

perception of some state *internal* to it, it was natural to give it a name with the prefix “intro-” (while the radical “spection” originally means no more than a certain kind of perception). Hence the broad sense of “introspection”: perception of some state internal to one’s body. Because of the narrow sense this word now has, it could be helpful to talk of “interoception”¹⁸ instead of introspection. I shall stick to that terminology for the rest of the paper.

Given the distinction we now have between introspection (knowledge of experience) and interoception (knowledge by experience), what should we think of the aforementioned disanalogy between the anti-introspection argument and the sceptical argument? Is it true that the anti-introspection argument rejects a knowledge claim which is a logical falsehood, while the sceptical argument rejects a knowledge claim which is only *prima facie* surprising? The answer is quite obvious: if the “anti-introspection” argument really is against “introspecting Free Will” in the narrow sense of introspection, then of course the argument is disanalogous with the sceptical argument, and is indeed conclusive. But its conclusion has no bearing whatsoever on the traditional “interoceptionist” stance.

If on the other hand the anti-introspection argument is read with the third meaning of introspection in mind (interoception), then the proposed disanalogy with the sceptical argument disappears: knowing by interoception that we are not fitted out with MD is by no means a *logical falsehood*, or a *grammatical mistake*. For sure, it might be a hard and *prima facie* implausible achievement, but that couldn’t count as a disanalogy with the situation of the sceptical argument, for it is also a hard and *prima facie* implausible achievement to know by perception that we are not brains in vats.

Therefore, the proper analysis of “introspection” in the case of Free Will is unable to render the anti-introspection argument less analogous to the sceptical argument than is the argument against knowledge from morality. To the contrary, it renders it, if anything, more analogous. So our previous conclusion still holds: if interoception cannot be saved by the analogy with scepticism, then a fortiori knowledge from morality cannot be saved by the more remote analogy with scepticism. Or in other words: if van Inwagen’s anti-interoception strategy succeeds, his own defence of knowledge from morality fails. Therefore, van Inwagen’s overall strategy in treating the Existence Question fails. This conclusion is not just an *ad hominem* argument since many philosophers of Free Will seem to have agreed with van Inwagen on the general idea that a libertarian epistemology based on morality was somehow better off than any epistemology having to do with our experience of Free Will.

5 The anti-interoception argument as an instance of the sceptical argument-type

The clarification of the notion of interoception has been brought to bear, so far, on the comparison between knowledge from morality and interoceptive knowledge of

¹⁸ The word was first introduced by Sherrington (1906) in a technical physiological sense, and this specialized meaning has been refined a few times since Sherrington. My use of the word is of course quite independent of the physiological terminology.

Free Will. And I have argued that this clarification shows the anti-interoception argument to be so close to the sceptical argument that it certainly is closer than the anti-knowledge-from-morality argument.

Now, I think we can push the point further, and bring it to bear on the more substantial question whether the anti-interoception argument, in and of itself, succeeds or fails. Indeed, the analysis of interoception as a particular kind of perception reveals much more than a close “analogy” between the anti-interoception argument and the closure-based sceptical argument—or rather should we say, the closure-based sceptical “argument-type”,¹⁹ for closure-based arguments can be multiplied ad infinitum, changing only the particular sceptical scenario: demons, brains in vats, martians etc., or the particular source of knowledge: sight, touch, memory, etc. What our clarification shows is that the anti-interoception argument is really an *instance* of that type of sceptical arguments. And if this is true, then refuting (or refusing) scepticism (the sceptical type of argument) directly *implies* refuting (or refusing) the anti-interoception argument, without the weaknesses or uncertainties involved in analogical reasoning.

We shall see that the anti-interoception argument is just an instance of the closure-based sceptical argument-type if we focus on the common form of closure-based arguments, whether they involve demons, brains in vats, or indeed Martians, and whether they involve sight, touch, or memory. The essential apparatus of the closure-based sceptical argument is the following:

- (i) x has an experience E as of p , to the effect that x could prima facie be claimed to know that p on the ground of experience E .
- (ii) some sceptical hypothesis SH just made up for philosophical purposes (and that has no positive support) is such that (necessarily) if that scenario were the case, then
 - (a) p would be false (and x is able to apprehend a priori that $SH \Rightarrow \sim p$) and
 - (b) x would have just the same experience E .

