Fichte’s Theory of Free Will

Rory Lawrence Phillips
UCL

Submitted for the degree of Doctor of Philosophy in Philosophy

2020

Declaration: I, Rory Phillips, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract:

Fichte's account of free will is at the heart of his philosophy. However, there exists no full-length attempt to come to an understanding of what Fichte's account is, and what it implies. I will therefore present Fichte's views on freedom as they appear in his major work, the 1798 *System of Ethics*, supported by a variety of other important texts. I will give an exegesis of Fichte's arguments, trying primarily to display the arguments clearly, but also offering some reconstructive remarks and defences of Fichte's views against various challenges. I then move on to think about issues connected with free will, mainly the issues of how evil actions are possible, the connections between free will and time, and the causation at play in free will. I conclude that Fichte has a coherent account of strong libertarian freedom, that has significant relation to his views not only on self-consciousness but also on the realm of nature, as well as God.

Impact Statement:

The primary impact of my research is inside academia, in the field of history of philosophy. We are in the midst of something of a "Fichte renaissance" and more gets written on Fichtean philosophy every year. However, there is much still to be done. Part of what needs to be done is sustained careful exegesis and presentation of Fichte's own arguments, especially those arguments for free will, which is a central concept for Fichte. It is this that my thesis will contribute to the current debates, in both Fichte's practical and theoretical philosophy. This has import not only for the philosophical public of the UK but also that of the USA and Germany (where scholarly societies to discuss Fichte's work are already established). It is also important to demonstrate to the wider philosophical audience of Fichte's merit where contemporary discussions of free will and related concepts are happening, as Fichte's oeuvre is a largely untapped resource where contemporary discussion is concerned. Bringing about these benefits will happen at both the level of teaching and at the level of research. My thesis has informed my teaching of Kant and German idealism, and the way I have structured the modules on those philosophers. I have begun to disseminate my research in papers, presented at numerous conferences, to both philosophers who are trained and well-read in German idealism and to those who are new to it. By talking to colleagues who are well-versed

in Fichte and German philosophy more broadly and to those who are interested more in contemporary debates about morality and freedom I can help to show Fichte's enduring relevance to the wider philosophical community. Impact outside academia could be brought about by popular-level articles for a number of journals or magazines. Showing the relevance and importance of Fichte's thought for wider cultural discussions of free will is therefore a possible benefit of this research.

Acknowledgements

There are many people to whom I owe substantial debts. Many of these people are members of the UCL graduate community or have been affiliated with UCL in some way. The person to whom I owe the most thanks in this department is my supervisor, Sebastian Gardner, who accepted me as an MPhil student when I began at UCL in 2014, and has been my supervisor since then, including his recent research leave. My views on Fichte have been shaped by his wealth of knowledge regarding Kant and the German Idealists, and our discussions regarding both issues of textual exegesis and philosophical significance were always stimulating and constructive. Many other members of UCL staff have at one point or another supported me along the road to submission, including Doug Lavin, Fiona Leigh, Lucy O'Brien, John Hyman, Mark Kalderon, Mike Martin, Rory Madden, Mat Simpson, Rob Simpson, and Tom Stern. Tom deserves a special mention for very kindly agreeing to help me with reading German. My time at Yale was a very constructive three months, for which I thank the UCL Doctoral School for their generous funding, and Yale University for sponsoring my visa. At Yale, I had regular meetings with Paul Franks, with whom I spent much of the time discussing Fichte's philosophy of religion. Thanks are also due to Stephen Darwall, John Hare, Michael Della Rocca, and Kenneth Winkler, for their supervision and feedback on a presentation I gave at Yale's early modern philosophy forum, organised by Matt Leisinger, to whom thanks are owed for stretching the boundaries of what counts as early modern for the sake of accommodating me (and Fichte). As well as staff, many of the former and current members of the graduate community at UCL and Yale, where I spent a term, have been and continue to be excellent philosophical colleagues as well as great friends. I thank particularly Showkat Ali, James Brown, Michael Burton, Claudia Cristelli, Niels Christensen, Nick Currie, Catherine Dale, Taylor Enoch, Daniel Ferguson, Lisa Grant, Stephen Harrop, Hugo Havranek, Kirstine La Cour, James Laing, Charles Jansen, Michael Markunas, Rowan Mellor, David Olbrich, Vittorio Sandri, David Sommer, Bex Rowson, Shunichi Takagi, Nikhil Venkatesh, Victor Weber. Some of this material has been presented at conferences. An ancestor of the sections on Fichte's philosophy of religion was presented at a conference on Fichte hosted by the VUB in Brussels. I thank Karin De Boer and David W. Wood for organisation and feedback. Some of the material

on Fichte's account of the natural world was formed as part of the conference entitled "Transcendental Philosophy and Naturalism", held at the University of Parma in May 2018. I thank Andrea Staiti, Sebastian Luft, and Konstantin Pollock for organising, and the participants of the conference for their feedback. Thanks are also due to Daniel Breazeale and Allen Wood, both of whom took the time to respond to my questions regarding Fichte and their interpretations.

The largest debt, however, is to my partner, Hannah Simpson, without whom not.

Table of Contents

0) Introduction

In order to understand Fichte's theory of free will, we have to understand his account of action in general. This means a foray into his theoretical philosophy, as action in general for Fichte characterises everything, but most importantly the mind. The mind is said to be a kind of agility, it is itself action, not a thing. We should not even say that the I is an active thing, but mere action. Roughly, Fichte thinks that what we know most certainly and fundamentally is that *I am*. But what does it mean for *I am* to be true? It means that I posit myself, or that I am aware of myself. The equivalence of "I am" and "I posit myself" means that for an I, its being and its activity, or its essence and existence, are the same. Therefore to speak of the I as an active being is to misconstrue its nature – it is an action.

Fichte's aim in theoretical philosophy is to provide what he calls a 'genetic' account of this. That is, he wants to give deductions which show how the I can proceed from the barest abstraction – "I am" - to an existent, spatio-temporal, living human being. Freedom makes up no small part in this sequence of deductions. Thus I intend to begin first with exegesis and explanation of what freedom in this context, the broadest of all, means for Fichte, before focussing on more specific issues in the realm of action. Naturally, the outcome at this more abstract stage will influence what comes later.

Throughout this thesis, I shall not be giving an overall account of Fichte's philosophical system, but only an account of his theory of free will. Given that Fichte is a systematic philosopher, each part of the system will of course be connected to every other, but it is also true that Fichte's theory of free will, or of self-consciousness, or of moral action, can be discussed and understood on its own merits. Therefore I will talk about other aspects of Fichte's system where appropriate to the discussion, without giving a comprehensive overview of the whole. This reflects something I have long thought to be true of Fichte. That is, just as Kant famously remarked that it was a 'scandal of philosophy' that the existence of the external world might remain in doubt, so Fichte would say that the real scandal was that free will is in doubt.

The question of whether Fichte believes in free will is a somewhat strange one. He affirms it in everything he wrote, frequently using superlatives such as 'completely', 'absolutely' and so on to describe this freedom. But without a proper understanding of the role that freedom plays in his theoretical philosophy we shall be at a loss to see how free action or free will itself works on the Fichtean model. I shall mainly be referring to the so-called 'early' Fichte – that is, the texts composed and published in between the Zurich period (1790-3) and the end of the Jena period (1799-1800). This period comprises the only version of the *Wissenschaftslehre* published in Fichte's lifetime, the *Grundlage des Gesamten Wissenschaftslehre* (trans:

Foundations of the Entire Science of Knowledge – abbreviated as SK) as well as *Der System der Sitten* (trans. *System of Ethics* - SE), *Der Grundlage des Naturrechts* (trans: *Foundations of Natural Right* - FNR), the 'later Jena' system, the *Wissenschaftslehre Nova Methodo* (trans: *Foundations of Transcendental Philosophy* WLnm), the important essay *On the Basis of our Belief in a Divine Governance of the World Order* (DG) and, what is perhaps his best known work, *Der Bestimmung des Menschen* (trans: *The Vocation of Man* - VM). I shall stick to the 'early' Fichte for the reasons that the Jena texts are generally thought to compose a whole, which is not true of the myriad later versions of the *Wissenschaftslehre*, and because Fichte's philosophical mindset appears to undergo some radical shifts in the later period, which would take me too far afield from the themes I want to discuss here.

This thesis is composed of six chapters. In the first, I look at Kant's views as they pertain to freedom across a number of different works. The objective of this section is to provide the reader with an idea of what salient points are picked up from Kant by Fichte, and what features of the Kantian account Fichte will go on to jettison. In the second, I outline Fichte's important and disputed thesis of the primacy of the practical, particularly with respect to practical reason. In doing so, I attempt to give an in-depth exegesis of the so-called "striving argument" from the 1794/5 *WL*. In the third chapter, I try to show that Fichte's most important work on freedom from the early period can be found in the first and second "Deductions" of the *System of Ethics*. This is done again by close textual exegesis of what I take to be the pivotal argumentative strands of those deductions. By the end of chapter three, then, I will have shown that 1) Fichte thinks Kant's views on freedom require amending, 2) that he argues for a strong notion of the primacy of practical reason, which orients the reader and informs his philosophical outlook, and 3) that he believes in strong libertarian freedom, which is informed by his philosophy of religion and his views on the natural world. In the second half of the thesis, I look at three interconnected issues arising out of this discussion. Chapter four contains a discussion of Fichte on "agent causation" and whether he can be said to hold an agent-causalist view (which many in contemporary philosophy think necessarily goes with libertarian freedom). I conclude that it is possible to think that Fichte is an agent-causalist in only one sense, though not in others. In chapter five, I think about Fichte's theory of evil, arguing that it is a coherent and plausible account of the source of evil, by contrasting Fichte's views with Schelling's, and thinking about how Fichte's view can account for the various kinds of evil one might find. In the sixth chapter, I consider Fichte's views on truth, as it pertains to free will. Because Fichte did not say much about truth as such, this section is more reconstructive than previous sections. I argue that Fichte is committed to "temporalism" about truth – that is, that the future is open and not fixed, subject to change at the hands of finite rational agents.

We should start by looking at the Kantian view, to see why this is unsatisfactory for Fichte. This will culminate in looking at Kant's view that freedom is a fact of reason, and to see how this acts as a spur to Fichte to develop his genetic account.

1) Kant's Theory of Freedom and Fichte's Response

The main texts I shall look at for Kant will be the Third Antinomy in the first *Critique*, and section III of the *Groundwork*. I take it that Kant's aims in these texts are, respectively, to argue that freedom and determinism, two types of causation, compossible, that is, there is a possible world which is both determined and contains free agents, and that, secondly, we must (normatively, not merely psychologically) take ourselves to be free in some sense. I shall end with some discussion of Kant's developed view in the second *Critique*, as well as looking at the *Religion Within the Boundaries of Mere Reason*, as this provides a more immediate setting for Fichte.

1.1) Third Antinomy

The third antinomy is Kant's presentation of a thesis and antithesis and an argument for each. The problem, of course, is that both arguments, according to Kant, seem sound, and so reason is stuck in a conflict. I shall present what I take to be the substance of the arguments now, and to see why Kant thinks each of them are sound. The thesis, that there must be two types of causality – natural and free – is argued for in the following way, via a *reductio*.

1) Assume that the only type of causality is natural, that is, everything which happens presupposes a preceding state from which it follows according to a law.
2) This preceding state must have the same conditions apply to it, and this follows *ad infinitum*.
3) But then there will never be a first beginning, and no completeness of the series.
4) But the law of nature is that nothing takes place without causes sufficiently determined a priori.
5) Given 1, it means that there must be sufficiently determining causes, but 2-4 show that there cannot be, which means that the assumption is contradictory, so there must be some other type of causality, namely, freedom, or the spontaneity of an absolute beginning.

It is not my task here to determine the soundness of Kant's arguments, but only to exhibit these arguments for the sake of Fichte. Suffice it to say that if this argument is sound then there seems to be a need for the possibility of absolute beginnings in the chain of causes (not absolute beginnings of time) – only these absolute beginnings could give the sufficient conditions for their causal series that proceed from them.

The argument for the antithesis is also a *reductio*, and is as follows:

1) Assume that there is freedom – the possibility of absolute or spontaneous beginnings
2) If this is so then the series will originate in the spontaneity as an absolute beginning, but the determination or choice of the spontaneity to originate will itself have an absolute beginning
3) But every action presupposes some state of the cause which is not yet active, and a beginning of this kind presupposes a state which has no causal connection with it.
4) Freedom is therefore not a kind of causality – it could never show up in experience

This argument purports to infer from the fact of a failure of transitivity on the part of transcendental freedom to it therefore not being a form of causation. But if it is not a form of causation, so the thought goes, it cannot enter into experience and cannot be said to be in accordance with laws. Having seen both the arguments of the antinomy, we should now proceed to where Kant tries to solve it.

Importantly, Kant tells us that we should not confuse what would be an answer to the question. He tells us that the solution of these problems (the antinomies as a whole) cannot be found in experience. This is because in experience we 'always remain involved in *conditions,* and come upon nothing unconditioned'.[1] In other words, our experience is such that any content we have would be of the wrong form to be evidence of the unconditioned. So the 'dogmatic solution' as Kant puts it (that is, trying to solve the problem by finding some object) is impossible. The only solution that is possible is the critical one, which 'does not consider the question objectively, but in relation to the foundation of the knowledge upon which the question is based'.[2] That is, the critical route is to investigate the knowledge that we have of the cosmological ideas of causation or time and space upon which the antinomies arise. Once we investigate that knowledge critically, Kant thinks, the antinomies will dissolve.

Kant distinguishes between transcendental freedom, which is the type of freedom that makes an appearance in the antinomies – the ability or power to begin a state spontaneously, from practical freedom, which is the will's independence of coercion from sensible impulses. Kant claims that the human will is not determined by sensible impulses because we have the ability

[1] A483/B511
[2] A484/B512

to weigh up reasons for and against actions demanded by our sensible desires, which animals are not able to do.[3] If one denies transcendental freedom, then practical freedom is also denied; that is to say that practical freedom presupposes transcendental freedom. This opens an avenue of argument present in *Groundwork* III, which I shall turn to later. For now, Kant questions whether it is a true disjunctive, or mutual exclusion, to say that something is produced by natural necessity or by freedom.[4] In other words, Kant asks the question of whether freedom could be compatible with natural necessity. Transcendental Idealism allows this because, on that picture, appearances have non-sensible grounds, which means that there is some non-sensible work going on behind the sensible appearances. So Transcendental Idealism opens the possibility that one and the same event, though in different (sensible and non-sensible) relations could be both a result of natural necessity and of freedom. This is then said to be something which is intelligible in its action and sensible in its effects.[5] The idea is not that there is some event which has both a non-sensible causal past and a sensible causal past. Instead, it is that there is one event which has an intelligible cause but whose effects are sensible. The relations this event has to future events are sensible and natural relations, but the relations it has to past events are not of this kind. Indeed it has no causal relations with past events.

Because Kant conceives as causal relations as necessarily nomological relations, this means that there must be some nomological relations that are sensible or empirical, and some which are intelligible. Because of the general machinery of Transcendental Idealism, this is not a problem in itself. The acting subject has an empirical character (or a set of empirical nomological relations) and an intelligible character (a set of intelligible nomological relations). We can then see that the intelligible character can act in the world, and it is both true that the act was an outcome of the empirical character in addition to other conditions and the laws of nature, and that the intelligible character acts, that is, one and the same agent is subject to empirical necessity and is intelligibly free.

The question is whether one and the same act could be in one sense related to past events via natural necessity, and in another sense not related to those events. Insofar as the act is an appearance, it is connected, insofar as the act is intelligible, it is not. There is an odd consequence of this view which Kant needs to defuse. The consequence is this. It seems to be as a result of the intelligible character alone that actions or agents can be said to be free. Given Transcendental Idealism's central thesis that everything is at once appearance and thing in itself, why is it not generally true that *all* natural events are at once free and determined? In other words, one might ask why Kant is restricting this duality of natural

[3] A534/B562
[4] A536/B564
[5] A538/B566

necessity and freedom to agents, given that everything has the dual nature of appearance and thing in itself, which seems to be the only requirement to have the natural necessity and freedom duality.[6] Kant's answer to this question would be, I think, twofold. The negative part of the reply would consist of the reminder that there could be no evidence that things in the world that are not agents genuinely have a non-sensible counterpart.[7] That is to say, that Transcendental Idealism is not committed to a one-to-one correlation between items in the phenomenal world and constituents of the noumenal realm, although it is generally committed to everything having a non-sensible counterpart. Only those things which are at once subjects and objects can be said to be in both realms in this correlative sense, and the only things we have any evidence for being subjects and objects are human agents. The positive part of the reply, which appears in the Critique of Practical Reason, is to say that we have reason to think that many of the constituents of the noumenal realm are agents.

Kant says that the causal connection of appearances (which seems to be distinct from an affirmation of the transitivity of causation), is a good principle of the understanding. This principle is 'not in any way infringed, if we assume, even though the assumption should be a mere fiction, that some among the natural causes [i.e. human agents] have a faculty which is intelligible only, inasmuch as its determination to action never rests upon empirical conditions, but solely on grounds of the understanding.'[8] This is somewhat different than before. Up until now in Kant's explanation and defusal of the antinomy, he has spoken as if the action *does* rest upon empirical conditions (in one sense) and *also* rests upon intelligible conditions. But now he speaks as if there is (in Sellarsian phrase) a space of reasons claim – that it is not, say, a mere desire that motivates me to act, but a desire that I took as giving me a reason to act, and that is a 'ground of the understanding'. Let's call the former style of Kantian claim the Duality of Causal Origins (DCO) theory. This latter Sellarsian style we can call Space of Reasons Causation (SRC). DCO is stronger than SRC. SRC seems to roughly correspond to the 'practical freedom' that Kant claims we have. The disanalogy is that Kant claims that were we not to be transcendentally free, we would not be practically free either. But it seems that even if DCO is false, SRC can still be true. We will see later Kant's argument for SRC in Groundwork III.

Kant now arrives at the main substance of his argument. This will be that ought implies can. He says that we always assume our reason has causality because of the imperatives we impose on ourselves and others. The ought in an imperative 'expresses a possible action the ground of which cannot be anything but a mere concept'.[9] This contains both the causation

[6] Perhaps a slightly irreverent way of putting it: Why is Kant not a Schopenhauerian?
[7] A546/B574
[8] A545/B573
[9] A547/B575

claim (that freedom is a type of causation) and the ability claim (that if we ought to perform some action that action is a possible one).[10] How is this then relevant. Kant seems to argue like this.

1) Given that I stand under obligations, I can fulfil these obligations
2) My empirical character is wholly determinate
3) So there must be something in addition which is free to choose to fulfil the obligations
4) This is intelligible character
5) Because it is intelligible, it does not stand under conditions of time, that is, it does not enter into the causal chain of effects
6) But it is still through reason that the sensible effects come to be

This is Kant's doctrine of noumenal choice, as it is sometimes called. The important point is that Kant thinks that because the intelligible character is non-sensible, it cannot stand under temporal conditions. This means that we cannot say things like 'Reason caused such-and-such' in the same way that we can say 'The wind caused such-and-such'. To think that we could unqualifiedly say that would be to put reason under the conditions of sensible intuition and make it into a form of nature.[11] This is cashed out by saying that reason is always a determining power, not a determinable one. We can legitimately ask why reason did not through its causality determine the appearances and sensible effects differently, though this question is unanswerable. Here we get to the point at which Kant sums up what he takes himself to have proved. It is not that freedom is actual (that would not have been possible as part of his transcendental philosophy). That much is clear, but what is peculiar is Kant then goes on to say that he has not even shown that freedom is possible. This is because we cannot infer from conceptual analysis or conceptual investigations alone the possibility of the causal powers of something. This is an example of what Kant calls real possibility. But he takes himself to have shown that freedom is not incompatible with nature, that is, he thinks he has shown its logical possibility. In other words, Kant thinks that he has shown that there is some possible world in which there exists both nature and freedom, but we could never identify that world with the actual world, as we cannot investigate by conceptual means alone the real causal powers of things.

[10] Of course, there are different senses of possible here. One version would be to read possible as 'permissible', which is just to say that if an action is obligatory then that action is *a forteriori* permissible. Another is to read it as a logical claim – that the action is logically possible. A third is to read it as a claim about our abilities – that it is within our power to perform it. The first (if obligatory then permissible) is sound, but uninformative. It would be a contradiction if an action was both obligatory and impermissible. It would clearly be fallacious to infer from that feature of the logic of deontic operators to the claim that ought implies can as standardly understood. For an argument that ought does not imply can, see Martin (2009).
[11] A552/B580

I shall now turn to some criticism of Kant's view. One problem is that Kant seems to suggest that we really are practically free, and practical freedom requires transcendental freedom, and so we are entitled to conclude that we are transcendentally free. But now he says in no uncertain terms that we do not have the warrant for this. Recall that Kant says 'For practical freedom presupposes that although something has not happened, it *ought* to have happened, and that its cause, in the appearance, is not, therefore, so determining that it excludes a causality of our will'.[12] Practical freedom presupposes the causality of our will, which is free either way to conform with what it ought to or not. So it looks like practical freedom presupposes transcendental freedom. But then the only way to escape the conclusion of transcendental freedom would be to deny that we can know that we are practically free.

One way to think about this is the following: What would the world be like if freedom did not exist? It seems intuitive in some way to think that the world would be rather similar to the way it actually is, assuming freedom exists. That is, we would still be confronted with the need to make choices and deliberate about what to do. The inference from being determined to fatalism (something like the view that one wouldn't have to choose what to do because it would happen anyway if it was going to happen at all) is not a good one. Practical freedom, however, seems to be just this deliberation and choosing about what to do and what one is obliged to do. Of course, it might be said that this is just seeming to be practically free, rather than actually being practically free. I am not sure the seems/is distinction could apply here. But we should leave this and move on to other criticisms for the moment.[13]

Two that seem to present themselves are an objection from the causal closure of the physical and a worry about overdetermination. The first is that Kant's account seems to violate the principle that any physical event must have a sufficient physical cause. It is not entirely clear that it really does so.[14] It is a feature of Kant's account that actions do have sufficient physical (or sensible) causes, but that they also have another type of causal origin. But then the worry about overdetermination looms. Why think that we are free at all (in Kant's sense of the causation by reason) if an entirely sufficient account of our actions could be given using sensible and physical events alone? Moreover, Kant seems to then posit a non-sensible causal origin, though whether this non-sensible origin is sufficient is unclear. It might seem to be sufficient because it stands in no causal (or even temporal) relation with prior events, but then it might be insufficient because it might be that reason needs an instantiation of some

[12] A534/B562

[13] I will return to this when discussing the Groundwork.

[14] Though Wood (1984) seems to think that Kant is committed to this. Kosch (2018) thinks that Kant is likewise committed, and it is one of Fichte's differences from Kant that Fichte denies this causal closure.

kind in the empirical world in order to affect it properly.[15] The intelligible character, being intelligible, has no sensible body, and therefore has no organ with which to affect the world. So it needs to be instantiated somewhere, in the empirical character, which is tied to a body that has causal powers.

A fourth worry that may arise here is to do with what I refer to when I say 'I' in the locution "I freely chose to X". am I referring to the empirical character, the intelligible character, or some third thing, an amalgam of them both? It seems that I could not be referring to the empirical character alone as that is an entity governed solely by the operations of natural laws (no different from an organ in the physical body, albeit it is governed by psychological laws instead of physiological). So am I then referring to the intelligible character alone? This is also unsatisfactory. Surely, if "I" has any reference at all, it is not to a shadowy metaphysical complement of me, but to the embodied and empirical and non-empirical duality – the self-conscious thing. So it seems that it should be construed as referring to the amalgam of the intelligible and empirical characters. Then we have a situation in which the locution "I acted thus" is shorthand for "I acted thus *in virtue of* the intelligible choice I made". Notwithstanding the worry about the reference of the second "I" here (which seems to need to refer to the intelligible character), this means that Kant has to supply an account of how the intelligible character becomes manifested or instantiated in the world as the empirical character. Apart from other concerns, the one most prominent for Fichte would be the necessary introduction of the doctrine of noumenal affection – the idea that the non-sensible world in some sense influences us in addition to the sensible world (so-called Double Affection). This replicates the worry about over-determination, but also introduces a problem more worrying for Fichte. That is the problem that I might be deeply or systematically misled over what I refer to when I intend to ascribe the performance of an action to myself. Insofar as Kant's account stands or falls with the metaphysics of Transcendental Idealism understood as a rigid and unfailing distinction between appearances and things in themselves, Fichte is going to see this as unsatisfactory.

This account of the logical compatibility of freedom and nature is therefore problematic. It is also unclear, even if these problems are solved, how the empirical and intelligible characters relate and how something that is not in time can affect something temporal. In my view, Fichte sees these as problems to be solved, partly through the jettisoning of the core aspect of Transcendental Idealism that Kant used to navigate through the problem – the distinction between appearances and things in themselves. For Fichte, this would cause unsolvable issues even if it turned out to be consistent. Fichte thinks that the real possibility of freedom is shown by intellectual intuition into one's own thinking. So Fichte thinks he can do without

[15] I think Fichte is sensitive to just this issue, and attempts to overcome it with his account of the relation of the will to the body.

Kant's central distinction. Fichte therefore wants an account which is monist (or at least not dualist in Kant's sense), which does not commit him to overdetermination. As for causal closure, Fichte will agree that Kant violates it in a problematic way because of things-in-themselves. But Fichte will also think that causal closure is false insofar as it is my willing that P which is the operative causal factor. He agrees with Kant that freedom is a type of causation and that it is causation via concepts. The most important thing we have learned over the last few paragraphs is that Fichte will want an account of freedom which preserves a monistic account of the causal locus. Having looked at the Third Antinomy and what it can show us about the task that Fichte saw, we should turn to the Groundwork for Kant's other argument about freedom – that we normatively speaking cannot think of ourselves as determined.

1.2) *Groundwork* Part III

Kant begins G III by stating that we can define free will in two ways – positively or negatively. The negative definition is that it is the property of causality such that the causality could be efficient 'independently of alien causes *determining* it'.[16] The positive definition is bound up with the nature of causes and their relation to laws in Kant's mind. Causality implies lawfulness – so for the free will, if it is to be a type of causality, it must be also an exemplification of lawfulness. Given the negative definition, this lawfulness cannot be alien to the will. All natural impulses are alien to the will, or heteronomous. So the lawfulness of the free will must consist in 'the will's property of being a law to itself'.[17] The statement that the will is a law to itself implies the principle that to will should act on no other maxim other than that it could be universal law – the principle of morality. So Kant concludes that a free will and a will under moral laws are one and the same.[18] This continues the ideas from the first *Critique* – that a free will is itself a type of lawfulness. Kant now proceeds to argue that we can fairly assume that we are all practically, as opposed to transcendentally, free. He says:

'…every being that cannot act otherwise than *under the idea of freedom* is just because of that really free in a practical respect, that is, all laws that are inseparably bound up with freedom hold for him just as if his will had been validly pronounced free also in itself and in theoretical philosophy.'[19]

[16] Kant *G* 4:446
[17] Kant *G* 4:446-7
[18] *G* 4:447
[19] *G* 4:448

This passage is Kant's statement in favour of regarding agents this way. But why think that we cannot act otherwise than 'under the idea of freedom'? This comes later in the paragraph:

'Reason must regard itself as the author of its principles independently of alien influences; consequently, as practical reason or as the will of a rational being it must be regarded of itself as free…'[20]

The thought here seems to be that the nature of reason entails that we could not coherently view ourselves as genuinely making choices and regard ourselves as determined to do so.[21] Kant is certainly not arguing that it is psychologically hard to do so, or even that it is as psychological fact impossible, but that whether there is a psychological impossibility notwithstanding, there is a logical difficulty. The logical difficulty is that I could not genuinely see it as my own choice if I did not also think that the future was genuinely open in some sense. Kant does not think he has shown anything substantive here[22], but merely that there is an incoherent set of attitudes, and the coherence is resolved by regarding ourselves as practically free – which has been defined in *Critique* as the independence of the will from sensuous impulses (the negative definition offered at the beginning of *G* III).

After recounting the details of his earlier theory, and the addition of the metaphysics of transcendental idealism, Kant seems to make a surprisingly strong claim. He says 'it is an indispensable task of speculative philosophy…to show that both [freedom and natural necessity] not only *can* very well coexist but also must be thought as *necessarily united* in the same subject'.[23] It is not enough, in other words, to show that it is possible that freedom and natural necessity are both in some sense (though a different sense) true, it also needs to be shown that for a self-conscious subject (presumably one who acts under the idea of freedom), natural necessity and freedom are united, and united in such a way that it couldn't be true that they were not. This seems to be alright for Kant to assert. He thinks he has already shown that they can coexist, and that any self-conscious agent thinks of themselves under the idea of freedom. The natural necessity of the empirical nature of the subject and the freedom of the intelligible nature of the subject come together in a unity to produce a self-knowing agent who is subject to the moral law in a world of sense.

A further peculiarity attaches to this paragraph. Kant says that showing the unity must be done, otherwise 'no ground could be given why we should burden reason with an idea which, though it may *without contradiction* be united with another that is sufficiently established, yet

[20] *G* 4:448
[21] This will surface again later in the TAAD.
[22] *G* 4:448-9
[23] *G* 4:456

entangles us in a business that brings reason into difficult straits in its theoretical use'.[24] The problem that Kant sees is that given the theoretical difficulties that we get ourselves in when we involve ourselves in discussions of freedom, it might seem worthwhile to enquire whether we have some grounds for attributing freedom co-existing with natural causality.[25] But the wording of the passage shows that Kant has in mind the necessity of thinking of freedom and natural causality as united, rather than the necessity of us having insight into the ground of this unity. This is then said shortly after to be impossible. Kant has outlined a way to show that freedom can be theoretically justified, and then says that this would be to overstep the boundaries of reason.[26]

This oversteps the bounds of reason because explanation requires one to appeal to 'laws the object of which can be given in some possible experience'.[27] To explain how freedom is possible would then require that one be able to show the object (freedom) that falls under the (moral) laws. Kant seems to be having a Humean thought. Just as Hume asked to be shown the causal connection – he could see the billiard balls and the collisions, but not the causal connection – Kant says that he can see the person and the movement of their body but not their freedom. So for Kant, free actions are inexplicable, and all philosophy can do is to show that it is not contradictory – a defence, not an explanation.[28] To sum up, Kant's position as stated in *G* III is that freedom united with natural necessity is, from a theoretical standpoint, possible and from a practical standpoint, necessary.

There are two main issues here. One is a continuation of the worry above regarding the reference of "I". Another is about a disparity in Kant's argumentative strategies in *Critique* and *G*. I shall start with the former. Kant says more than once in *G* III that the intelligible self is the 'proper self'. At one point he says 'in it [the intelligible world] reason alone, and indeed pure reason independently of sensibility, gives the law, and, in addition, that since it is there, as intelligence only, that he is his proper self (as a human being he is only the appearance of himself)…'[29] At another he says 'the law interests because it is valid for us as human beings, since it arose from our will as intelligence and so from our proper self…'[30] This is a related, but slightly different worry to above. The above concern was that Kant has to think that the correct reference of "I" in "I acted freely" is the union of the natural and non-sensible. Recall that we said that "I acted freely" for Kant is a condensed form of "I acted thus in virtue of the intelligible choice that I made". Here Kant tells us that our proper self is a purely intelligible

[24] *G* 4:456
[25] *G* 4:455
[26] *G* 4:458-9
[27] *G* 4:459
[28] *G* 4:459
[29] *G* 4:457
[30] *G* 4:461

thing that has made an intelligible choice – this is therefore the reference of the second token of "I". But he says that purely intelligible beings that make purely intelligible choices would have their wills perfectly autonomous – that is, in accordance with the laws of morality. Given the obvious fact that my will is not always in accordance with what I would choose were I to be intelligible only, what exactly is my 'proper self' doing?[31]

Clearly Kant thinks of the 'proper' self as an agent considered as a thing in itself. One way to get him out of this would be the following. The intelligible character wills the moral law, but because empirical things have natural principles or laws governing them as well, this intelligible will is corrupted or made impure. So there is an intelligible choice which is transmitted by the grounding relation that things in themselves have to appearances, and the transmission, because it needs to be instantiated in something that is alien to reason (i.e. nature) is corrupted on the way. This is possible, but Kant's account of radical evil that he develops in 1793 is not this. That account relies on the intelligible choice itself being for evil. It could then be possible that Kant in 1785 had this view, but this was not his considered mature view. Another way to get Kant out of this issue is to point to the passage where Kant invokes the 'practical use of common human reason'.[32] He says that 'not even the most hardened scoundrel' when confronted with examples of good virtues 'does not wish that he might also be so disposed', and 'wishes to be free from such inclinations, which are burdensome to himself'.[33] Thus our hardened scoundrel (and everyone else) shows that 'with a will free from impulses of sensibility he transfers himself in thought into an order of things altogether different'.[34] In other words, Kant seems to be conceiving of sensuous impulses here as something alien to the will, and we can be meaningfully said to conceive of ourselves as ourselves without them. So our 'proper self' is a kind of idealised self that we locate in the world of things in themselves, given that we know that such a world exists (though can know no more about it), and given that such a world has no natural impulses, which are alien to reason, it follows that the self we locate in that world would indeed truly be us, in some ways more truly us than we currently are.[35]

Here is then the problem. If this is what Kant means, then he can give no account of evil apart from the transmission-fault story I gave above. But that locates evil in alien causes, not in freedom. This means that there are no free evil acts, which, if moral responsibility requires freedom, means that there are no evil acts. Kant does remedy this with his later accounts, but

[31] This comes to a head with Kant's account of radical evil. With that account, Kant seems to discard the principle that any intelligible choice would be in accord with the moral law.
[32] *G* 4:454
[33] *G* 4:454
[34] *G* 4:454
[35] This suffices to make sense of the text at hand, I think, but Kant is still going to have to give an account of free evil action, which requires amending this account.

we should move on to the second worry that I mentioned, regarding the disparity in argumentative strategy.

Recall that in the *Critique*, Kant seems to claim that practical freedom (understood as freedom from sensuous determination) implies transcendental freedom. Now, in *G*, Kant seems to be arguing that we can legitimately regard ourselves as practically free. So why does he not just say that given that, we can legitimately regard ourselves as transcendentally free? That is, why is it that Kant neglects to argue thus:

1) We can legitimately regard ourselves to be practically free – that is, we must always act under the idea of freedom
2) If we must always act under the idea of freedom, then we can regard ourselves as transcendentally free
3) So we can regard ourselves to be transcendentally free.

One explanation is that the assertion of practical freedom is a seems claim, not an is claim. Kant does not claim that we are practically free, just that we can legitimately see ourselves as such. One might question whether there is genuinely a seems/is distinction here. It is a familiar point from Kripke (1980) that some judgements (of pain, for example) have a self-authenticating character – If I sincerely judge that I am in pain, then I am in pain. Similarly, one might say that if I seem to myself to be practically free, then I am practically free – it couldn't be otherwise. If practical freedom is just understood to be freedom from sensuous determination, then this obviously does not work. Kant's insistence on the opacity of the self – not merely that self-knowledge is hard, but that knowledge of the self as it is in itself (as opposed to as it appears) is impossible - means that this could not work.

What about a different way to think about it? Suppose that practical freedom, or acting under the idea of freedom, means that we generally deliberate or decide what to do.[36] One might argue the following: There is no seems/is distinction for deciding or deliberating. I cannot seem to decide what to do without actually deciding what to do. So I cannot seem to act under the idea of freedom without actually being under the idea of freedom.[37] I think that Kant would say that this too falls to the opacity criticism. Whilst some determinists may want to keep the ability to say things like "I decided to…" it's not clear that they can – surely all decision would have been made independently of me at all.[38] Another way of saying this is that if determinism is

[36] This is how it is generally taken by commentators, notably Korsgaard (1996)
[37] I offer this as a suggestion of a possible move, though it seems that the analogy with pain does not work: the pain example works because the phenomenal character of pain is constitutive of pain; decision-making plausibly has a phenomenal character but this doesn't appear to be constitutive.
[38] Van Inwagen's Consequence Argument is relevant here – if my 'decision' is the inevitable consequence of a series of events which extend beyond my birth in accordance with natural laws (which is how Kant would conceive of it), then my so-called decisions are illusions.

true, I might be the occasion for a causal series but not the locus of that series, which I would need to be if transcendental freedom is actual. But given that to assert this would be to go beyond the bounds of reason, Kant cannot assert this. Kant's conclusion is that for practical purposes (which are necessary) agents need to think of themselves as able to stand under practical laws, or that reason is itself practical.[39]

I am now going to summarise what we have learnt from the Groundwork. Kant has kept his earlier stance on the metaphysics of transcendental idealism as the key to the problem of how freedom and nature can coincide. What he has added is a different argument for the (practical) necessity of taking ourselves to be free. This practical necessity is then a warrant for thinking of ourselves as having the capacity of freedom united to the natural necessity in us. How this unity is possible is unknowable – free acts cannot be explained *as* free; all explanation is in terms of laws of nature. The point which is likely to be attacked is the idea that we cannot but take ourselves to be free when acting.[40] As discussed above, one way to defend this would be to say that there is no seems/is distinction, but Kant cannot appeal to that because of his view of the opacity of the self. Fichte will make similar observations to Kant, but they will not in the last analysis rest on an account like this. Fichte also thinks that we must 'take ourselves' to be free, though he does think that the dogmatist will, as far as theory goes, have an account of this which makes the 'taking' an illusory one, and so Fichte's account could not rest on this alone. Fichte must therefore locate freedom deep within the agent, in order to safeguard it. We should proceed to look at the second *Critique*, to see the Fact of Reason argument.

1.3) The *Critique of Practical Reason*: The Fact of Reason

The famous Fact of Reason appears in the second Critique, where Kant appears to have a change of mind. In the Groundwork, Kant tries to provide a deduction of freedom, which answers the question *quid juris* regarding it. Here, however, Kant seems to reject such a project in favour of a much more direct appeal to common rational moral cognition.