With those two features, the sceptical argument proceeds as follows:

- (1°) if x could know by E that p , then x could know by E (plus a priori reasoning) that SH isn't the case.
- (2°) x couldn't know by E (plus a priori reasoning) that SH isn't the case.
- (C°) x couldn't know by E that p .²⁰

¹⁹ Or, to follow Pritchard (2005), the “template closure-based argument”.

²⁰ As discussed in footnote 10, the exact presentation of the sceptical argument type isn't the most classical one. This is due especially to premise (1°) which conjoins the Closure intuition on the one hand, and the “no antecedent knowledge of $\sim SH$ ” intuition on the other hand. Correspondingly, there are two different ways of refusing (1°), by denying Closure, or by accepting antecedent knowledge of $\sim SH$. I discuss the possibility of antecedent knowledge—for both the sceptical argument and van Inwagen's argument—in the next paragraph. It may be important to remember here that this different grouping of the premises is only a matter of presentation, as long as the set of possible replies to the argument remains unchanged.

The brains-in-vats argument we presented earlier takes any of us as “ x ”, the whole set of our perceptual experiences as “ E ”, “we have hands” as “ p ” and “we are brains in vats” as “ SH ”.²¹

But obviously van Inwagen’s anti-interoception argument is also of that form: we just have to substitute my experience of freedom for “ E ”, “I am free in performing action a ” for p , and the Martian device scenario for “ SH ”.

Maybe I have been too liberal in my formulation of the common form of closure-based arguments? Maybe I shouldn’t have talked of “experience E ” in general, but only of “perceptual experiences” or “external perceptual experiences”? Well, if that were the case, we would have to say that closure-based arguments against memory rely on an essentially different intuition or argument form than the brains-in-vats arguments. But obviously they have just the same form, and this form is the one we just expressed. Just consider our argument type with the following parameters: let “ x ” be myself, let “ E ” be the set of my memorial experiences, let “ p ” be the proposition that I had coffee for breakfast this morning, and let “ SH ” be the hypothesis that a very powerful demon created me one second ago with just the set of memorial experiences which I happen to have. What we get is obviously a new instance of just the same sceptical argument. What is essential to this kind of closure-based sceptical arguments is that they threaten what I’ve called “knowledge *by* experience” (knowledge the content of which is the very content of the experience that grounds it).²² And we’ve seen that interoception was just an instance of that kind of knowledge.

I now come to a problem already alluded to in footnote.²³ In the formulation here given, the sceptical argument may seem to rely on the problematic contention that not only knowledge in general, but also particular knowledge modes are closed under (known) entailment. That is, it may seem to imply that whenever I know something *from source E*, then I know (or can come to know) the consequences *from source E as well*. This would be unfortunate, since the non-closure of knowledge modes is widely accepted.²⁴ So this might be sufficient to discredit both the argument type just mentioned, and van Inwagen’s instance of that type. [Remember here that van Inwagen explicitly says that “if it were right (that we can find out whether we have Free Will by introspection), we could find out *by introspection* (sic) whether we were fitted out with Martian devices” (van Inwagen 1983, p. 204), emphasis mine.] But it seems to me that this problem with van Inwagen’s formulation is at the end of the day no more than a minor distraction.

A first thing to notice here is the difference between “knowing from source E ” and “knowing from source E plus a priori reasoning”. If anyone said that the former

²¹ More precisely, in order to satisfy the conditional (ii)-(b), SH must read: “we are brains in vats manipulated to have exactly the perceptual experiences E ”.

²² Namely, closure-based arguments which involve sceptical scenarios. This not the case for all closure-based arguments; for instance, “lottery arguments” do not involve sceptical scenarios (see Hawthorne 2004).

²³ See footnotes 10 and 20.