Freedom is said to be the only idea (in Kant's technical sense of idea) that we can know *a priori*. This is because it is the condition or *ratio essendi* of the moral law. By contrast, the moral law is the *ratio cognoscendi* of freedom. The moral law is then the way we know we are free, but the moral law consists in freedom (in the sense of autonomous or self-legislative activity). So the Fact of Reason is simply that we feel ourselves to be morally constrained. I shall now turn to the passage in question.

---

[39] *G* 4:458

[40] See Farrer (1958) for a defence of this, albeit one without reference to Kant. It is of course a key aspect of Korsgaard's (1996, 2009) account that we must take ourselves to be free.

For Kant, freedom and the moral law (or 'unconditional practical law') reciprocally imply one another.[41] Kant is concerned with the order of our cognition when it comes to freedom or the moral law. He argues that freedom could not come first in the order of discovery, so it must be the moral law that precedes. There are two arguments in the text, one with respect to the architectonic of reason and one from experience. The architectonic argument is essentially the following:

1) Theoretical reasoning has no need of the concept of freedom
2) So something other than theoretical reasoning must have introduced this concept, which is practical reason
3) The introduction takes the form of the moral law, or a purely practical law
4) So the fact that we have the idea of freedom shows that we have prior cognition of a moral law

Kant goes on to say that experience also shows this conceptual order to be right.[42] He illustrates this with the famous example of a man who, on being asked whether he give false testimony against another on pain of death, he knows that, even though he may not want to die, he could choose to do so in order to comply with his moral duty. Therefore he judges that 'he can do something because he is aware that he ought to do it and cognises freedom within him, which, without the moral law, would have remained unknown to him'.[43] So instead of Kant deciding to give a philosophical deduction of freedom, he instead opts to an appeal to the order of discovery of the concepts of moral law and free will in order to show that the only ground on which we can know that we are free is that we know that there are actions we could perform which are morally required of us, and that we could perform them because they are morally required.

There are two important remarks to make here. Firstly, on the reason why Kant thinks that the order of discovery has to come this way, and secondly, on why it is appropriately entitled a fact of reason. The reasons Kant gives as to why we cannot discover freedom first, as shown in the architectonic argument, are a few. He says that freedom would never be needed by theoretical reasoning (as per premise one of the architectonic argument), and that we could never learn it from experience, because 'experience lets us cognise only the law of appearances and hence the mechanism of nature, the direct opposite of freedom'.[44] He also says that we cannot be immediately conscious of freedom 'since the first concept of it is negative' and that if positive freedom were presupposed, then we could infer from freedom to

[41] KPV 5:29
[42] KPV 5:30
[43] KPV 5:30
[44] KPV 5:29

a moral law (and this would be analytic) but that it requires an intellectual intuition.[45] What does it mean to say that the first concept we have of freedom is a negative one, then? It seems that what Kant has in mind is this: freedom is a negative concept insofar as the concept is of the independence of the natural law of appearances, or the law of causality.[46] So the only concept of freedom we have to begin with is a concept of some action or process being uncaused, or having a beginning that is preceded only by other events in time, not in a causal series. The positive concept of freedom is what we have after we learn of the moral law. This is because freedom informed by morality becomes not merely the independence of causes, but the law-governed performance of moral actions. In other words, Kant's reasoning is that if we had the positive concept of freedom, we could infer from this to the moral law, but it would be a trivial inference – which is why he says it would be analytic. To have this positive concept of freedom from the beginning, one would have to have an intuitive insight into the nature of rational beings – an intellectual intuition – which on Kant's view, we cannot have.

The second remark to make was on why it is appropriately called a fact of reason. Kant's answer may be quoted in full:

'Consciousness of this fundamental law may be called a fact of reason because one cannot reason it out from antecedent data of reason, for example, from consciousness of freedom (since this is not antecedently given to us) and because it instead forces itself upon us of itself as a synthetic a priori proposition that is not based on any intuition, either pure or empirical…'[47]

So we do not have some given or more basic data to work with from which we can infer that there is an unconditional practical law, but neither is it based on any intuition. In other words, it itself (consciousness of unconditional practical law) is given. But it is not given through the senses, or through the imagination. We simply comprehend it. From there, we can infer that we have the freedom to conform to it or not.

The most important thought here, from a Fichtean point of view, is the idea that either one cannot, or one need not, give a transcendental deduction of the moral law. It is as though Kant has given a metaphysical deduction, instead. As we shall see in chapter 3, Fichte does think that it is possible to give transcendental deductions of the moral law, when the deduction is also a genetic one. Because of the genetic or developmental nature of Fichte's deductions, there is also the question of whether he agrees with Kant's view on the order of discovery between the moral law and free will. One more thing to say is that, for Fichte, because the

[45] KPV 5:31
[46] KPV 5:29
[47] KPV 5:31

moral law and free will are reciprocally implicatory concepts, it is not possible to have consciousness of the one without the other (because possession of both is a necessary condition on self-consciousness).[48] This still leaves open the question of conceptually explanatory relations, however. This is where Fichte would make a decisive break with Kant. Fichte of course accepts the possibility and the actuality of intellectual intuition.[49] At every moment of consciousness I am aware that I am self-identical, and that this awareness of a subject and an object being identical is at least one of the various meanings that Fichte gives to intellectual intuition. As far as insight into freedom goes, Kant says that it requires intellectual intuition because freedom is an idea of reason. For Fichte, freedom is not only, as it is for Kant, the keystone of the system of reason, but it is the motor, the driving force, of all aspects of a rational being's life, both as thinker and as agent. The very idea that rational beings could be as divorced from insight into their nature, as Kant thinks they are, is anathema to Fichte.

To conclude this section, we can see that Kant's thinking takes a turn which Fichte, in his mature thinking, would not approve, even if his initial reaction to the Critique of Practical Reason was overwhelmingly positive.[50] This is because Kant's sidestepping of the question of a deduction of the moral law runs counter to Fichte's systematic ambitions. Fichte would also charge Kant here of artificially separating consciousness of freedom from consciousness of the moral law, which Kant needs to do because he thinks of freedom as an idea of reason, a part of the architectonic that Fichte will reject. We should now turn to Kant's account of free will and related ideas in the *Religion*. We will do this with special attention to Kant's view on evil and on the distinction between free will and free choice, or *Wille* and *Willkür*.

1.4) *Religion within the Boundaries of Mere Reason* & *Metaphysics of Morals*: Radical Evil and the Will

It is commonly thought that Kant's distinction between *Wille* and *Willkür* arose as a response to criticisms that Kant's view would make it impossible to act wrongly and freely. This may not be true, but either way, Kant clearly thought that the distinction needed explicating.[51] *Wille* can be translated as will, or willing, whilst *Willkür* is a choice, closer to the Latin *arbitrium*. This roughly correlates with two broad understandings of free will, one more prevalent prior to the

[48] For Fichte it would even be misleading to speak of two concepts – they are one and the same.
[49] Though, as many have pointed out, it is not always clear whether Kant and Fichte use this term in the same way.
[50] As seen in an oft-quoted letter where Fichte states that he has been 'living in a new world' since reading it.
[51] Guyer (2017) provides contextual remarks on the debate between Kant, Schmid, Reinhold, and Ulrich.

Enlightenment, and one more prevalent after. The first (the *Wille* understanding) is that to be free is to be correctly oriented toward the good, so I am unfree when I go wrong with respect to actions (but still blameworthy). This is bound up with theological views about being in 'bondage' or 'serfdom' to sin, or evil in general.[52] The second (*Willkür*) is about whether I am free with respect to an action or not, irrespective of whether that action is good or bad. This is just the question of whether my choice is free.

The distinction is important because on a Kantian view whereby morality and freedom are closely tied together, we need to be able to say that one acted wrongly and freely. Kant sometimes speaks as if *Wille* is properly speaking neither free nor unfree (that is, the predicate-pair free/unfree does not apply) but *Willkür* is free. Kant says that because the will is not directed toward actions (that is the job of choice) it cannot be called either free or unfree. Kant says the will 'which is directed to nothing beyond the law itself, cannot be called either free or unfree, since it is not directed to actions but immediately to giving laws for the maxims of actions (and is, therefore, practical reason itself)'.[53] Even though our will is identical with practical reason and is therefore a lawgiving capacity, people do, as Kant notes 'choose in opposition to [their] (lawgiving) reason'.[54] So from the Metaphysics of Morals, which is Kant's final explanation of the distinction between *Wille* and *Willkür*, we can see that the will itself is practical reason, but because the human being is not merely an intelligible being, but also a sensible one, the human being has the ability to somehow act contrary to their own lawgiving reason. This capacity of choice is free in the sense that it has the ability to elect between two options, but it also requires grounding on some prior principle. The prior principle is then either autonomous (i.e. free) or not. Kant therefore has both conceptions of freedom, the Anselmian, according to which freedom is to be correctly aligned to the good, and the Enlightenment, in which freedom is capacity for choice. Our capacity for arbitrary choosing is informed by the principle that we accept, and therefore we can freely choose evil. Kant says that we 'cannot comprehend how this is possible' even though 'experience proves often enough that this happens'.[55] Whilst Kant does have a story about evil which I will turn to shortly, this functions more as a name for the phenomenon here than an explanation of it. In other words, to use the theory of radical evil to attempt to comprehend why people act contrary to their own lawgiving reason assumes the fact it tries to explain.

[52] For examples, see St. Anselm's (2000) *De Libertate Arbitrii* and Luther's (1979) *On the Bondage of the Will*.
[53] Kant MM 6:226
[54] Kant MM 6:226
[55] Kant MM, 6:226. Kant could be talking about weakness of will – knowing the better yet choosing the worse – but it seems that the parenthetical 'lawgiving' is supposed to bring to mind something more fundamental. Kant calls weakness of will the 'frailty of human nature' *Religion* 6:29

Radical evil is Kant's reinterpretation of original sin or the concupiscence toward sin.[56] In Kant's view, every rational agent has made a noumenal or intelligible choice to put self-interest or self-love ahead of the demands of morality. This intelligible basic choice issues in empirical actions which reflect that intelligible choice – crucially the intelligible choice cannot be unmade or reversed by empirical choices. Similarly to the doctrine of original sin, Kant thinks that there is a radical, or root, cause of evil within human nature. But contrary to at least some interpretations of the doctrine (e.g. Calvin's), which subscribe to what is sometimes called "Total Depravity", Kant does think that there is a seed or equally basic root of goodness in the human will. But, as I said above, Kant does adhere to part of the traditional Christian view that a necessary condition of becoming better is by the work of the Holy Spirit or through God's grace.

The will needs to be oriented toward a proper apprehension of the good, but if this is not so, then moral conversion is needed, which can only be achieved via God's grace. But we can still say that the arbitrary choice of the individual is free, so they can still be said to be responsible and/or blameworthy for the actions they perform, even if their will is set or immutable (at least immutable with respect to their own action – surely it would not be immutable with respect to God).[57] But the person themselves cannot change their own will in such a way as to change their fundamental orientation towards the good, because of radical evil. This also shows why radical evil could not be used as an explanatory device to the question in the Metaphysics of Morals as to why agents will choose in opposition to their own reason. The theory cannot explain this because it relies on a prior act of choice to put self-love above morality – which, being the source of evil actions, cannot be explained by a theory of why evil actions happen. Evil remains in some sense inexplicable. The puzzle here, however, is that Kant does say at the beginning of the relevant section in the Religion that he could give a formal proof of the propensity to evil, but given the facts of evil are so obvious to us it would be redundant. This suggests, at least, that there is some further principle from which we could derive evil.

In a sense, Kant gives a 'metaphysical' deduction of evil, without the 'transcendental' deduction, which would show that we have the right to apply the concept to the world. This is probably because Kant has already given reasons for thinking that we have the right to think in moral terms (the fact of reason), and so anyone who would want to deny the right to apply the concept of evil to the world would either a) have to show that the fact of reason was

[56] Kant seems to mean radical evil to fulfil the job of the theologian's concept of concupiscence – he uses the word at 6:29.

[57] This is remarkably close to Luther's distinction between the necessity of compulsion and the necessity of immutability (Luther 1979, 119f, 211). For Luther, an agent who sins does so freely (provided they are not compelled to do so by a coercive force) because they voluntarily will that sinful action. But their will is of such a character as to be perverted by sin.

absolutely illusory, b) have to argue that there actually exists no evil. The former is a large burden for any theory to bear, and would need to be accompanied by a general position according to which such things are generally illusory (otherwise morality would be receiving special treatment for no justifiable reason).[58] The latter would have to provide a much less parsimonious explanation for the apparent existence of evil, and it seems very close to being self-defeating – if the existence of evil is wholly an illusion then it is surely a terrible fact that we labour under such an illusion.[59]

We should move on to discuss particulars of Kant's presentation. An important point is that he says that the investigation is to whether the human being is good or evil 'by nature'. This, he tells us, means only 'the subjective ground…of the exercise of the human being's freedom in general…antecedent to every deed' but this subjective ground 'must, in turn, itself always be a deed of freedom…Hence the ground of evil cannot lie in any object *determining* the power of choice through inclination, not in any natural impulses, but only in a rule that the power of choice itself produces for the exercise in its freedom, i.e., in a maxim'.[60] When Kant uses the phrasing 'by nature evil' or any similar phrasing, he therefore only means to point to the fact (which Fichte affirms as well) that the human being has a certain level of control over what kind of being they are, and this control is expressed through the use of maxims, or subjective principles of action. Kant goes on to say that when we say the human is by nature evil or good, 'this only means that he holds within himself a first ground (to us inscrutable) for the adoption of good or evil (unlawful) maxims'.[61] Here we see that for Kant, whilst the adoption of the first good or evil maxim (which is the ordering of morality and self-love) is itself inscrutable, the adoption of later maxims can be explained with reference to the basis of this more primary maxim.

The major question which needs to be asked about Kant's views on the relationship between freedom and evil is the following. Where is the choice for the evil maxim located? There are two options, it seems. Either it is located in the empirical part of the human being, or it is in the non-empirical or intelligible part. It might seem that, given Kant's views on freedom that we have seen in the first two *Critiques* and the *Groundwork*, that the choice is at the noumenal or intelligible level. But this does not sit well with Kant's long-held doctrine that the will is identical with practical reason, i.e. the source of moral authority or normativity. If we were just *Wille*-

[58] I have in mind such positions as eliminative or reductive naturalists (who would indeed argue that consciousness and many attendant things are illusory, or at least lead to illusions) or perhaps Nietzsche (who has a general view according to which it is not only morality which is illusory, but also selfhood and free will). In the First Introduction, Fichte imagines the dogmatist taking a similar route.
[59] It must be said that Kant uses the phrasing 'formal proof' rather than 'deductions' so my analogy may be inapt. Kant *Religion* 6:32-3. Wood has pointed out that recent attempts to fill the so-called gap in Kant's argument fail to grasp Kant's point that the innateness of evil in human nature is an empirical effort, to be completed by future anthropology. (2014, 55-6).
[60] Kant *Religion* 6:21
[61] *Religion* 6:21

having beings, as opposed to sensible beings as well, then we would not feel the moral law as an imperative. So it seems odd to locate the source of evil at this level. Perhaps then it is located at the empirical level. But Kant tells us that the ground of the evil cannot be placed 'in the sensuous nature of the human being, and in the natural inclinations originating from it' (6:34-5). It also cannot be placed in what Kant calls a '*corruption* of the morally legislative reason, as if reason could extirpate within itself the dignity of the law itself' (6:35). This seems to correspond to locating evil at the intelligible level – also called by Kant an 'evil reason'.

Neither of these are appropriate ways of grounding evil. Kant appears to say that, for a couple of reasons, the origin of evil, or the choice for the evil maxim, is inexplicable, and cannot be precisely located in either of the two aspects of the human being. He says, 'We must not however seek an origin in time of a moral character for which we are to be held accountable, however unavoidable this might be if we want to *explain* the contingent existence of this character' (6:43). His thought here seems to be that if we investigated the origin of evil enough, we would perhaps be drawn to the conclusion that it all stemmed from natural causes or from childhood dispositions, not fully (or even at all) under the control of reason. Therefore, given that Kant thinks we have good grounds to affirm our freedom with respect to evil actions, we should not seek to give such explanations, and should rest content with the proposition that they are in a sense inexplicable or incomprehensible (which he affirms at 6:39-40).

So the question of at what level the choice for the evil maxim is to be located is a misguided one. The choice for the evil maxim is a way of expressing the point that evil must stem from evil (6:43), and to point to the fact that ultimately, because human agents are free 'all the way down', their evil actions must be seen to be freely chosen and represent a value-ordering system which subordinates the moral to the non-moral.

It is a major point of difference between Kant and Fichte that the former does not give a theoretical proof, and seems to think that there is no *a priori* proof, but the latter thinks not only is there an *a priori* proof but that he can give it. Secondly, whilst for Kant evil is inextricably bound up with society (Wood 2014, 49), for Fichte the root of evil is much more 'individualistic'. It is of course true that Fichte recognises that social conditions can bring new vices into play, or make existing vices worse, and it is true that Fichte does think that there are arguments which establish that I could not be self-conscious unless I was conscious of rational beings outside me. But the ground of evil lies in the individual *qua* natural being, rather than individual *qua* social being, for Fichte. We shall see this in more detail in chapter 5.

So what have we learnt from this brief overview of Kant? We know that Fichte wants to maintain a Kantian view on freedom, or at least sees himself as firmly inside the Kantian tradition. This means that our interpretation of Fichte should uphold at least the following criteria. Firstly, and most importantly, any interpretation of Fichte's views on free will out to be

able to explain how the rejection of the noumenal self or things-in-themselves affects free will. Fichte sees Kant's views as either in line with his own (if Kant does not really subscribe to things in themselves) or to be changed in light of the rejection of things in themselves. The context around at the time, with the publication of Reinhold's 1789 *Essay on a New Theory of the Human Capacity for Representation,* was concerned with remodelling Kant's philosophy with a first principle. Fichte subscribed to this project, at least for a time, and partly, we might think, in order to make the subject self-sufficient, to provide sure grounds for freedom. In other words, in order to remove noumenal affection or the noumenal self from Kant's system, Fichte attempts to ground the system on a first principle, which is independently supposed to be an attractive accomplishment anyway. Given that it is a hallmark of Fichte's idealism that it rejects orthodox Kantian conceptions of things-in-themselves (whether it has room for these in some other capacity, as some have argued),[62] we should be able to explain how Fichte's views of free will are articulated to accommodate this.

It is not that Fichte sees Kant's views as inherently contradictory with things in themselves, but that Kant has stopped short of the real goal, the complete self-sufficiency of the subject and the wholesale rejection of things in themselves, which has meant that Kant's views are consistent, but deeply unsatisfactory, for they provide no systematic deduction of free will. Secondly, it should be able to do justice to two desiderata. It should be able to accommodate Kant's insight of the Fact of Reason, but also the aims of systematic transcendental philosophy. That is, it should explain how it is that Fichte thinks that free will is in a sense obvious, and how the philosopher can provide a deduction of free will. Thirdly, it ought to be able to show how rational agents can freely choose evil, because Kant's view is confronted with a problem – in order to explain how agents have chosen the evil maxim, it requires there be a ground of evil already in the agent, which is contrary to the theory. Fourthly, it should do justice to Fichte's insistence on the primacy of practical reason. This has been a major feature of recent Fichte scholarship, and rightly so. We need to be able to see, however, how it meshes with his views on free will.

1.5) A Note on Reinhold and the Early Fichte

It would be remiss of me to move on without taking into account the controversy that developed when Fichte was converting to Kantian views, and the central players involved. Whilst I do not intend here to provide any historical accounts of the transition from Kant to Fichte,[63] there are issues that need addressing. It is well-known that various philosophers, including Schmid,

[62] For example Beiser (2002), Rockmore (2010).
[63] Such an account is provided by Beiser (1987). A compact assessment of this history insofar as it relates to free will is Guyer (2017)

Ulrich, and Reinhold, engaged in a controversy after the publication of Kant's *Groundwork,* which centred on Kant's supposed view that one was only free when one acted morally, and unfree otherwise. This of course would entail that all evil actions were unfree, and thus punishment of them would be *prima facie* unjust. This, in Kant's mind, was because of a confusion between *Wille* and *Willkuer*. *Wille* (Will) is practical reason, or the source of the normativity of law. *Willkuer* is 'arbitrary choice', which is our capacity to act in conformity with the moral law or not. Kant had made a distinction between these two clearly in his late work, which we have just looked at. A popular narrative, however, has it that Reinhold, in his *Letters on Kantian Philosophy* argued that Kant had made just the mistake of arguing that immoral actions could never be free. But Reinhold takes himself to be defending Kant against others – chiefly Schmid, and the system of 'intelligible fatalism'. "Intelligible fatalism" is one of those terms which seems to mean a different thing depending on who one asks. We shall see later that Fichte has some things to say about such views. Intelligible fatalist views are associated with Schmid, who purportedly thought that Kantian morality meant that the moral law acted as a sort of noumenal (hence, intelligible) cause on us to act or not act in conformity with it (hence, fatalism). It is worth pointing out that Fichte might well have regarded Kant's own views as committed to this, because of the necessity of noumenal selves in Kant's picture. This would then function as a good motivation to jettison the very idea of noumenal selves. Kant claims, in the late *Metaphysics of Morals*, that strictly speaking the will is neither free nor unfree, only choice is free. This is just a way of pointing out that an agent can always choose not to conform with the law, even if they recognise that law as binding (because of evil, or weakness of will).

Fichte also signals development in his early reviews of Creuzer's *Sceptical Reflections on the Freedom of the Will* and Gebhard's *On Ethical Goodness as Disinterested Benevolence*. In the former, Fichte notes that Creuzer argues that Kant is mistaken on free will because Kant's theory 'violates the law of logical ground' or the principle of sufficient reason. Fichte goes on to note that Reinhold has already refuted the objection, but not gone far enough, as he has not 'indicated nor overcome the basis of the misunderstanding in question'.[64] This basis for overcoming the misunderstanding turns out to be Fichte's conception of self-activity, the acts of determining by the intelligible I, and its relation to the determinate being of the empirical I. In the Gebhard review, Fichte hints that he has found a way to move beyond Kant's Fact of Reason. He says that in the Kantian system, as opposed to the eudaimonic system of morals:

[64] Fichte (2001a, 292)

'ethical feeling would be the effect of a kind of reason which would, in this function, stand under no condition other than the condition of its own nature (namely: the condition of absolute unity and hence of uniformity) – i.e. an effect of practical reason'.[65]

Fichte goes on to say that practical reason 'can neither be described as a fact nor postulated in consequence of any fact whatsoever; instead, it must be proven. It must be proven *that* reason is practical.'[66] The proof is also said, portentously, to possibly 'also provide the foundation of *all* philosophical knowledge'. Fichte gives a rough outline of how he envisions the argument. It goes: 'The human being is given to consciousness as a unity (as an I). This fact can be explained only by presupposing something in human beings that is simply unconditioned; we must therefore assume that there is within human beings something simply unconditioned. What is simply unconditioned, however, is practical reason'.[67] I do not intend to reconstruct this argument here. The upshot, however, is that we can see Fichte's thought moving decisively in the direction of his mature ideas regarding the primacy of practical reason, and the need to move beyond Kant. Indeed, we can see these reviews as Fichte working out his views in the context of the vigorous debate surrounding Kant's theory. I shall now say something about Fichte's first published work, the *Attempt at a Critique of All Revelation*, because it contains an early statement of a theory of the will.

The most interesting, for our purposes, parts come when Fichte discusses two versions of freedom, once which is wholly negative (transcendental freedom) and one which is positive (empirically free choice). Here we see that Fichte is probably responding to some of the debate which we talked about above.[68] Fichte describes the 'absolutely first expression of freedom through the practical law of reason' as not meaning choice between alternatives at all 'since the law allows us no option but rather commands by necessity'.[69] This first expression of freedom, transcendental freedom, means 'complete liberation from the coercion of natural necessity', and that this law allows no choice, and 'it determines only in *one* way'.[70] However, empirical freedom does allow for choice between being determined by the moral law or being determined by the sensuous drive, and so whilst moral beings like God are not free in the sense of choosing between alternatives, finite rational agents are free in this way. Here Fichte seems to be foreshadowing some of his later commitments. Firstly, that the moral law applies everywhere and that there is only one morally right thing to do in any given situation. Secondly,

[65] Fichte (2001b, 305)
[66] (2001b, 305)
[67] (2001b, 305)
[68] That Fichte is doing so is perhaps hinted when he says 'This analysis…has had also the secondary aim of clearing up some obscurities in the critical philosophy in general and opening a new door by which those who had hitherto been unacquainted with it or had opposed it might enter into it' (32, p22)
[69] Fichte *Attempt* 31, p21
[70] 32, p22

that freedom of choice is essentially bound together with finitude. However, there are some differences. For example, Fichte's mature theory of evil (the subject of chapter 5) differs from the idea that we choose to be determined by a sensuous impulse (though Fichte's mature account of evil can incorporate this). Secondly, Fichte's mature account involves a key distinction between formal freedom and material freedom, which is absent from the *Attempt*. However, the *Attempt* does highlight Fichte's ongoing commitment to the reality of freedom and the primacy of practical reason in explaining this reality.[71] It is to this doctrine of Fichte's that I now turn.

2) The Primacy of Practical Reason

The purpose of this chapter is twofold. The first is to outline the transcendental argument against determinism, and why Fichte would subscribe to the conclusion of this argument: namely, that determinism is a self-defeating belief. The second aim is to show Fichte's argument for the primacy of practical reason, which establishes Fichte's right to do philosophy according to the idealistic method. It is Fichte's conviction that we could only prove the reality of free will once we orient ourselves with practical reason as primary over theoretical reason, but this needs arguing itself first of all.

2.1) Idealism versus Dogmatism

The transcendental argument against determinism (TAAD), is a cluster of arguments which purport to show that determinism can be proved to be self-defeating.[72] I intend to show here that Fichte is committed to such a position, and that this position can help us to get a grip on the major argument of the later Jena *Wissenschaftslehre*, the *WLnm*. I shall sketch out what I take to be Fichte's view, and then compare it with some more recent formulations. Fichte's strategy in the introductions to the *WL* is to set up the terrain between his own idealism and the major opponent, dogmatism. These are all-encompassing philosophical and metaphilosophical views, and therefore one might be either an idealist or a dogmatist; there is no middle ground. The most fundamental issue between the systems is the correct starting point of philosophy. For Idealism, it is the I, or the self. For dogmatism, it is the thing – the

[71] There is a wealth of material devoted to the development of German idealism. Any list of these should include at least Ameriks (2000) Beiser (1987, 2002) Foerster (2012) Franks (2005), Pinkard (2002). It is however outside the purview of this thesis to recount or reconstruct that development.
[72] Variants of the argument are given by Haldane (1927), McTaggart (1934), Lucas (1961, 1970) Wick (1964), Hasker (1973), Plantinga (2000), Slagle (2016), Lockie (2018).

thing-in-itself, as Fichte says.[73] However, one issue that comes to the fore, and one that certainly occupied Fichte, is not with the starting points of the systems, but with the results of those systems. Fichte thinks that dogmatism not only characterises the clear majority of other philosophical systems (apart from Kant – only on Fichte's rather unorthodox reading of Kant – and Beck), but also that Spinoza represents the best exposition of dogmatism.[74] Spinoza's philosophy ends up with a denial of some of the basic and most certainly known facts of life – that we are selves, that selves are free, and that we have veridical experience. For Fichte, the dogmatist, if they are consistent, denies these – claiming that self-hood is an illusion, freedom is unreal, and that our ordinary experience of the world does not reflect the reality of things. This is, Fichte thinks, an intolerable position. Fichte appeals to the extreme unintuiveness of such a position, and argues that only idealism can provide the security of selfhood, freedom, veridical experience, and, perhaps most importantly, morality.

Thus, Fichte's move can be seen to be something of an appeal to intuition regarding the results of a system of dogmatism. Intuition, Fichte thinks, is on idealism's side. Idealism has never been consistently or fully carried out, so Fichte says that we should wait and see whether it can be carried out, and whether it can deliver on the promises of its intuitiveness to safeguard ordinary thought in a way that consistent dogmatism cannot. We must take care, however, when reading these passages. Though Fichte speaks of 'idealism' in contradistinction to 'dogmatism' he does not regard 'dogmatism' as straightforwardly identical with 'realism' (whatever one makes of such a view). There are, we find out, dogmatist idealists (such as Berkeley), and dogmatist realists. Dogmatism is not straightforwardly a view about the existence or non-existence of the external world, or even the independence of the world from the mind, but an account of both the starting point of philosophical reflection and a view of what Martin calls 'objective reference'.[75] The dogmatist is fundamentally committed to the view that the principle of causality is 'a sufficient explanatory resource for the theory of objective consciousness'.[76] Understanding this point is crucial, because the claim that idealism has greater intuitive appeal than realism is, to both Fichte and our contemporary ears, implausible. Rather, the intuitive appeal of idealism is as opposed to dogmatism, which has as determinate forms of itself some forms of realism. But the intuitive appeal of dogmatism, especially when one considers the end results, is, Fichte contends, lacking.

---

[73] Fichte, '[First] Introduction' I:427-8, 12-3. It must be noted that Fichte probably does not have the Kantian conception in mind. Fichte's imagined opponent is someone like Spinoza, someone who we would now identify as a naturalist of sorts. Fichte wants to designate all transcendental realists as dogmatists, and Kant *under one interpretation* - one that Fichte does not favour. Fichte uses the locution to highlight that it is about independent objects that the dogmatist is concerned about.

[74] There are questions, of course, about Fichte's interpretation of Spinoza, and how much his reading was influenced by Jacobi and the "Pantheism controversy", but I leave these aside for now.

[75] Martin (1997) 18

[76] Martin, (1997) 41

The appeal to intuition, however, is not sufficient on its own.[77] What Fichte wants is for the audience of the debate between idealism and dogmatism to, as he says: 'Attend to yourself; turn your gaze from everything surrounding you'.[78] Fichte claims that if we make a serious attempt at this, to think ourselves alone, then we will come to see that a theory that tells us that we ourselves are illusions, such as Spinoza's, is not the philosophical theory that could make sense of how we find ourselves in the world.

Thus stated, Fichte's move is simple – he asks the would-be dogmatist whether they are content with a philosophical system that tells them that everything (or almost everything) they think they know is illusory, including, crucially, their own self-conceptions and conceptions of what is moral. Of course, this will not convince the committed dogmatist. Those who already sign up to such a program, and are therefore either aware of these results but uncaring, or aware and think of them as *virtues* are not going to be swayed by Fichte's appeal to intuition.[79] I shall look further at what Fichte might say to the committed dogmatist in the next section of this chapter. But for now, it is enough to note that the dogmatist is committed to determinism, naturalism, and in general, focused on getting philosophical explanations out of objects and causal relationships between different kinds of objects.

2.2) The Transcendental Argument Against Determinism

Fichte argues that a correct account of consciousness would have to answer the question of how I know I am different from other objects. In other words, Fichte thinks that a good account of consciousness needs to take into account the fact that self-conscious thought picks out something distinct and determinate. If we had no account of the way in which this works, then we would not be able to distinguish where I end and where nature begins. We can think of the problem that Fichte sets out to address first as the problem of how selves as a distinct kind of thing are individuated. The key to this is found in action, but we need more background to Fichte's argument first.

One variant of the TAAD is given by Fichte in the Second Introduction to the *Wissenschaftslehre*. This argument which comes toward the end of Fichte's exhibition of the

[77] For example, G. E. Moore (in his review of one of the Kroeger translations of Fichte), (Oct. 1898) takes the view that dogmatism and idealism are equal, no matter what. Moore says that Fichte 'has only to say on his side that his presupposition enables him "actually to construct a philosophy", and "has intuition on its side;" both which advantages may be equally claimed by the other side' (96-7).
[78] '[First] Introduction' I:422, 7
[79] A clear and recent account of Fichte on how the change from dogmatist to idealist comes about is Kemp (2017)

dilemma between idealism and dogmatism, states that the dogmatist is contradictory in presupposing mechanism. He says:

'What they say stands in contradiction with what they do; for, to the extent that they *presuppose* mechanism, they are the same time elevate themselves above it. Their own act of thinking of this relationship is an act that lies outside the realm of mechanical determinism. Mechanism cannot grasp itself, precisely because it is mechanism. Only a free consciousness is able to grasp itself'.[80]

This seems to be a bad argument. It clearly begs the question against the dogmatist.[81] The argument, in effect, concludes that determinism (which is the only game in town for dogmatism) undermines itself because it uses a free act to claim that there are no free acts. But of course, whether there are free acts is exactly what is in question. Fichte, however, shows cognisance of this. He says that this is 'just where the difficulty arises, because this is an observation that lies completely outside of their field of vision', and furthermore that it 'must necessarily be incomprehensible to them'.[82] So, Fichte would be begging the question if he intended the argument to refute the dogmatist, but he does not do so; instead, he intends the argument for 'others who can see and who stand watch'.[83] Fichte is aware of the type of argumentative strategy used in the TAAD, but refrains from deploying it against the dogmatist, insofar as he recognises that the actual argument he runs begs the question against the dogmatist.

However, there are resources, I think, in the texts of the later Jena period, namely the *WLnm* and the *System of Ethics*, to construct a better version of the argument.

Fichte begins by claiming that consciousness rests upon the distinction between what is subjective and what is objective, and that the goal of philosophy (or at least, one goal) is to show how the unification is possible. One way in which unification is possible is if subjective becomes like objective (representation, or in the words of the GWL, objective activity), and the other is if objective becomes like the subjective, in action.[84] Fichte claims that this unobjectionable distinction between the subjective and the objective can only be made if the efficacious activity of the self is recognised.[85] The only condition under which I can distinguish myself from the world, as a representor of that world and not merely as a consequence of the

[80] Fichte, 'Second Introduction' I:510,95 emphasis original.
[81] It may well be true, at least from Fichte's perspective, that all variants of the argument beg the question in this way. I do not want to adjudicate here whether, if this is Fichte's view, it is correct.
[82] 'Second Introduction' I:510, 95
[83] 'Second Introduction' I:510, 95
[84] *SE* IV:1, 7.
[85] Guyer (2015, 142)

world is if I recognise that representation is a type of activity. As Guyer says 'Fichte's thought is that if I were entirely passive, just a mirror for changes going on in an entirely objective world, I would have no way of distinguishing myself from that world…'[86] It is 'only by conceiving of itself as acting on the world that the self can distinguish itself from the world'.[87]

The upshot is that to think of one's self as entirely determined by the world is to think of oneself as not a self at all. The argument is often taken, as we shall see below, to establish that to believe determinism is to believe that there are causally sufficient conditions for any given belief, which make any justifying reasons epiphenomenal to that belief. Fichte's addition to the TAAD is to claim that to think of oneself that one is fully determined by the natural world is to deny the distinction (at least, any deep distinction) between the objective and the subjective, which is a condition of making the judgement that one is wholly part of the natural world in the first place, because this is a condition of the possibility of any consciousness and therefore any given belief.[88] We can see here that Fichte is committing himself to the principle that Kant gives in the third chapter of the Groundwork, that reason must look upon itself as the author of its own principles, independently of alien influences.[89] On this line of thought, Wood says that the dogmatist has 'faith...adopted wilfully and contrary to experience'.[90] Acting (and indeed judging) requires myself to think of myself as free, as 'even as a theorist I must take it insofar as I deliberate about what hypotheses to test, how to test them, and what conclusions to draw from the evidence'.[91] It is fitting that Wood mentions deliberating about what hypotheses I should test, because *that* is exactly what Fichte does. In following the transcendental method in Kant, which in turn is modelled on the scientific method, Fichte deliberates what hypothesis to test, and that is something that is only possible because we should take what Idealism holds seriously.

Fichte goes on to argue that the activity must be governed by a law, which cannot itself come from nature, otherwise the agent would merely be, albeit in a more sophisticated manner, a mirror.[92] But we should consider the value of the TAAD first. It seems to me that Fichte is taking a TAAD-like step as his first step in the *WLnm* and *System of Ethics*, and that this can function by way of a more complex motivation for idealism as opposed to dogmatism. Fichte's version of the TAAD is that belief in the truth of determinism is self-defeating because such a belief would mean a denial of the necessary conditions for the possibility of the belief being valuable *qua* judgement of the true, as opposed to, for example, *qua* production of a

[86] Guyer (2015, 142)
[87] Guyer (2015, 142)
[88] *SE*, IV:5, 11
[89] See Kant *G* 4:448, 96
[90] Wood (2000, 100)
[91] Wood (2000, 100)
[92] We shall see below that this makes an appearance as the pure will.

mechanism which reliably increases reproductive success. This necessary condition is the direction of fit from subjective to objective. I shall now look more closely at canonical formulations of the argument and criticism of it.

Hasker takes the key claim in the argument to be 'For a person to be rationally justified in accepting a conclusion, his awareness of reasons for the conclusion must be a necessary condition of his accepting it'.[93] So, he asks, if my awareness of these reasons is not necessary for the belief, then that is saying 'that I would believe just the same *whether or not* I knew of any supporting evidence – and how could a belief held in this manner be rationally justified?'[94] Determinism then claims that 'my acceptance of any conclusion has its sufficient explanation in a chain of physical causes which determine the state of my nervous system'.[95] So, determinism claims that the awareness of the reasons is irrelevant, because the sufficient condition for any belief is given by causally prior antecedents. The awareness of the reasons is therefore not necessary, and so the belief cannot be rationally justified.