²⁴ Indeed, it is sufficiently uncontroversial for Dretske to take it as a premise in an argument against general closure. See Luper (2005) on that point.

is closed under known entailment, this would indeed amount to a closure of the knowledge mode corresponding to source E. But saying that the latter is closed is really a claim about competent a priori deduction, and not about the particular knowledge mode corresponding to source E. Since van Inwagen obviously had in mind the deductive step from “we are free” to “we are not fitted out with MD”, it is clear that he never meant to assume anything like closure of knowledge modes, but only closure of Competent Deduction.

Now, a second remark is needed because it is still true that the most plausible versions of Competent Deduction Closure (including principle (CD-Closure) above) are unable to deliver *directly* premise (1°) and its instances. In order to get (1°) directly, a stronger principle would be needed:

(Strong CD-Closure) If x knows that p , and x is in a position to competently deduce q from p (due to a priori considerations), then x is in a position to know that q in this way.

Unfortunately, (Strong CD-Closure) doesn't seem to be true in general. There are cases in which I know a proposition p , and I can entertain the deduction from p to some entailed q , but I already know q and couldn't have come to know that p in the first place if I hadn't antecedently known q . In such cases, the deduction from p to q couldn't constitute a new way for me to know that q .²⁵

Because of such cases (of essentially antecedent knowledge of the consequent), the most defensible versions of Competent Deduction Closure provide us with deductive knowledge *only* for those consequences (of our known beliefs) which we don't antecedently know.²⁶ This qualification is of great importance for sceptical arguments because one possible response to scepticism is precisely to appeal to some antecedent (non perceptual, presumably a priori) knowledge of anti-sceptical propositions. If we do have such antecedent knowledge of \sim SH, then (CD-Closure) is trivially satisfied (by falsity of the antecedent), and doesn't deliver the result that everyday perceptual knowledge would set us in a position to acquire further deductive warrant of \sim SH. In other words, it doesn't support premise (1°):

(1°) if x could know by E that p , then x could know by E (plus a priori reasoning) that SH isn't the case.

²⁵ Suppose a reliable friend tells me that Mr John X is truthful and highly reliable regarding his family (the Xs), and I come to know this proposition (q) upon my friend's testimony. Suppose, furthermore, I hear Mr John X saying that *all* members of his family are truthful and highly reliable regarding the Xs, and I come to know this new proposition (p) upon Mr John X's testimony, whom I know to be trustworthy because of my knowledge that q . Now, it is obvious that p entails q and I could easily bring my attention to the deduction of q from p . But it would be really odd to say that I would thereby acquire a new way of knowing that q —a way relying ultimately on my knowledge that p (plus a priori reasoning)—for I wouldn't know that p in the first place if I didn't *antecedently* know that q .

²⁶ In the principle (CD-Closure) of this paper, this qualification is expressed by the condition that the knower x must be “in a position to *acquire* the belief that q ”. If x has antecedent knowledge of q , he is not in a position to acquire new belief (let alone new knowledge) that q , and therefore the principle is trivially satisfied for x doesn't satisfy the antecedent of the conditional.

Therefore, if we have antecedent knowledge of \sim SH, premise (1°) may be false without falsifying Competent Deduction Closure. This shows that premise (1°) really bears in itself two different assumptions essential to the sceptical reasoning: one is the epistemic closure intuition, and the other one is the impossibility of having antecedent knowledge of \sim SH. For that reason, there are two different ways to resist premise (1°)... and therefore three ways to resist the closure-based argument type:

- (a) resist premise (2°), i.e. accept that deduction from everyday perceptions is an acceptable way to come to know the negation of sceptical scenarios.
- (b) resist premise (1°) by accepting antecedent knowledge of anti-sceptical propositions.
- (c) resist premise (1°) by denying closure. This means accepting that we can know everyday perceptual propositions even though there is no way we can know some of its obvious logical consequences.

What solution is the most plausible? It is not the purpose of this paper to answer this question. What is important for us is only this: whatever solution is adopted for the sceptical argument-type, an instance of that solution will immediately follow to rebut van Inwagen's anti-interoception argument. If easy deductions like "this is a hand, therefore I am not a brain in a vat" are epistemically legitimate, then so are deductions like "I am free, therefore I am not fitted out with MD". If we have some antecedent, a priori, way to rule out brains-in-vats scenarios and demons scenarios, then surely we can also rule out Martian device scenarios. If Competent Deduction Closure fails, then there is no particular problem in "intercepting" Free Will even though we can't know that we are not fitted out with MD.