We can phrase this in terms of epistemological luck, as Allison does, in one version of the argument he presents.[96] According to the epistemological luck argument, 'if materialism were true, then, although we might still have true beliefs, and even good reasons for these beliefs, the contingency of such a state of affairs would be enough to prevent these beliefs from counting as knowledge'.[97] For all we know, Allison continues, we might, like badly constructed machines, 'be wired to churn out false "beliefs" without any possibility of correcting them'.[98] This last way of putting the point shows the obvious flaw of this version of the TAAD. We can account for the so-called 'luck' by merely explaining it as a result of the evolutionary process.[99] Allison tells us that what he terms the 'Recognition argument' is the stronger variant of TAAD. This is broadly the argument used by Hasker. In Allison's words 'if, as materialism assumes, there are causally sufficient conditions for my belief that *p*, then my reasons (whether or not they are good, i.e., justifying, reasons) are not necessary conditions'.[100] This provides a nice counter to the evolutionary challenge to the epistemological luck argument. That challenge claimed that we could make sense of the truth-tracking sense of our beliefs by virtue of certain applications of the theory of natural selection. But the problem now becomes the following: if that is the case, then it seems as though any value our beliefs may have *qua* states of mind that aim at truth and not *qua* products of well-functioning and well-adapted mechanisms that

[93] Hasker (1973, 178)
[94] Hasker (1973, 178)
[95] Hasker (1973, 177)
[96] Allison (1996, pp.92-106, esp. 99-100). Whilst Allison speaks of materialism, it is fair to say, I think, that Kant and Fichte would have regarded the one as entailing the other, and the transcendental argument in question here can be made to work against both.
[97] Allison, (1996, 99)
[98] Allison, (1996, 100)
[99] Such an explanation might take the form of the view defended by Millikan (1984).
[100] Allison, (1996, 100)

increase reproductive success seems to be entirely epiphenomenal. *That* is what is so problematic.

The upshot of this is that an argument that purports to show the truth of determinism necessarily fails because belief in the conclusion of the argument has causally sufficient conditions, which do not mention the reasons or awareness thereof. So, determinism cannot be rationally believed to be true.[101]

One general criticism is the criticism raised against all transcendental arguments, made famous by Stroud.[102] Stroud argues that at most, a transcendental argument can give merely what we must think to be the case in order to be coherent, but not what actually is the case. This criticism is sharpened if we consider the Quinean view that any belief can be held come what may as long as we are able and willing to make radical enough changes elsewhere in our system of belief.[103] If we have to believe something only to be coherent, and we can accommodate the lack of this belief by making other changes which also maintain coherence, then it looks like transcendental arguments, and *a forteriori*, the TAAD, are either invalid or unsound.

The obvious response to this challenge is that what is sought by the argument is not that determinism is false, because to establish claims of that nature, of the 'inner nature of things' is not the goal of transcendental philosophy. What the argument claims to establish, and, if sound, would establish, is that there is something self-defeating in belief in determinism. It is surely possible that determinism is in some way true, and Fichte would want to uphold that this is a possibility, but what the TAAD purports to show is not that it is false (either necessarily or actually), but that it cannot be coherently believed, even if we take into account the Quinean point that we can in principle accommodate any belief within the system of beliefs.[104] Another way to get to this same point is to respond on Fichte's behalf by saying that this Stroudian response actually begs the question against the idealist – it assumes that there is a principled

[101] Lucas occasionally states the conclusion much more strongly than this, that it shows that determinism is necessarily false. But that is because his version of TAAD is based upon considerations arising from the correct application of Gödel's incompleteness theorem to a logistic system that models a perfectly determinate human being. Examining this particular argument in detail would take me too far afield. See Lucas (1970) sections 21-29, esp. 24-26. Though, having said that, he admits at the end that his argument fails if it can be shown that the physical system is essentially inconsistent, but that this is 'complete conceptual nihilism' (163), because it would mean that human beings are fundamentally inconsistent in some way, as this is the only way to escape the Gödelian argument. But see Dennett (1981) 256-266 for a critical response.

[102] Stroud (1968)

[103] This is found in Quine's (1980), esp. 42f.

[104] Presumably these Quinean considerations would lead one to the state of complete conceptual nihilism that Lucas wants to avoid, it would be like the cognitive equivalent of Ivan Karamazov's famous line in The Brothers Karamazov that since God does not exist everything is permitted; so the conceptual nihilist will say that since all the norms of belief are flexible everything is permitted.

and accessible distinction between the way things really are and our access to those things, something that Fichte denies.

One argument that is raised more specifically against the TAAD can be found in Armstrong. Armstrong says that we should just identify the causes of my belief with my reasons for holding that belief. He claims that 'the premiss of my inference must be the cause (in the appropriate way) of the belief that I arrive at'[105] and in so identifying the reason (in this case a premise in an inference) we actually guarantee at least some reasonableness of one's beliefs, because they can be rationally explained, or as he puts it 'causality is in fact a logical precondition of the possibility of inferring'.[106] There are two basic strategies for defending the TAAD here. One is to claim that the type of explanation in why one believes as one does is not a causal explanation. A variant of this strategy is found in Wick. Wick claims that when we ask questions of the reasonableness of belief, we never want to inquire as to the generation of the belief-state, instead, our question concerns the content of the belief.[107] The former, he says, would only make sense if we asked a question of the form 'why do you know that', which he says seems to imply that we regard some sort of malfunction of a psychological mechanism has occurred.[108] Wick's position is to try to show that the type of explanation we seek in explanation of why someone holds a certain belief just is not amenable to causal explanation in the way Armstrong requires.

The other way of responding to this challenge is to try to show that this identification may be fine, but then the determinist has merely pushed back his self-defeat one step further. This is shown in the following example. If the determinist accepts the identification of reasons with causes, or at least, the right kind of causes, then presumably there is going to be a neural structure of some kind which realises the content of the belief. There is then one item in question with neural properties and content properties. But the question is now that it seems that the determinist wants to say that the item in question is causally efficacious by virtue of the content properties. But this is, as Plantinga points out, 'extremely difficult'.[109] This leaves the determinist in the position of a semantic epiphenomenalist, which would equally establish the conclusion of the TAAD. The argument still stands in the face of this objection. Recall, however, that the argument never pretends to establish the falsity of determinism, merely the

[105] Armstrong (1993, 200)
[106] Armstrong (1993, 200)
[107] Wick (1964, 533)
[108] Wick (1964, 532). Harman (1970) claims that explanation by reasons is not explanation by causes (849). For Harman, we can conceive of a psychologistically indeterminate thing instantiated in a physically deterministic thing, and reasons may be identified with causes (851), but the explanatory work is not done by the causes. This is one of the central thoughts to the strategy that is taken here.
[109] Plantinga (2004, 603) Plantinga makes his remark in the connection of beliefs being causally efficacious in the production of action, but I think that the case can be assimilated to the case of the production of beliefs.

self-defeating nature of the claim. One more criticism must be looked at, one which is particularly pertinent to Fichte – that of the argument begging the question.

This objection is formulated best by Churchland.[110] Churchland argues against one specific instance of the TAAD, found in Popper. She claims that the argument presented assumes the truth of what it tries to prove, and thus is question begging. A supposedly parallel argument from a biological vitalist (one who believes that life is essentially different from non-life) is designed to show this. The vitalist argues that the anti-vitalist cannot be right because in order to be right they have to have vital spirit, which is just what they deny.[111] Similarly, the indeterminist is thought to argue that the determinist cannot be right because they use their freedom to claim or think that they are not free. Notwithstanding whether or not this argument works against the variant of the TAAD found in Popper (Churchland's immediate target), it cannot be said to work for the version that Hasker or Allison present. This is because the Recognition Argument does not assume the truth of what it tries to prove, but assumes the truth of determinism instead. Fichte's argument in the Second Introduction seems to indeed beg the question in this way, though Fichte is, as shown by what he says after stating the argument, aware of that. This means that there must be some other reason that he states it, especially as the argument appears in the second introduction, which is directed to those who already follow a philosophical system. I suggest that as well Fichte is using the TAAD not as an argument *per se* but as an apt occasion to invite the dogmatist to reflect. Only that will be sufficient for the dogmatist to change their minds. This answers any questions we may have regarding the exegetical worth of attributing this argument to Fichte. As we have seen, Fichte claims that neither dogmatism nor idealism can be proved or refuted. My reading here preserves that claim. But we should remember that the TAAD will not argue that determinism can be shown to be false, but that it is self-defeating, or in Lucas' phrase, ends in 'complete conceptual nihilism', albeit only from Idealism's vantage point.

But there is still the issue of how this comes to be an argument against the dogmatist. Wood views Fichte as giving a TAAD-like argument but emphasises that Fichte's insistence on faith shows that Fichte never meant this to be a conclusive anti-dogmatist argument. Faith in this context does not mean anything of a question that cannot be decided by reason or belief against evidence.[112] What it does mean is that the idealist is always aware that they can only reasonably engage with the dogmatist as the idealist conceives them, and not as they may actually be. As Wood says, 'there is nothing in the grounds for this conviction that dogmatists would have to acknowledge as providing *them* with a convincing reply to the thought that the

[110] Churchland (1981)
[111] Churchland (1981, 100)
[112] See for example, Wood (2016, 72)

appearance of freedom is only an illusion'.[113] What this means is not that the dogmatist could never change their mind at all, but that they must be the ones to change their own minds. This is true to some extent for all philosophical changes – I may indeed change my mind based on an argument that you present to me, but I would first have to consider the argument myself, independently of your presenting it. In the present case, however, the choice between a dogmatic and an idealist philosophical outlook is inextricably linked with selfhood and personal choice, such that it becomes a more complex change. Not only do I have to convince myself of a view that I was previously not convinced of, but I have to perform some kind of act of will in order to fully realise the view. This latter is why this change of heart is different from others. One can present me with an argument that I later assent to on the basis of my own thought, but one cannot present me with an act of will. I need to perform such a feat myself. Choosing idealism or dogmatism is therefore akin to what L. A. Paul (2014, 2015) has recently called 'Transformative experience', insofar as the choice can only be vindicated from a post-choice vantage point. I would suggest that this may be one reason why even sophisticated presentations of the TAAD might well end up falling prey to criticisms that they beg the question (though that is speculative, and I will not attempt to adjudicate that issue here).

There only remains the question of the advantage in explanatory power that idealism is supposed to have over dogmatism. Because Fichte thinks that there are only two first principles available to any philosopher (the thing-in-itself and the I-in-itself), there are only two philosophical systems possible. As stated above, it is clearly question-begging for the idealist to assert that the I-in-itself explains something that the dogmatist cannot, because on the dogmatist's picture, the thing (freedom, or the self) does not need explaining, but explaining *away*. But one might appeal to considerations of faithfulness to experience (or intuition), and point out that the feeling of ownership of actions, *inter alia*, is a desideratum for a theory of the self, and any theory that can sufficiently accommodate this feeling without turning it into a sort of delusion or falsehood has the advantage, *ceteris paribus*. Fichte would probably say that the dogmatist's view of the self as a kind of delusion reflects poorly on the fact that they reify the self – in effect, if a theory tells us that the most certainly known facts about ourselves are falsehoods, then so much the worse for that theory. We have now seen Fichte's argument for idealism both directed toward the beginner and the convinced dogmatist. Fichte's view is that there are reasons to pursue the idealist programme. A major feature of this programme, one which is certainly in the background of the TAAD and the quarrel with the dogmatist is the primacy of the practical. It is to this that I now turn.

[113] Wood (2016, 76)

2.3) The Different Senses of the Primacy of The Practical

Commentators on Fichte are agreed that Fichte does in some sense affirm the primacy the practical, or the primacy of practical reason at least in the Jena period, which is my focus here. Disagreement revolves around the question of 1) to what sense Fichte is so committed, 2) to what extent he is committed, and 3) what philosophical work PPR is supposed to do. There are also two separate doctrines of Fichte that one might wish to designate by the phrase 'the primacy of the practical'. There are also several senses of PPR, some of which can be jointly affirmed. It is necessary both to disambiguate the two separate doctrines, though they are connected, and to disambiguate the various senses of PPR. I start with the former task.

The 'primacy of the practical' can mean a thesis in philosophy of mind regarding the structure of the mind, or the ordering of our mental faculties, which is PPR. It is important to dwell for a moment on Fichte's conception of a mental faculty. They are (as in Kant) not separate "institutions" relating to one another, but only different applications of the same power.[114] For Fichte, it is very important that we do not reify or objectify these faculties. They are not things, but acts. It is helpful to consider the context in which Fichte was developing this view. Aenesidemus had raised sceptical objections to Kant and Reinhold and their putative usage of an inference from the application of a power of X-ing to a thing's having the power to X. In Reinhold this came to the fore in the guise of the faculty of representation.[115] Aenesidemus raises the challenge: how is this inference different from the dogmatic inference from one's having an idea of an object to there being an object in the world?

In order to avoid this challenge, Fichte conceives a middle path. On the one side, we have the realists, represented by Reinhold, who will claim that the inference from a power to a thing that realises the power is legitimate. On the other, we have phenomenalists like Aenesidemus, who claim that a faculty is nothing other than its manifestations. The crucial move that Fichte makes, as I see it, is to agree with the phenomenalist that the faculty is nothing over and above its manifestations, but then argue that the phenomenalists counterfactual – 'were there to be no manifestations, there would be no faculty' is true, but never satisfied, because the antecedent is always false. This gives Fichte a conception which is at once non-reifying but also non-phenomenalist.[116] This is hinted at when Fichte claims that if one thinks that I have to exist in order to do things is 'maintaining that the I exists independently of its actions', which

[114] Kant says that 'there can, in the end, be only one and the same reason, which must be distinguished merely in its application'. *G* 4:391, 47.
[115] As developed in his (2011)
[116] See also Breazeale (2013) 408ff. This is also Henrich's point when he says that for Fichte, dreams are 'one of the elementary experiences of freedom' (because the dream is a result of the activity of the I, even when we might ordinarily think of ourselves as inactive), (2003, 213).

is to treat consciousness as a type of object.[117] Thus a claim about the relative primacy or non-primacy of one faculty to another is regarding the fundamentality of one or other of the manifestations of the powers of the I.

The primacy of the practical can also mean the ontological thesis that is most explicit in the essay that appears at the close of Fichte's time at Jena in the essay *On the origin of our belief in a Divine Governance of the World-Order.* The thesis appears in that work when Fichte famously says the 'only kind of reality that pertains to your or exists for you…is the ongoing interpretation of what your duty commands, the living expression of *what* you ought to do, just because you ought to do it. Our world is the material of our duty made sensible. This is the truly real element in things, the true, basic stuff of appearance'.[118] As I said, this is the most explicit, and striking, presentation of the thesis.[119] But Fichte has also affirmed this thesis in the *System of Ethics,* in which it again takes a more assertive tone, as he says '…all of the I's cognition is determined by its practical being – as indeed it has to be, since this is what is highest in the I. The only firm and final foundation of all my cognitions is my duty. This is the intelligible "in itself", which transforms itself by means of the laws of sensible representation into a sensible world'.[120] This statement actually contains both theses that could be thought of under the rubric of primacy of the practical. The first part of the passage, i.e. the phrasing 'since this [practical being] is what is highest in the I' is the thesis that practical reason, or something about our natures as active or practical beings, is more fundamentally constitutive of our essence. The later part of the passage is very similar to the phrasing of the DGWO essay.

This position also seems to survive the atheism controversy, as we see in Schopenhauer's lecture notes from Fichte's 1811-2 Berlin lectures – 'How generally is a knowledge possible? For knowledge is intuitive perception of the absolute. The absolute becomes visible through the action of the I which indicates that underlying it there is some clear concept of arrangement which is directly accompanied by approval through the law; the feeling of this approval is in ordinary language called conscience. Therefore is there an intuitive perception of true being? Does such a perception enter into knowledge? Yes, by means of action according to the moral law. *Therefore there is nowhere any truth and reality except in moral action*'.[121] It must be said, however, that the correct interpretation of these lectures is controversial and will depend on many factors which I will not discuss here. *Prima facie*, however, it seems as though this ontological sense of the primacy of the practical is a view that Fichte is committed to

[117] *FTP,* K29/H29, 112.
[118] *DG* 185, 150
[119] This is certainly one of the prominent passages that is often cited in characterising Fichte as an 'ethical idealist'.
[120] *SE* IV:172, 164
[121] Schopenhauer (1988, 37, 83). Emphasis original

throughout his philosophical career. I will leave discussion of this thesis here, as I am primarily concerned with the correct interpretation of the view regarding the order of our faculties.

As Fichte is a transcendental philosopher, it is plausible that any claim to primacy will turn out to be a transcendental one. We should take a moment here to review what a transcendental version of the theses would be. With the ontological thesis, a transcendental interpretation would appear to be something like this: It is a necessary condition of the possibility of the natural world that there be moral properties (oughtness). For PPR, the thesis would be: it is a necessary condition of the possibility of our having theoretical reason that we have practical reason. There are other senses, such as a temporal priority, thinkability priority, legislative priority, and methodological priority. I shall examine each sense in turn.

Perhaps the most obvious sense of PPR is the temporal sense – that practical reason is primary because it appears at an earlier stage in a thinker's lifespan than theoretical reason. Whether or not this is true is plausibly a matter for empirical psychology. Fichte is not interested in this sense of PPR. He is interested in transcendental conditions for the possibility of self-consciousness, not the development of a person's mental life. In any event, it seems clear that Fichte thinks that there is no time at which a rational agent or a person is manifesting the practical power but not the theoretical power, as this mean there would be a radical distinction or split between the practical power and the theoretical power. But Fichte agrees with Kant that reason must be in a unity, differentiated only by its applications. So we should move on to consider other forms of priority.

It appears that Fichte thinks that it is because we have a practical power of reason that we have theoretical reason. In other words, we are not primarily orientated to the world as knowers, but actors. As Kinlaw says, 'theoretical reason is subordinate precisely because theoretical knowledge of the world is simply one way – and not the primary way – of relating to the world'.[122] Fichte says that whilst his investigation, the 1794/5 Wissenschaftslehre, begins with the exposition of theoretical reason and then proceeds to an exposition of practical reason, this is not a reflection of their respective positions in our mental architecture. He says 'reflection must set out from the theoretical part; though it will appear in the sequel that it is not in fact the theoretical faculty which makes possible the practical, but on the contrary, the practical which first makes possible the theoretical'.[123] Firstly, this is a pretty clear statement

[122] Kinlaw (2002, 142). Interestingly, Kinlaw thinks that this position commits Fichte to the ontological thesis of the primacy of the practical – that the moral world is 'more real' (as he puts it) than the natural world (154n4). I am not sure that such an entailment holds, though the positions are no doubt related in some way.

[123] *SK* I:126, 123 All reference to this text have the translation amended from 'self' to 'I', as it is both less misleading and a better translation of Fichte's words (*Das Ich)*. Also, in the first (and only published) chapter of the *Attempt at a new presentation of the Wissenschaftslehre*, Fichte explicitly says that although the word 'self' 'presupposes the concept of the I, and everything that is thought to be absolute within the former is borrowed from the concept of the latter', that 'it seems to me that in a scientific exposition one should employ the term that designates this concept in the most immediate

of an unequal relationship between the practical and theoretical faculties. Secondly, it appears this is a transcendental claim, as he says that the practical faculty *makes possible,* that is, is a necessary condition for, the theoretical faculty. But straight after this, Fichte says that the theoretical must come first in the order of explanation because the thinkability of the practical depends on that of the theoretical.[124] It is, however, unclear as to what exactly he intends to mean here. Obviously, there is one sense in which the thinkability of anything depends on the theoretical power of reason – theoretical reason, in the broadest sense, just is the ability to think things, including reason and its powers. Though this is not true for Kant, who has a distinction between various theoretical powers – reason, the understanding, and judgement, it is not clear that Fichte wishes to retain this division. Or, if he does, it is downstream from Fichte's concerns here, I think.[125] But this is true with or without a philosophical investigation of the fundamental principles of theoretical reason. Merely by having theoretical reason would this be the case.