Though I (like van Inwagen) favour the answers which preserve epistemic closure, my general conclusion is only that anyone who rejects the sceptical argument-type (for whatever reason) should also reject (for just the same reason) van Inwagen's anti-introspection argument because it is just an instance of that type.

6 Conclusions and upshots

We have now reached the two intended conclusions. First, van Inwagen's anti-interoception argument is inconsistent with his overall treatment of the Existence Question, and more precisely with his defence of the knowledge-from-morality stance. For all his argument can show, the interoceptionist stance is no worse off than the knowledge-from-morality stance (it is rather, if any different, better off!). Some readers may want to read this conclusion in its negative form, namely that the knowledge-from-morality stance is "yet worse off" than the (already very badly off) interoceptionist stance. In any case, this goes against an assumption common to various philosophers of Free Will that *if* there is any hope for a libertarian stance, it could only be found in indirect strategies like the knowledge-from-morality stance.

Second, I argued that van Inwagen's argument is inconclusive for anyone not prepared to endorse general scepticism about perception. So not only does his argument fail to place the interoceptionist stance in a worse epistemic position than

the knowledge-from-morality stance; it doesn't even place interoception in a worse epistemic position than everyday perceptual knowledge.

Some readers may find that I haven't managed to establish those conclusions. But in this last section, I want to address another kind of reaction, which we could express as follows.

Your discussion does succeed in showing that van Inwagen's argument, in its unfortunate historical form (with the sceptical MD scenario), doesn't pose a real threat to the libertarian interoceptionist stance. But a charitable reader should go beyond the arguments as explicitly stated and address their best reconstructions. And there *is* indeed a minimal revision of van Inwagen's argument that remains unscathed by anything you've said. Namely:

(1r) If we could know by interoception that we are free, then we could know by interoception (plus a priori reasoning) that indeterminism is true.

(2r) We cannot know by interoception (plus a priori reasoning) that indeterminism is true.

(C) We cannot know by interoception that we are free.

Since *this* argument is obviously successful against libertarian interoceptionism, all you have proved at the end of the day is that van Inwagen was a bit careless in knocking down a position which almost nobody wants to hold anyway.

A first thing to notice is that this reaction changes nothing to the first conclusion we drew, i.e. the comparative conclusion that interoceptive knowledge isn't worse off than knowledge-from-morality. Indeed the "revised argument" here proposed renders the two arguments (anti-interoception and anti-knowledge-from-morality) yet closer, and therefore yet more likely to stand or fall together in an analogical reasoning with the sceptical argument. If this were van Inwagen's original intuition, it would devastate yet more clearly his own knowledge-from-morality stance.

For the same reason, this revised argument cannot be plausibly presented as a "charitable reconstruction" of what van Inwagen had in mind. Notice that what gives the revised argument a special force is that the threatening alternative in it isn't a pure invention made up for philosophical purposes and lacking any kind of positive support; it isn't a good old "sceptical scenario" like the MD scenario. And therefore, it is not clear that this argument can be considered as just one more instance of the sceptical argument type.²⁷ But, van Inwagen shouldn't be happy to concede that the parameter of disanalogy with sceptical arguments which renders interoceptive knowledge unrescuable is precisely the presence of a non-farfetched alternative, for this very same parameter would yet more obviously undermine his

²⁷ For sure, it keeps a strong *analogy* with sceptical arguments—this analogy is indeed what van Inwagen uses to defend knowledge-from-morality from the same threat of Determinism. But again, mere analogy may not be enough, and it may very well be that some plausible reply to the sceptical argument type is impossible or implausible for this argument.

knowledge-from-morality stance.²⁸ For that reason, we should treat separately the anti-interoception argument with sceptical alternative (which is van Inwagen's) and the anti-interoception argument with Determinism alternative. Those are two completely different kinds of alternatives and generate completely different kinds of replies.