There is another, less literal, sense of how one thing can make another 'thinkable'. This is the philosophical sense in play whenever it is asserted that an account of X depends on a prior account of Y. The description and explanation of Y give the conceptual tools necessary to give an account of X. Here the account of Y would provide the thinkability of X. For Fichte, then, if this is what he means, theoretical reason must come first in the order of philosophical explanation because a correct account of the nature of theoretical reason and the I's acts insofar as they are theoretical provides the correct conceptual framework for analysing practical reason and the I's acts insofar as they are practical.

A more specific, and characteristically Fichtean, way we might cash out the thinkability dependence in this way is the following. In philosophy, according to Fichte (at least in the early Jena WL), it is necessary to start with absolutely certain first principles and then work from there.[126] It may be that the thinkability priority is merely a re-statement of this view. The highest principle of Fichte's system is the I's self-assertion or self-positing – this is the unconditioned

---

and proper way'. See 'Chapter One: All consciousness is conditioned by our immediate consciousness of ourselves', in *Introductions and other Writings*, I:530n, 115n.

[124] *SK* I:127, 123

[125] See Fichte's remarks at *SK*, I:282, 248ff where he seems to claim a very broad sense of theoretical reason. The marks for something's being related to the theoretical faculty are that the thing is 'subjected to its [theoretical reason's] laws of presentation' and that in the theoretical part of the *WL*, we are 'concerned with knowing' *SK* I:285, 251. See also Breazeale's remark (2003, 257), that Fichte is not trying to 'reinvent the wheel' with regard to Kant. It is therefore possible that Fichte does not feel as though he needs to, at this fundamental stage in the *WL*, to divide the theoretical faculties into more fine-grained powers, as Kant does, even though he accepts that division at some more determinate level.

[126] There is an ongoing controversy in Fichte scholarship about the extent to which Fichte is committed to something like foundationalism or coherentism in epistemology. I cannot go into the details of this controversy as it has developed, but the key essays for it are Breazeale (1996) (who takes Fichte to deny foundationalism, in the most obvious sense), Rockmore (1996) (who takes Fichte to be a foundationalist).

and absolutely necessary act.[127] Fichte begins the theoretical portion of the 1794 *Wissenschaftslehre* by giving this principle. It would be unphilosophical to begin with anything else. So the thinkability of practical reason depends on the thinkability of theoretical reason because that is the path that philosophical reasoning must follow.

A parenthetical remark we can draw on from the same page complicates the issue. Fichte says that 'reason in itself is purely practical, and only becomes theoretical on application of its laws to a not-I that restricts it'.[128] Here it is necessary to point out that Fichte's overall project at this stage is to exhibit the consequences of the highest principle of philosophy – that 'I am I'.[129] The not-I is a principle that Fichte draws out. The I determines itself, and in doing so, posits something which is completely opposed to it, a not-I.[130] Fichte's claim regarding the relative priority of theoretical and practical reason here is therefore that theoretical reason is only a relational property of reason itself, and it only has this relational property in virtue of positing a not-I. But the positing of a not-I is a necessary act of the I.[131] Reason does not, then, become theoretical, it is always theoretical. This is because reason, the I, always stands opposed to a not-I. There is no time at which reason is not opposed to a not-I. We can say, as a way of making sense of Fichte's parenthetical remark, therefore, that theoretical reason is a necessary, though still fundamentally relational power of reason, whereas practical reason is a necessary and essential (or intrinsic) power of reason. But this is compatible with Fichte's claim that theoretical reason must come first in a philosophical account of reason.

We then see that there are two senses of PPR (temporal and thinkability) that Fichte denies. However, he affirms a transcendental primacy of practical reason – it is a necessary condition for theoretical reason. We can give some textual support for distinguishing necessary and essential. In the System of Ethics, Fichte says: 'The theoretical powers pursue their own course until they hit upon something that can be approved. They do not, however, contain within themselves any criterion for the correctness of the latter; instead, this criterion lies in the practical power, which is what is primacy and highest in human beings and constitutes their true essence'.[132] According to what Fichte says here, practical reason is a rational agent's true essence, but he does not commit himself to any statement regarding the non-necessity of theoretical reason. Indeed, throughout the section from which this passage is taken, Fichte reaffirms that theoretical reason is necessary because the practical power by itself has no power of cognition.

---

[127] *SK*, I:95, 96
[128] *SK* I:127, 123
[129] *SK* I:95, 96
[130] *SK* I:97, 123
[131] As will become clearer, the acts of the self in this context are non-temporal transcendental stages.
[132] *SE* IV:166, 157

But the point must still be clarified. Fichte later warns us that it is possible to think of the relation between the I and the not-I in the wrong way, and end up with what he calls 'intelligible fatalism'. This is the position according to which what Fichte calls 'pure activity' relates to an object 'in itself and as such'.[133] Pure activity is the theoretical power of positing oneself and thinking of oneself. – it is the type of thinking where the object of thought and the thinker are identical.[134] For Fichte, in order to guarantee the freedom of the I (which is non-negotiable), the pure activity must remain at some distance from the not-I. This is because, Fichte tells us, if the pure activity were to relate to an object, then it would be finite, because objects are finite. Pure activity is infinite, because the I is infinite. Such is needed to guarantee the freedom of the I. So the theoretical power of the I cannot relate necessarily to a not-I, indeed, it needs a 'special absolute act of connection',[135] which is only given to us by the positing of the not-I, which happens, as Fichte tells us, absolutely, or without any ground; by this I take Fichte to mean that the positing of the not-I is a free act of the I. This also means that the positing of the not-I is an unprovable assumption, that nobody can prove to one another on rational grounds.[136] The fact that I posit a world distinct from myself, or opposed to myself, is a 'fact of consciousness'.[137] So whilst the positing of the Not-I by the I in some sense *must* occur, it is not the case that it is *constrained*.

To sum up, it seems that practical reason is identical with 'reason as such', and theoretical reason is a power that is involved only upon a certain condition being fulfilled, but that condition (i.e. the positing of a not-I) is always fulfilled. But the fact that it is always fulfilled does not mean that it is a directly necessary connection between the positing of the I and the not-I – there are other types of activity that provide that connection.

A fourth sense of 'primacy' comes to the fore when Fichte is reviewing the argument that he makes in section III of the 1794/5 *WL*. The argument, which I shall provide an exegesis of below, is intended to show that we have practical reason, as Fichte thinks that this is something that has only ever been assumed. Fichte thinks that the only way to prove that we have practical reason is to prove that practical reason is the basis of theoretical reason. He says such a proof 'must be carried out agreeably to theoretical reason itself, and the latter should not be ousted from the case by mere decree'.[138] In other words, Fichte is telling us that his thesis of the primacy of practical reason is a *theoretical* one, by which I mean that the

[133] *SK*, I:263n, 232
[134] See for example *SK*, I:96, 97. This seems to me to be related to the 'intelligible fatalism' defended by Schmid and criticised by Fichte in *SE*. For Schmid, our consciousness of the moral law entails that we do our duty, so we are in a sense determined by this to act. In the theoretical sense of intelligible fatalism under discussion here, the view would be that in thinking myself, I think the world as well.
[135] *SK*, I:263, 232
[136] *SK*, I:252-3, 223
[137] *SK*, I:252, 223
[138] *SK*, I:264, 233

thesis is by no means meant to lead to practical considerations *overriding* theoretical ones. So, what we get is this: Theoretical reason maintains the role of the legislator – theoretical reason maintains full jurisdiction over what is acceptable and unacceptable – and what is acceptable is acceptable on rational or theoretical grounds *alone.* Therefore, there is another sense in which Fichte denies PPR – the sense in which practical reason is thought of as the legislator in the domain of theory – which is a quasi-pragmatist thesis. So, Fichte affirms the transcendental sense of PPR, and denies the temporal, thinkability, and legislative senses.

Some passages from the later work seem to be ambiguous on the legislative sense, however, so it requires more attention. For example, let us consider the passage from the *System of Ethics* quoted above, namely: '…all of the I's cognition is determined by its practical being – as indeed it has to be, since this is what is highest in the I. The only firm and final foundation of all my cognitions is my duty.'[139] One issue that arises is that to say that cognition is to be 'determined' by my practical being, and that the 'only firm and final foundation' of my cognition is 'duty' seems to put in jeopardy my hypothesis that for Fichte, theoretical reason maintains the role of legislator, as well as the coherency of Fichte's doctrine. Another issue is that later in *SE*, Fichte appears to say that the practical faculty is not what is highest, but the cognitive faculty is, thus reversing his position earlier in the text. He says: 'What is primary and highest in a human being – though not what is most noble in him – is cognition, the primordial matter of his entire intellectual life.'[140] So it appears that Fichte in one passage claims that the practical faculty is highest, and in another claims that the cognitive faculty is the highest. This, I suggest, is due to the fact that at this point in the SE, Fichte has changed from thinking of transcendental conditions to empirical human beings in situations of empirical willing, theoretical cognition of objects come first, and so present what is 'highest' or 'primary' – as an empirical episode of willing has to be preceded by cognition of its object.[141]

That suffices for a reply to the latter of the two criticisms noted, but what of the former? Firstly, we should note that by 'determine' Fichte by no means intends that practical reason can override theoretical reason. He says that theoretical reason is 'formally' determined insofar as practical reason (or the ethical drive) set theoretical reason in motion to search for the concept required for any action.[142] Theoretical cognition is also determined 'materially' but only insofar as Fichte thinks that all cognition, even the most abstract is 'at least indirectly related to our duties' and carries with it cognition of the objects 'purposiveness' (though not necessarily for *me*) or the objects 'final end'. Whatever we think of this doctrine, i.e. that all cognition is indirectly related to our duties, it makes perfect sense for Fichte. If reality is fundamentally

[139] *SE*, IV:172, 164
[140] *SE* IV:344, 325
[141] *SE* IV:255, 244
[142] *SE* IV:172, 163

moral, as he seems to think, then of course cognition of objects is related to my duty, because the objects that my cognition refers to are my duty made manifest, as he says.

What practical reason does not do is intercede in a theoretical decision. This would happen if, in Fichte's words, practical reason 'provided the material' for theoretical reason, which it cannot do. Confusion may arise regarding 'being materially determined by' and 'providing the material for', but as long as we keep those firmly separated, as Fichte does, we should not be misled into thinking that Fichte thinks that our practical reason in some sense tells us what to believe.

Further support for my view here comes from Fichte's normative views. For Fichte, it would be immoral to let one's practical reason override one's theoretical reason; he says: 'A material subordination of the intellect to the moral law is therefore impossible….Though I may not give in to certain inclinations and pleasures because this runs counter to my duty, it is not the case that I must will not to cognise certain things because this might perhaps run counter to my duty'[143], though cognition is 'formally subordinated' insofar as the highest end is self-sufficiency, or freedom. All this means, however, is that Fichte conceives it as a duty to further our cognition, and to never 'subordinate your theoretical reason as such, but continue to inquire with absolute freedom, without taking into account anything outside your cognition'.[144] So theoretical reason maintains, within its domain, primacy in the legislative sense, though practical reason, by virtue of its transcendental primacy, determines, both formally and materially, theoretical reason to this domain.

An issue connected with this is how to interpret Fichte's view that philosophy must begin with the I and its free act of self-positing because I 'ought' not to begin anywhere else. It is not, he says, that I cannot think further, but that I ought not to. I would suggest that this assertion is a result of what Wood calls the 'requirement that our conception of ourselves and our activity must be a systematic conception, self-consistent, not self-undermining, and capable of being presented in a coherent transcendental system. Such an incoherence is not merely something we (psychologically) *can't* believe; it is something we (normatively) *must not* believe'.[145] Whilst Wood says this in the context of a theory of freedom, I believe it is equally applicable here. Likewise, Breazeale says that I have the ability to doubt the validity of the moral law and question my belief in my freedom, but 'I could do this only if I were willing "to destroy my own inner self" that is, only if I were willing to sacrifice what I hold dearest'.[146] Breazeale calls this a methodological primacy of the practical, as practical reason here provides the warrant for the choice of philosophical starting point.[147] Even though Fichte does seem to claim that there

[143] *SE*, IV:217, 206
[144] *SE* IV:218, 207
[145] Wood (2016, 35)
[146] Breazeale (2003, 255)
[147] Breazeale (2003, 254)

is at least one case in which moral claims or practical claims may override a theoretical claim (inasmuch as I ought not to proceed beyond the free act of the self-positing I), I take it that this represents an extension of, and a further motivation for, his TAAD step. We could call this a methodological primacy of practical reason, as Breazeale does, but only if we keep firmly in mind that once the project (that is, of the system of idealism) has been started, it is a wholly theoretical exercise – which Breazeale also emphasises. By way of an interim summary, Fichte affirms the transcendental and methodological senses of PPR, and denies the temporal, thinkability, and legislative senses.

Consideration of these points leads to a criticism of interpreting Fichte according to my view. Namely, that he insists, especially in the *WLnm*, but also in the earlier *WL*, that thinking and willing, or real activity and ideal activity, or theoretical and practical reason, are necessarily bound together, as we have seen. But having them necessarily bound together does not mean that there cannot be a priority in some other sense.[148] Fichte does claim at numerous points, for example in section 4 of the *WLnm,* but also in other portions of the text, that the I is a unity of practical and theoretical powers. For example, 'The I is neither the intellect nor the practical power; instead, it is both at once…If we want to grasp the I, we have to grasp both of these; separated from each other, they are nothing at all'.[149] Zöller, for example, claims that the ideal and real powers of the subject form a 'structure of intricately related *co-original* moments that collaborate in the constitution of subjectivity'.[150] Similarly, Breazeale claims that accounts that stress Fichte's primacy of practical reason thesis obscure 'what is arguably his single most important accomplishment as a transcendental philosopher: namely, his demonstration of the *inseparability* of knowing and willing, theory and practice, within the original constitution of the I and at every moment of empirical consciousness'.[151] This is the so-called equiprimordiality thesis. Whilst it is clear that Fichte does indeed think that the two are both required for I-hood and empirical consciousness at all, it seems to not take into account the subtlety of his position, and indeed his many other remarks, to say that they are on an equal footing. Indeed, as I have attempted to show above, Fichte can think that they are equally required, but in different senses, thus maintaining a priority.[152] That disparity rested on a distinction between a thing's being necessary, and a thing's being essential, or constitutive. I suggested that it is because theoretical reason is necessary but not essential that Fichte maintains the disparity.

---

[148] As an analogy, something may be necessarily identical with itself considered under some other aspect, and yet the one have some priority over the other. An example is that Water and H20 are commonly considered necessarily identical, but surely H20 has some explanatory priority and is more fundamental. Likewise even if theoretical reason and practical reason are faculties of the very same thing – the I (though of course there is no 'bare particular' of the I 'standing behind' theoretical and practical reason) – practical reason might be more fundamental. See Mann (2015, 34f).
[149] *FTP*, K54/H46, 152
[150] Zöller (1998, 73).
[151] Breazeale (2013, 406)
[152] This is also seen by Neuhouser (1990, 8)

The line that Breazeale, and Zöller take on PPR seems to stem, at least partially, from Aenesidemus. At the very least, there is an interesting parallel with Aenesidemus to be considered. At the end of the review of *Aenesidemus*, Fichte explains that Aenesidemus objects to Kant's moral theology. Fichte says:

'Aenesidemus' protests against this mode of inference are based upon his deficient grasp of the true difference between theoretical and practical philosophy. These protests are summarised approximately in the following syllogism: Until we decide *whether* it is *possible* to do or to refrain from doing something, we cannot judge *that* we are commanded to do it or to refrain from doing it. But whether an action is possible or impossible is something that can be decided only according to theoretical principles. Therefore, even the judgement *that* something is commanded is based upon theoretical principles. That which Kant first infers *from* the command has got to be already shown and decided *before* any command at all can be rationally accepted. It is far from being the case that the recognition of a command can provide the basis for the conviction that the conditions for its fulfilment actually do exist. On the contrary, that recognition can only follow upon this conviction. One can see that Aenesidemus is assailing the actual foundation of Kantian moral theology, namely, *the primacy of practical over theoretical reason*'[153]

To summarise, Aenesidemus objects to Kant's usage of the principle that "ought implies can", because in order to know whether we should do something, we need to know that we can do it, or as Fichte says, that we have the conviction that the conditions for its fulfilment actually obtain. Fichte is probably alluding to Aenesidemus again when he says, in the *WLnm*, that 'One commonly says, "I cannot will unless I first possess a cognition of the object I am willing." This, however, is not true, for there is also another kind of willing, one that provides itself with its own object and to which, therefore, no object is given in advance'.[154] I want to note the similarity in Aenesidumus' criticism and the interpretations of Zöller and Breazeale. Both Aenesidemus on the one hand, and Zöller and Breazeale, on the other, seem to commit the error of thinking that there is only one sort of willing. Zöller and Breazeale claim that practical reason cannot have primacy in any real sense because all the exercises of practical reason require a use of theoretical reason. Fichte claims that this is only true for a restricted sphere – of empirical willing. I therefore submit that Zöller and Breazeale are incorrect to say that there is no significant sense of the primacy of practical reason in Fichte. I shall examine their claims in more detail now.

[153] Fichte, 'Review of *Aenesidemus*' (*RA*) in (1988[1794]) I:21-2, 74-5, final emphasis added, all others original.
[154] *FTP*, K143/H135, 293

The review of Aenesidemus also contains a condensed argument for the primacy of practical reason against the charge that has been raised. Though the argument appears in an extended and fuller format in the 1794/5 Wissenschaftslehre, it is helpful to look at Fichte's early attempt to solve the problem. The problem is that the primacy and indeed the very existence of practical reason in any substantive sense is in question if it is true, as it seems to be, that we need to know what we are willing in order that we will it. We can call this the 'knowledge condition' on willing. So the problem is that the knowledge condition seems to render indefensible any account of the primacy of practical reason, which is necessary for the Kantian enterprise.

Fichte attempts to resolve this problem here by claiming that practical reason and the moral law are not primarily aimed at producing actions, but only producing endeavours or strivings towards action.[155] Fichte then elects to represent 'the elements of this mode of inference in their highest abstraction', and says that 'If, in intellectual intuition, the I *is because* it is and *is what* it is, then it is, to that extent, *self-positing,* absolutely independent, and autonomous. The I in empirical consciousness, however, the I as intellect, *is* only in relation to something intelligible, and is, to that extent, dependent.'[156] There has been an opposition set up between two elements of rational beings. There is the I *qua* reason in general, and the I *qua* my own empirical consciousness. These two, however, are supposed to be in some way united. 'Since, however, the I cannot relinquish its absolute independence, a striving is engendered: the I strives to make what is intelligible dependent upon itself, in order thereby to bring that I which entertains representations of what is intelligible into unity with the self-positing I.'[157] We should note that Fichte here probably intends to use 'intelligible' to refer to the Not-I, insofar as it is an object of thought. My own empirical consciousness is dependent upon the not-I – Fichte has said previously that the correct interpretation of Kant's refutation of Idealism is that 'the consciousness of the *thinking* I…is possible only under the condition that there be a not-I which is *to be thought'*.[158] This striving to bring the self-positing I in harmony with the empirical I is 'what it means to say that *reason is practical*. In the pure I, reason is not practical, nor is it practical in the I as intellect. Reason is practical only insofar as it strives to unify these two.'[159]

Fichte's response to Aenesidemus is then that practical reason can exist and be primary insofar as the I in any of its finite determinations (i.e. individual persons) is *not* in harmony with

[155] *RA*, I:22, 75. Fichte probably has in mind the opening pages of Kant's *Groundwork*, where Kant makes numerous claims regarding the good will, among them the claim that the good will is good independently of whether it is efficacious or not.
[156] *R*A, I:22, 75
[157] *R*A, I:22, 75. Fichte uses the word '*entsteht'* which Breazeale translates as 'engendered'. It has a temporal meaning, but it seems clear that Fichte is here trying to argue that this is a logical order not a temporal one.
[158] *R*A, I:22, 75
[159] *R*A, I:22, 75

the core of subjectivity, the pure I. It is the nature of Reason to give grounds for things, and thus the unity would be one in which the I determines itself and thus determines all that is not-I. However, Reason finds that this is unfulfilled, and thus demands its fulfilment. So, Fichte tells us, practical reason is founded 'on the *conflict* between the self-determining element within us and the theoretical-knowing element. And practical reason itself would be cancelled if this conflict were eliminated'.[160] This last way of putting the point can be misleading, however, as it sounds very much like practical reason is consequent on theoretical reason. Indeed, this is Breazeale's interpretation of this passage. Breazeale says: 'From this it is surely obvious that practical reason always presupposes theoretical reason, inasmuch as the very task of practical reason is to overcome the conflict between the intellect and the pure I'.[161] But this is peculiar, as Fichte's words seem to imply that he is thinking that this is compatible with thinking of practical reason as primary in the transcendental sense. He says 'far from practical reason having to recognise the superiority of theoretical reason...'[162] Another peculiarity is the absence of discussion of the knowledge condition – that was Aenesidemus' primary criticism, which Fichte set out to rebut, but he seems to have forgotten it; at least, it is not clear how this conclusion relates to the knowledge condition.

In answer to the latter question, Fichte seems to think that the knowledge condition is successfully rebutted when he draws a distinction between what the moral law immediately directs itself to and what it mediately directs itself to. The former is 'the constant endeavour toward an action', and the latter is an action. Fichte regards actions as 'of course something which must be governed by the laws of this world', presumably because they involve the movement of natural bodies. Fichte's response is to grant to Aenesidemus that the knowledge condition holds for actions, but not for the 'constant endeavour toward an action'. Because Fichte seems to conceive of the knowledge condition, which is opposed to the principle that "ought implies can" as an attack on PPR (at least as it is found in Kant), then a rejection of the knowledge condition is a defence of PPR.

With regard to the former problem – that of whether Breazeale's view is correct, it must be said that when Fichte says 'Reason is only practical insofar as it strives to unify these two [the self-positing independent I and the intelligent dependent I]', he would not assent to the view that there is a time at which the conflict, which must be resolved by unifying these two, is non-existent. From this it follows that reason is always practical. But the question remains as to whether Fichte here assents to the transcendental sense of PPR. That is, does Fichte in the

[160] RA I:23-4, 76

[161] Breazeale (2013, 411). Whilst Breazeale takes this as evidence of his view – that theoretical and practical are equiprimordial or co-original, it is not clear that this passage actually does that. If practical reason, if as this passage seems to suggest, is consequent upon a conflict of theoretical reason, then that would not amount to equiprimordiality either.

[162] RA I:23, 76

review believe that practical reason is a necessary condition for theoretical reason? In other words, is it the case that without there being a conflict then I would not be able to know anything at all? Fichte speaks of the 'intelligent I', which would suggest a negative answer. If that is the case, then we see that at this stage, Fichte had only saw himself as setting out a defence of PPR insofar as it appears in the guise of "ought implies can" – a denial of Aenesidemus' knowledge condition. But, as we shall see below, the argument in the *WL* also claims that practical reason exists only because of a contradiction between the independent and dependent aspect of the I. The argument in Aenesidemus has a more limited scope, and there are no claims regarding the transcendental nature of the primacy, but the mere fact that Fichte claims that practical reason is a result of a conflict in the I does not mean that it cannot be so prior.

The equiprimordiality thesis may be interpreted in different ways. At its most basic, it is possibly a 'formal' sense – the result of the joint assertion of 'if P then Q' and 'if Q then P' where P is practical reason and Q is theoretical reason. This rather bare conception is acceptable, and as far as I can see, what Fichte thinks. The problem is that this seems to come too late. When Fichte asserts this equiprimordiality, it appears to be in the context of discussions of the powers of real empirical agents. That is to say that it appears primarily in the context of acting. Breazeale does note this, but he does not think it is an issue. For example he says, 'If thinking is to be "real", it must have an object, which can be grounded only in *feeling*, and hence in some hindrance to the practical power of the I. If willing is to be rational (i.e. if it is to be the activity of an *I*), it presupposes a theoretical acquaintance with the world within which one strives, as well as a determinate goal provided by "ideal thinking". The I must posit – i.e., know – itself as willing'.[163] This is all correct, as far as I can see, but it comes too late in Fichte's program. Of course in order to will that the world be X I must have a concept of X. But this assertion of the knowledge condition is not sufficient to establish the falsity of PPR. This would only establish the falsity of the temporal sense of PPR, which we have already rejected as not capturing Fichte's meaning.

However, there is a sense in Fichte that what he calls 'real activity' (i.e. doing) is prior, and then 'ideal activity' (representing) is posterior. The only sense that I can make of this is that ideal activity is to some extent 'parasitic', as it were, on real activity – Fichte calls ideal activity 'constrained and arrested and can occur only subsequently to a real activity'.[164] Whilst it is true that Fichte says that the I only is what it posits itself to be (a position which Breazeale takes as establishing the equiprimordiality thesis), it is not clear what Fichte is supposed to mean here.[165] That is, it could be that there is some bodily movement or act of will which occurs *prior*

[163] Breazeale (2013, 411)
[164] *FTP*, K52/H47, 148
[165] For example, see *FTP*, K48/H44 141.

to my thinking of it, because I do not self-ascribe such a bodily movement. Then, and only after this movement occurring, I self-ascribe it, and thus it becomes mine. In Fichtean spirit, this should not be thought of as an analysis of constitutive conditions for something to count as an action, but rather a genetic account of how agents are formed. It is quite possible that this is what Fichte has in mind as an account of how the powers of real and ideal activity work.[166] It is analogous to Fichte's account of intuition and concept. Fichte, like Kant, affirms that we cannot have intuitions without concepts and cannot have concepts without intuitions. But intuitions do in a real sense come first. It seems as though Fichte's view is that we have an intuition of an object brought to us by a feeling and then by self-ascribing that intuition we conceptualise it. The self-ascription is not *optional* but is certainly secondary. I believe that something akin to this is what Fichte has in mind when he says that '…the ideal activity would be a product of the practical power, and the practical power would be the existential foundation of the ideal activity…[the ideal activity] is that which witnesses the practical'.[167] But if something like this is right here, then we have rejected the equiprimordiality thesis in the strong sense that Breazeale and Zöller affirm.

I now turn to an examination of what I take to be the canonical form of Fichte's major argument, and the ramifications of the argument, and Fichte's position on PPR, for how his theory of how consciousness of freedom and consciousness of the world are linked.

2.4) "Without a striving, no object is possible"

The most famous passage where Fichte seems to affirm PPR has become known as the expression of the 'striving thesis' Fichte says '...*in relation to a possible object*, the pure self-reverting activity of the self is a *striving*; and as shown earlier, *an infinite striving* at that. This boundless striving, carried to infinity, is the *condition of the possibility of any object whatsoever*: no striving, no object'.[168] Fichte appears to be stating that the condition of the possibility of objects is the striving of the I. This thesis is actually the result of a lengthy argument which I shall attempt to reconstruct below. Briefly, the idea is that the I as infinite activity demands that all reality should be posited absolutely through itself. Fichte, soon after enunciating the striving thesis claims just this, and says that this demand 'is called – and with justice – practical reason'.[169] So it seems that Fichte affirms PPR because he affirms the striving thesis, which states that objects are conditional upon there being a striving of the I, and this striving of the I is an expression of reason's practical interests.

---

[166] I am thus in sympathy with Martin (1997, 134).
[167] *FTP*, K49/H45, 142.
[168] *SK*, I:262-3, 231
[169] *SK*, I:264, 232

The striving argument blends together transcendental and dialectical strategies.[170] Broadly, a transcendental argument is one that argues from a premise that is generally committed to, then to necessary conditions for that premise. A dialectical argument is one that argues from what appears to be a contradiction to a synthesis of the propositions in the contradiction. I take it that Fichte is interweaving these methods in this argument, and there are portions of the argument that are dialectical.[171] On the whole the argument is a transcendental argument because it is designed to show a necessary condition for the possibility of X – though it is dialectical because the condition for the possibility of X is a synthesis of two contradictory assertions. Before I begin exegesis of the stages of the argument, I should say that whilst Fichte talks about an antithesis or a contradiction, it is not clear what this contradiction or antithesis is at this stage. Matters are further complicated because Fichte mentions two antitheses. The first contradiction is between roughly the propositions that A) "The self is independent of the not-I" and B) "The self is dependent upon the not-I".[172] For convenience, I shall from now on refer to this as the Dependence Contradiction. We are told soon after that this contradiction can only be resolved by thinking of the relation between the self *qua* independent activity and the not-I as one of cause to effect.[173] We are also told that this causal relation itself contains a contradiction, but it is the only way possible to resolve the Dependence Contradiction, so Fichte sets himself the task of resolving what I shall call the Causal Contradiction: A) The I is to exert causality on the not-I, and B) The I can exert no causality on the not-I.[174] There is yet a third contradiction, between A) The I is infinite and unbounded, and B) The I is finite and bounded.[175] This is actually more like a re-statement of the Dependence Contradiction, but in terms of finitude, because for Fichte, only an infinite thing can be independent. Fichte tells us that such a contradiction, were it to hold, would reveal a contradiction in the very nature of the I – that it posits itself absolutely and posits itself as opposed to a not-I, which are supposed to be mutually exclusive. With the different contradictions laid out, I shall now proceed to an exegesis of the proof-structure of the argument.

The key to the proof is given when Fichte says 'Merely by positing something, it [the I] posits itself…in this something, and ascribes the latter to itself. Thus we have only to find a difference in the mere act of positing in these two cases [positing itself as infinite or finite], and our

[170] I am in partial agreement with Martin (1997), who claims the argument to be transcendental. (134) But Martin does not, as I do, view the detail of the arguments for the major transitions as dialectical.
[171] Fichte regards this method as different from what he calls the 'synthetic method' of the theoretical sections. The synthetic method is outlined at *SK*, I:115, 113
[172] This is explicitly stated in terms of a dependence relation at *SK*, I:249, 220
[173] *SK*, I:250, 221
[174] *SK*, I:254, 225
[175] *SK*, I:254, 225

problem is solved.'[176] Thus Fichte's first premise is that there is a sense in which the I is finite, and a different sense in which the I is infinite, or, that the I posits itself as in some sense limited, and in some sense not. This is just the second half of the third principle of the whole *Wissenschaftslehre* – that the I posits itself as partially limited by a not-I.[177]

There is a slight peculiarity to this premise. The peculiarity is connected with what I said above in the previous regarding intelligible fatalism. Fichte tells us that the positing of the not-I is 'a mere hypothesis; *that* such a positing occurs, can be demonstrated by nothing other than a fact of consciousness, and everyone must demonstrate it for himself by this fact; nobody can prove it to another on rational grounds.'[178] The *Wissenschaftslehre* tells us that *if* experience is actual, then certain conclusions regarding the structure of human knowledge and mental architecture follow. The *Wissenschaftslehre* does not tell us *that* experience is actual.[179] Fichte tells us that there might be an attempt to 'trace some admitted fact to this highest fact on grounds of reason; but such a proof would do no more than persuade the other that by admission of some such fact he had also conceded this highest fact'.[180] Indeed, earlier in the theoretical part of the work, this is exactly what we find Fichte doing. He says that the positing of the not-I must be an original act of the I because in order to make any knowledge claim or representation of the world whatever there must be something 'whereby it [the object of my knowledge-claim or representation] discloses itself as something to be presented'.[181] But, he says that is 'something that no object can teach me; for merely in order to set up something as an object, I have to know this already; hence it must lie initially in myself, the presenter, in advance of any possible experience'.[182] In other words, this premise, as is appropriate for a transcendental argument, is assumed to be acceptable to all.

The second premise (of the transcendental portion of the argument) is (roughly) that if the I posits itself as limited by a not-I, then the limitation takes the form of resistance. By resistance Fichte means that the I's activity is bounded by something. In order to explicate Fichte's thinking here, we need to look at the distinct types of activity that the I engages in. The first type of activity is called pure or infinite activity. This is the activity that the I engage in when it thinks of itself or posits itself. This activity is also called self-reverting activity. Fichte calls this activity 'infinite' because it is inherently self-reflexive. There is also objective activity, in which

---

[176] *SK*, I:255, 226
[177] *SK,* I:106-7, 106-7
[178] *SK*, I:252, 223
[179] Note the similarity between this formulation of the procedure of the *Wissenschaftslehre* and Niethammer's formulation of Kant's philosophy (that Kant says, 'if experience is, then…') – that Ameriks is very sympathetic to. It thus seems peculiar that Ameriks finds Fichte to be too dogmatic and not critical enough. See Karl Ameriks, *Kant and the Fate of Autonomy* (Cambridge, Cambridge University Press, 2000, 64ff)
[180] *SK*, I:252-3, 223
[181] *SK*, I:104-5, 105
[182] *SK*, I:105, 105

the I thinks something that is not itself. Objective activity is activity of the I that is intentional, and the intentional object is the not-I (or some determinate part of it). Objective activity is characterised by a resistance, what Fichte would later describe as a 'feeling of necessity'.[183] In other words, thought does not have the same power that it has over pure activity. Thinking about objects is constrained in various ways – I must respect various aspects of the object's form, structure, and content, in order to represent it faithfully. Fichte notes with a certain satisfaction that even the etymologies of the (German, but the same point applies to the English) words prove his point – they *object* to my activity.[184] Therefore to say that the limitation that the I posits as applicable to itself via the not-I takes the form of resistance is to say that the activity of the I is resisted in various ways; the activity (when directed toward objects) is not infinite. In other words, the positing of the not-I as limiting the I limits only because it resists the I's activity. To sum up slightly more formally:

1) The I posits itself as limited, that is, the I posits a not-I
2) The limitation on the I takes the form of resistance to the I's activity

The move after this is not clear. This is where the dialectical work of the argument begins. We have seen that the relevant contradictions (Dependence and Causal) have been stated, and now that Fichte has got this far with the transcendental method, he needs to switch to a dialectical method. The dialectical sub-argument motivates the move from these premises to the conclusion. Indeed, without this sub-argument, the transcendental argument would be straightforwardly invalid. It is now therefore necessary to clarify the line of argument that guarantees the move from two to three.

Recall the two types of activity (pure and objective) that we have just considered – the I is supposed to be absolutely self-determining and infinite, but at the same time is constrained by the world and is thereby determined, not determinant. These two types of activity are distinct, but in Fichte's words, there must be a 'bond of union, whereby consciousness is conducted from one to the other...'[185] Fichte is here noting that of course we have consciousness of ourselves (infinite activity) and consciousness of objects (objective activity). The problem is therefore how to reconcile these two. The I must have both activities, but it will

[183] '[First] Introduction' I:424, 8. The distinction between the representations accompanied by a feeling of necessity and those accompanied by a feeling of freedom means that for Fichte, experience must be something thicker than mere apprehension of sense data. Experience is identified with the representations that are accompanied by a feeling of necessity. This means that Ameriks is wrong to criticise Fichte on the grounds of starting with the Cartesian ideas and trying to argue outwards. In fact, Fichte follows Kant's strategy (which Ameriks thinks is the superior) of starting with empirical judgements.
[184] *SK*, I:256, 227. The etymology is helpful: in German *gegenstand* (= stand against), English 'object' (from *ob-jacere*, 'thing put before, present, oppose, cast in the way of')
[185] *SK*, I:256, 227

not do to, as Kant did, merely state that the powers of the mind might be joined in a 'common root' but that is a question that we can and should ignore.[186] That is not systematic, not scientific, and ultimately, for Fichte, not philosophical.

We must therefore find what joins these two activities as modifications of the same self (which is itself a type of activity). Fichte claims that the bond of union that links the finite and infinite activities is the causal connection – the causal connection asserted in the thesis of the Causal Contradiction. The pure activity of the self should be the cause of the objective activity of the self, or 'the self should determine itself to the second activity by means of the first'.[187] It would then follow, by transitivity, that pure activity would be mediately causally related to the not-I. Clearly, however, he is not thinking of a temporal causal relation. For one thing, time is, according to orthodox transcendental idealism, a form of intuition by which parts of the not-I is represented (to use a mixture of Kantian and Fichtean terminology). The positing of the not-I must precede, in a non-temporal sense, any temporal relation (such is a consequence of the statement above – that no object can teach me what it is to be presented). So when Fichte here says, cause and effect', it would be better to think of the logical 'ground and consequent' or 'determinant and determined' relation. Fichte follows this by unpacking how this is to be, by 'enter[ing] more deeply into the meaning' of the positing of the not-I.[188] Fichte then recalls the Causal Contradiction. We can formalise this somewhat with the following:

3) The pure activity of the self and the objective activity of the self must be 'one and the same...there must...be a bond of union'.
4) This can only happen if the pure activity of the self and the objective activity of the self are related as cause to effect.
5) But the Causal Contradiction holds.