So, for the purposes of this paper, my two main reactions to the "revised argument" are that 1) it is irrelevant to the *comparative* conclusion that introspection is no worse off than knowledge-from-morality or other indirect strategies, and 2) though it may constitute a more problematic threat to libertarianism than van Inwagen's argument, it cannot be considered a mere revision of van Inwagen's intuition, and therefore, it is also irrelevant to the positive conclusion that van Inwagen's argument raises no worry for the libertarian introspectionist stance.

Now, of course, if some other argument than van Inwagen's obviously defeated introspectionism, there would be little interest in responding to van Inwagen's argument in the first place. So, even though the proper treatment of the revised argument requires a separate treatment, it may help to give here at least a first indication that the task of responding to it isn't totally hopeless. Fortunately, I think that the preceding discussion gives us some resources to give such an indication. We can see a first insight of a possible solution if we pay attention to the fundamental difference between van Inwagen's argument and the revised argument. As we've seen, the fundamental difference lies in the nature of scenarios. The most important aspect of sceptical alternatives is that our experiences are *in principle* unable to reliably track their being the case. Fortunately, the correlative weakness is that alternatives with such a feature seldom come up as live options, and sceptics have to *make them up*.²⁹ So the "strength" of those alternatives is that, very intuitively, we

²⁸ There is a subtler and more important reason to think that the choice of the MD alternative (as opposed to a Determinism alternative), far from being an effect of carelessness, has an essential role in van Inwagen's strategy. Here is why. When the MD scenario first appears in van Inwagen's book (i.e. in the discussion of the Paradigm Case Argument (see van Inwagen 1983, chap. 4.2), it isn't chosen first and foremost as a sceptical scenario, but rather as a scenario that precludes free will under any defensible construal of the term, *be it compatibilist or incompatibilist*. This feature makes it a good test case to eliminate theories or arguments that are unacceptable even by compatibilist lights. It seems to me that the same kind of dialectical intention is at play here: van Inwagen probably intended to dismiss introspection altogether, for compatibilists and libertarians alike. For if he succeeds in doing so, then the compatibilist is bound to follow him in accepting an indirect epistemology of free will, like knowledge-from-morality. Now, on such common ground, it is not clear at all that the libertarian has a harder epistemological time than the compatibilist, because indirect epistemologies are admittedly fairly mysterious for anyone. But if the compatibilist is still in a position to claim an introspective epistemology *for his* kind of free will, then maybe he will be in a less mysterious position (after all, could we not know by introspection, for instance, that our first order desires fit our second order desires, or that they fit our deep self?) Therefore, dismissing introspectionism *for everyone* sets van Inwagen in a far better position in the epistemological discussion with compatibilists. And only the MD scenario allows him to do that.

²⁹ See, for instance, the way in which Pritchard defines sceptical alternatives: "Since it is clearly the case that we are able to rule out some error-possibilities, in order to meet this challenge the sceptic therefore needs to further argue that there are error-possibilities within the class of relevant error-possibilities that are ineliminable—that is, error-possibilities which agents are unable to rule out. To this end, the sceptic typically adduces a special kind of error-possibility—viz. sceptical hypotheses. Crucially, this type of error-possibility has the special feature that, intuitively, no agent could ever rule it out" (Pritchard 2005, p. 24).

cannot properly rule them out... but their weakness is that, being totally made up scenarios, lacking any kind of positive support, it is not so clear that they *have to* be ruled out, at least in any strong sense of the term. The Determinism scenario has exactly the reverse strength... and the reverse weakness! It is a live option with (at least some minimal) positive support, and this gives strength to the idea that a libertarian would have to rule out this alternative in some way or other... but it also has the reverse defect of not being “ineliminable in principle”, or non-trackable by mere stipulation, because the nature of the “scenario” here is settled by our evidence from the external world, not stipulated for our dialectical needs as is the case for sceptical scenarios. For that very reason, the question whether our experiences could or couldn’t track Determinism has to be *argued for*, not stipulated. This, I think, suggests a general strategy for how to tackle the new argument: one can put into doubt the assumption that our interoceptive experiences couldn’t reliably track determinism. If we cannot exclude the possibility of such a tracking relation, why should (2r) be considered obviously true? There seems to be at least some room for discussion here.

Of course, I do not take this hint, as it stands, to be a full answer to the new argument. A full answer would need a paper of its own. My intention is only to suggest that rebutting van Inwagen’s sceptical argument against interoception isn’t just a local victory in a hopeless war; it is rather a first battle to clear the way for others.

To sum up the main upshots of this paper: I have tried to convey the general idea that there is no good reason for libertarians to despair of the traditional interoceptive epistemology and switch to an indirect one (like knowledge from morality); first because such strategies are not better off than interoception; and second, because it is far from clear that interoception faces decisive challenges. One of the rare explicit arguments actually proposed by an influential philosopher—that of van Inwagen—is actually a failure, and discussion of this argument also gives a hint of how to approach further issues.

Acknowledgments I would like to thank Peter van Inwagen for the stimulating discussion which gave birth to this paper. It happened during the “Metaphysics of Persons” Conference in Saint-Jacut which was funded by the Société Française de Philosophie Analytique (SOPHA). I am also very thankful to Cyrille Michon, Emmanuel Baierlé, Andrew Bailey, Ide Lévi and Ghislain Le Gousse for helpful advice.

References

- Avnur, Y. (2012, April). Closure reconsidered, 12(9). <http://hdl.handle.net/2027/spo.3521354.0012.009>.
- Balaguer, M. (2010). *Free will as an open scientific problem*. Cambridge: MIT Press.
- Bartha, P. (2009). *By parallel reasoning: The construction and evaluation of analogical arguments*. Oxford: Oxford University Press.
- Campbell, C. (1957). *On selfhood and godhood: The Gifford lectures delivered at the University of St. Andrews During Sessions 1953–54 and 1954–55* (Rev. and Expanded). London: Allen & Unwin.
- Cohen, S. (2002). Basic knowledge and the problem of easy knowledge. *Philosophy and Phenomenological Research*, 65(2), 309–329.
- DeRose, K., & Warfield, T. A. (Eds.). (1999). *Skepticism: A contemporary reader*. New York: Oxford University Press.
- Donagan, A. (1987). *Choice: The essential element in human action*. London: Routledge & Kegan Paul.

- Hawthorne, J. (2004). *Knowledge and lotteries*. Oxford: Oxford University Press.
- Hawthorne, J. (2005). The case for closure. In M. Steup & E. Sosa (Eds.), *Contemporary readings in epistemology* (pp. 26–43). Malden, MA: Blackwell.
- Jordan, J. N. (1969). Determinism's dilemma. *The Review of Metaphysics*, 23(1), 48–66.
- Kane, R. (1996). *The significance of free will*. New York: Oxford University Press.
- Lehrer, K. (1960). Can we know that we have free will by introspection? *The Journal of Philosophy*, 57(5), 145–157.
- Luper, S. (2005). The epistemic closure principle. In *Stanford encyclopedia of philosophy*.
- O'Connor, T. (2000). *Persons and causes the metaphysics of free will*. New York: Oxford University Press.
- Pritchard, D. (2005). *Epistemic luck*. New York: Oxford University Press.
- Reid, T. (1788). *Essays on the active powers of man by Thomas Reid*. Edinburgh: Printed for John Bell and G.G.J. & J. Robinson London.
- Sherrington, C. S. (1906). *The Integrative action of the nervous system*. New York: C. Scribner's Sons.
- Strawson, G. (1986). *Freedom and belief*. Oxford/New York: Clarendon Press/Oxford University Press.
- Swinburne, R. (2011). Dualism and the determination of action. In R. Swinburne (Ed.), *Free will and modern science* (pp. 63–83). Oxford: Oxford University Press.
- van Inwagen, P. (1983). *An essay on free will*. Oxford: Clarendon Press.
- van Inwagen, P. (2008). How to think about the problem of free will. *Journal of Ethics*, 12(3/4), 327–341.