Fichte therefore attempts to resolve this contradiction by focusing more deeply on the meaning of the proposition that the not-I is the only candidate for the content of the act of positing the not-I, or counterpositing, as he sometimes puts it. One paragraph represents an attempt to do this in terms of limitation in general – without any determinate limits. He says that so long as a mere object in general is posited (i.e. the whole not-I), the requirements (that the pure and objective activity be related in the right way) are satisfied. This is because there are boundaries, but the boundaries are determined by the absolute spontaneity of the I. Thus Fichte claims that on this view, the self 'is finite, because it is to be subjected to limits; but it is infinite within this finitude because the boundary can be posited ever farther out, to infinity. It

[186] See for example, Kant, *Groundwork*, 4:391, 47, also *Critique of Practical Reason*, 5:91.
[187] *SK*, I:256, 227
[188] *SK*, I:258, 228

is infinite in its finitude and finite in its infinity'.[189] If this sounds self-contradictory, that is because Fichte intends for us to take it that way – he is not actually asserting this. This is a consequence of a proposed solution to the antithesis that Fichte will (shortly) reject. He says that this leads to an 'absolute confinement' which is in contradiction with the 'absolute infinite nature' of the self. In short, it will not do if we merely think of the positing at such an abstract level as 'object in general' because this leads us to think of the I as finite and infinite in the same senses, and this is contradictory.[190]

Fichte then tells us that not only does the I posit a not-I in general, but it posits objects, that is, it posits determinate parts of the not-I which are thought by virtue of our objective activity. Our objective activity that has as its intentional object the not-I needs to have specific parts of the not-I as its object. In other words, the I posits activity (found as resistance) in objects. The resistance found in the objects of objective activity is itself a type of activity. This activity must be opposed to some activity of the I, but an activity which is non-identical with the activity of positing the object in the first place, otherwise they would annul each other, cancel each other out.[191] Fichte then claims that 'therefore, as an object is to be posited, and as a condition *of the possibility of such positing,* there must be another activity (=X) occurring in the self, distinct from that of positing'[192] The thought behind it being a transcendental condition seems to be that if there were not a second activity of the self then there would be no positing, because the object and the positing would cancel each other out. This seems actually quite right if we transpose it into non-Fichtean language, but we have to bear in mind that objects (at this abstract level) are essentially bundles of resistance. If I had a certain intentional content, and the object of that content resisted my intentional activity, I would not get any content in the first place, and therefore I would know no object.[193] It might be said that here Fichte has just got the picture inverted. It is not, we might think, that insofar as the I's activity is resisted is there an object, but only insofar as there is an object is the I's activity resisted. The latter seems like a more natural standpoint. In Fichte's defence, I can say two things. The first is to reiterate the point above that the concept of an object is something that no object can teach me; the second

[189] *SK*, I:258, 228

[190] It seems that this is a charge which Hegel raised against Fichte, perhaps taking Fichte's final position to be identical to the position rejected in this paragraph. For example, he says 'The ego clearly posits an object, a point of limitation, but where the limitation is, is undetermined. I may transfer the sphere of my determination, and extend it to an infinite degree, but there always remains a pure Beyond, and the non-ego has no positive self-existent determination' Hegel, 1995, 498)

[191] *SK*, I:258-9, 228

[192] *SK*, I:259/228 emphasis original.

[193] It seems plausible to regard a paradigm case of this a case of the object resisting conceptualisation. Then of course Fichte would have to deny the existence of non-conceptual content. If there is no non-conceptual content, then resisting conceptualisation is a case of not knowing an object at all, and so the positing would never happen in the first place. It would take me too far afield to get into the question of Fichte on non-conceptual content; all I needed to show here was that sense can be made of this claim.

is that this is less of an argument and more of a re-statement of the position that Fichte wants to defeat: i.e. dogmatism (or any form of naturalism).

So, we get:

6) In positing the not-I, the I posits an activity in objects.
7) This activity is opposed to some activity of the I, which is non-identical with the act of positing the object, and this is a condition of the possibility of such positing.

Fichte then tells us that the only candidate for the other activity of the I is the pure activity of the I. He gives three requirements that the other activity must satisfy – it must be not eliminated by the object, or co-positable with it; secondly, it must be 'absolutely grounded' in the I, because it is independent of any object and all objects are independent of it. Thirdly, it must be infinite, because objects are capable of being posited 'out to infinity', and the activity must match this.[194] That is to say that the object must be opposed to the pure activity of the I.

8) The object is opposed to the pure activity of the I.

From this, Fichte concludes that the objective activity of the I is made possible by, that is, is caused by (in the special non-temporal sense outlined above) the pure activity of the I. He says 'Only *insofar* as this activity is resisted, can an object be posited; and so far as it is not resisted, there is no object'.[195] Fichte now takes himself to have shown that (5) is false – that is, it is not contradictory that the pure activity of the I be related to the objective activity of the I, and thus to the object itself, as cause to effect – he has resolved the Causal Contradiction. There must be a connection between the infinite activity of the I and the object that is posited by the objective activity of the self. This connection is mediated by the objective activity of the self, which is related immediately to the object, and the object in turn is related mediately to the pure activity of the self. Indeed, this point could be drawn out with an analogy with Kant's transcendental unity of apperception (which Fichte draws a parallel with as well). For Kant, the transcendental unity of apperception means that something can only be mine if it is possible that the phrase "I think" accompany the judgement. If we view Fichte's problem here through the lens of this Kantian vocabulary, the problem is that the objective activity of the self could only be the intentional activity of the I if it were shown that there is a real connection between that activity and the pure activity of self-consciousness.[196]

---

[194] All of this is in the same paragraph, S*K,* I:259, 229
[195] *SK*, I:259, 229
[196] Kant might object to this parallel by claiming that Fichte's investigation into the 'real connection' or the 'common root' of the powers of the mind is illegitimate, and even unnecessary, on the basis that we know already that it is at least possible that there is a connection, because the power of self-

So now Fichte has resolved the Causal Contradiction, he now turns to this pure activity in its relation to the object. Recall that the resolution of the Causal Contradiction was supposed to yield the key to the resolution of the Dependence Contradiction, which is the overall aim of the striving argument – to eliminate the dependence of the I *qua* intelligence. The Dependence Contradiction is resolved via the Causal hypothesis because the I *qua* intelligence is dependent upon the not-I, but if it were shown that the not-I were to be in turn dependent upon the I, then the I would be, by transitivity, wholly self-determined and independent. In order to show this, Fichte examines the connection between the pure activity of the I and the object (or the activity of the object). He first notes that they are 'perfectly independent of each other and utterly opposed', but 'they must nonetheless be connected'.[197] This again is a necessary condition on the object being posited (which we know to be true without argument), which, as Fichte says, happens absolutely, without any ground, or as he says 'by the act of positing merely as such', rather than, say, positing it *as* thus and so. In other words the positing of the not-I is non-inferential. In this sense the positing is 'absolute'. There is, however, a different, though related, sense in which the positing of the not-I is absolute. It is absolute because is based on a connection that depends 'entirely on the I'.[198] I take it that he means by this that there would be no connection if the I was not active and resisted by the object, or that the I is the keystone in the connection, and without it no other element can play its role.

9) The object is absolutely connected to the infinite activity of the I.

For Fichte, to say this is to say that they are 'posited as absolutely alike'.[199] It is not clear what this is supposed to mean. One way of thinking about it is this: To say something is absolutely connected is to assert its identity, but to assert identity just is to assert absolute likeness. But the activities in question – the activity of the object, and the pure activity of the I thinking itself, are obviously not the same. So, Fichte says, we can only say that their likeness is demanded – they ought to be absolutely alike.[200] Fichte moves on rather quickly to discuss the derivate question of which should conform to which. However, the notions of likeness and absolute connection require more explanation. It clearly cannot be that to assert absolute connection is asserting identity which is further to assert likeness. If this were right then the absolute

consciousness and consciousness of objects are two types of consciousness of the same mind. Be that as it may, and Fichte has no need to disagree, it could equally well be said in response to Kant here that (using Kant's own conception of philosophical deductions) we are not answering the *quid facti* (the question of fact – does a connection exist?) but the *quid juris* (the question of right – with what justification do we think such a connection?). For a recent paper detailing Fichte's relationship to these two questions, see Bruno (2018).

[197] *SK*, I:259-60, 229
[198] *SK*, I:260, 229
[199] *SK*, I:260, 229
[200] *SK*, I:259, 229

connection would be cancelled entirely by the observation that the I and not-I are not alike. But the absolute connection is maintained, albeit in a different form. Fichte does not renounce the claim of absolute connection, only the claim of (actual) absolute likeness. So the 'likeness' needs to be thought of in a different way. I suggest that we look closer at Fichte's claim that absolute connections are to be thought of as connections that depend entirely on the I. This means that for relations of likeness there is a sufficient ground of such likeness in the I. All the relata have their distinct parts to play, and none can be reduced to the other, but the relation is absent unless the I is present. It is clear that the not-I should conform to the I; his words are that 'what is required is the conformity of the object with the I'.[201] In a footnote Fichte claims this is the key insight of Kant's notion of the categorical imperative – that all reality should be like the I.[202] This footnote contains the key to thinking this notion of absolute likeness. It is that we have reached the limit of what reason is able to determine directly or absolutely. Reason is confronted with the difference of the activity of the object, which it is unable to determine directly. It therefore transforms its thought that the object be absolutely like the activity of reason into a demand that they become so.

There is also a slight ambiguity in the text. In this section, which forms the latter half of the dialectical argument, Fichte speaks about an absolute connection between the pure activity of the self and the object, or not-I. Before, however, he spoke about an immediate relation between the object and the objective activity of the I, and a mediate relation between the object and the pure activity of the I.[203] I do not think this is any real inconsistency. It is not obviously contradictory to say that the relation between the I and the not-I is mediated, but absolute. It is mediated insofar as it has a third term which the relation must pass through, as it were, but it is absolute because the relata form a unity, albeit one that is not actual, but ideal. It is to say, at least in this context where the relata are not absolutely alike, that the connection is mediated, and the relata ought to be absolutely alike, and one should conform to the other.

10) The object which is posited absolutely (and related mediately to the pure activity of the I) should conform to the pure activity of the I.
11) If the object does not conform, then the conformity is demanded.

Here we are at the end of the dialectical portion of the argument. Fichte takes himself to have shown that the pure activity of the I is the necessary condition for the objective activity of the I. He also then takes himself to have shown that the object that is posited must conform to the

[201] *SK*, I:260, 230
[202] My reading of the argument therefore differs from Seidel, for example, who takes Fichte to mean the Absolute I by 'possible object' in the phrasing of the conclusion. This is, in my view, not borne out by the text. See Seidel (1993, 112f)
[203] *SK*, I:257, 227

pure activity of the I, as the likeness between them is 'absolutely demanded'. But we got to this position by a dialectical argument showing the various oppositions that are present, between pure and objective activity; between pure activity and the object. Fichte therefore concludes with the third transcendental premise, the thesis that only a striving being is capable of finding the resistance required for the act of counterpositing (positing the not-I), or that the necessary and sufficient condition of this resistance is that the subject is a striving subject. This is to say that without a striving, no object is possible. This is because we have shown how the object must be absolutely related to the I, in order for the positing of the not-I to be possible. This requires that the I and the not-I be absolutely alike, which they are not. The demand that they are absolutely alike, and the awareness that this demand is not fulfilled, is the striving. This striving is in turn a condition on the possibility of objects in the first place, because were we not striving beings, there would be no resistance, and objects are only related to us by resistance. Striving is what makes possible my theoretical faculty, and my agency. It is, as Martin says, not the content of any action, but a 'pre-intentional condition on intentional action' and indeed on knowledge.[204]

12) The necessary condition for the possibility of limitation by resistance (and all positing of a not-I, that is, all intentional experience of a world) is that the I is a striving I.
C) Therefore, the I is a striving I.

It is interesting to note that Fichte says, 'in relation to a possible object, the pure self-reverting activity of the self is a *striving*'.[205] What he means is that the relation or connection between the pure activity of the I and the object, is one in which the I is striving to make the object like itself. In other words, the I seeks that the world be fundamentally the mirror of itself. Pure reason seeks that the world be reasonable, and in seeing that it is not, strives to make it so. As Fichte says in an essay contemporary with the 1794/5 lectures, 'Since he is a being that represents and, from a certain perspective…must represent things as they are: so, through the fact that the things which he represents do not harmonise with his drive, he falls into a contradiction with himself. Thus the drive to work on things so that they agree with our desires, so that actuality corresponds to the ideal. Man necessarily proceeds to make everything, as well as he knows it, reasonable'.[206] Fichte warns us that this should not be taken as an argument for the proposition that the pure activity of the I is necessarily related to an object, without any special act of connection. To take it as so would be a mistake, and would mean

[204] Martin (1996, 30).
[205] *SK*, I:261, 231
[206] Fichte, (1996[1795]), 122. This contains the vocabulary of drives which is missing from the presentation of the striving argument given in *SK*.

'intelligible fatalism'.[207] We should instead take it as saying that '*if* it is so posited [as in relation to an object] it is posited as a striving'.[208] Let us now take stock of the position so far.

The major import of this is that this practical relation to the world is primary. Fichte has argued back from the fact of conscious intentional experience to a necessary condition of this experience being that finite human agents are striving beings, that is, are primarily orientated toward the world in practical, not theoretical terms. He has shown this by relating our intentional activity to our self-consciousness, and then by relating our self-consciousness essentially to consciousness of resistance, and a striving to overcome that resistance. In Fichte's words, a proof that reason can (and must be) practical 'can be achieved no otherwise than by showing that reason cannot even be theoretical, if it is not practical; that there can be no intelligence in man, if he does not possess a practical capacity; the possibility of all presentation is founded on the latter'.[209] The most pressing questions now are these: firstly, to what extent do the striving argument and the resulting thesis support PPR? Secondly, does the striving argument work?

The way of putting the point seen in the essay on language is helpful, because it allows us to see what is on the one hand the biggest problem with the argument and at the same time allow us to make more sense of the argument. As it stands the argument is not valid because the move Fichte makes from (9) to (10) is illegitimate. Why does Fichte assume (other than his general methodological position) that the object should conform to the subject?

There are at least two possible answers to this.[210] One is given by the Kantian notion of the interests of reason. The other comes by reflection on Fichte's usage of the word 'absolute' to describe the connection. I shall start with the latter. I said above that Fichte conceives of absolute connections as connections that depend entirely upon the I. If a connection depends entirely on the I then it seems that there must be something about the I that makes the connection hold. This might mean something along these lines: if the object lies outside the scope of the infinite or pure activity, then it is not an object at all. This is the familiar transcendental idealist point that something must be knowable in order for it to count as an object. In Fichtean terms, knowable would here be perhaps cashed out in terms of resistance. Being knowable is not, one might think a 'real property' of objects, but it is a function of their possible relationship to the I. In this sense, the whole connection, and the relata of the connection, rely fundamentally on the I. if this is right, then to be in an absolute connection is to be in a unity whose character derives from the I, so it would indeed be natural to posit the relata as 'absolutely alike'. I take it that by this Fichte means something like that the objects

[207] *SK*, I:263n, 232n
[208] *SK*, I:263, 232
[209] *SK*, I:264, 233.
[210] The answers are not mutually exclusive.

would be fundamentally reasonable, or made in reason's image, as it were. However, as Fichte swiftly points out, this is not the case, and thus, as Breazeale says, reason's original pretentions are turned into an injunction, to make the world reasonable.[211]

The other avenue to make Fichte's inference at the end stronger is perhaps a more speculative detour through his views on the interests of reason. In the section on interests of reason in *SE*, Fichte begins by noting a fact – that some occurrences interest us, and others do not.[212] This interest is immediately related to what I want and 'cannot be produced through any rational grounds'.[213] This is the non-inferential character of interests. Interests arise with a feeling of harmony or disharmony of the subject with the thing in question. However, given that interests are felt, and I can only feel myself, the harmony or disharmony would have to lie within myself.[214] That is, I feel myself to be harmonious or disharmonious. Fichte then claims that all interests are mediated through my interest in myself, and this interest in myself has its origin in a drive. This drive is the drive toward harmony between the original and the actual I.[215] This is called the pure drive – the drive to 'activity for activity's sake, a drive that arises when the I internally intuits its own absolute power'.[216]

The general idea is that the conditions for self-consciousness are found in human action, and the conditions for action as such are to be found in freedom, and freedom is to be understood as a law that governs the subject independent of the object.[217] It must be, for Fichte, that the moral law itself is a condition for action, although, to avoid denying the obvious existence of immorality, Fichte must also maintain that the moral law cannot cause our actions.[218] The highest form of the moral law is the categorical imperative, which Fichte restates as 'Always act as if you were to give laws to yourself for eternity'. In other words, the categorical imperative demands unity of the self.

I think we can reconstruct the argument that yields the needed inference as follows.

1) Reason has interests (as proven by the argument in SE) (assumption)
2) One of these interests is the interest in unity, which is because of a drive for unity (specification of 1)
3) The drive for unity is a result of the moral law, as unity is the moral law's highest command (specification of 2)

[211] Breazeale, (2013b. 345-6)
[212] *SE*, IV:143, 136
[213] *SE*, IV:143, 136
[214] *SE*, IV:143, 136
[215] *SE*, IV:143, 137
[216] *SE*, IV:144, 137
[217] Guyer (2015, 139)
[218] Indeed, he must also say this if his account is going to fit in with his account of human existence as fundamentally striving.

4) Therefore, unity is morally demanded (from 3)
5) Unity requires the determination of the not-I by the I (specification of 3 and 4)
6) Therefore, the determination of the not-I by the I is morally demanded (from 4, 5)

This argument is somewhat speculative and controversial, and may not be what Fichte intended. However, it could be said that these conceptions of the interests of reason and the drive for unity supplement the striving argument, by showing how it follows from the beginning of the WL that reason issues an injunction to itself to be unified. In other words, persons are essentially striving, and this is because all persons, *qua* rational beings, feel the force of the conflict between the disunity of their self and the drive for their self to be unified – and the drive to be unified is prompted by tacit recognition by Reason itself of its being subject to the moral law. This has hints of freedom already in it, and we have seen that Fichte regards his theoretical philosophy to be important for free will. However, the primacy of practical reason is only necessary preparatory groundwork for the doctrine of free will that we later find. It provides Fichte with the warrant for pursuing the idealist framework. Free will is also necessary because without it, the striving argument would prove only that rationality is a fundamentally tragic state – striving without control over one's situation.

3) The Theoretical Exploration of the Concept of Freedom

In this chapter, I will interpret Fichte's arguments for freedom as they are found in a core text of the later Jena period – the *System of Ethics*. Whilst there has recently been interest in Fichte's account of freedom in various ways, by various authors, for example Goh (2012), Wood (2016), Kosch (2018), these authors focus on presenting Fichte's often difficult views, rather than attempting to make sense of Fichte's arguments as they are presented in the text.[219] Fichte clearly thought of the deduction as a single argumentative thread, and I intend to, as far as possible, present this thread and articulate it.

3.1) Freedom in the first Deduction of the System

I have now shown that Fichte's thesis of the primacy of the practical is a central part of his views in the Jena period. Now I will move on to discuss the arguments for free will that we find in the System of Ethics. Fichte begins *SE* by investigating the I – the very foundation of selfhood. Fichte claims that the basis of the self is the will, or, as he puts it, that I only find myself as willing.[220] We should therefore begin our exploration of Fichte's theoretical use of the theory of freedom by looking at the theory of willing. Willing, he thinks, is something

[219] Beiser's (2002) book does so, from pages 323-333, when it comes to Fichte's deduction of the external world.
[220] *SE* Theorem 1 IV:18, 24

primary, that is, unable to be explained on the basis of something else. This means that willing is something absolute – it can only be explained on the basis of itself. What Fichte probably means by this is that the term 'willing' is indefinable in terms of something that itself does not make reference to something at the same level as willing. So if one were to define willing, it would always have in that definition a reference either to willing itself (in which case the definition would fail) or reference to, e.g., spontaneity, intention, action, trying, all of which can be defined in terms of willing in some way (in which case the definition would be circular and not explanatory). Another way of thinking about what Fichte means, which is perhaps closer to his own thought, is that willing appears to be a basic phenomenon – it does not appear to be constructed out of more basic constituents of thought – at least, not ones that are more basic and of a different kind. By this I mean that one might be able to zoom in, as it were, on an episode of willing, but all one would find there would be more willing. This is in line with the Leibnizian-Wolffian way of thinking of definition. Leibniz says that 'we also have a distinction between *nominal definitions*, which contain only marks of a thing to be distinguished from other things, and *real definitions*, from which one establishes that a thing is possible'.[221] Fichte says 'It is here presupposed that one knows what *willing* means. This concept is not capable of a real definition, nor does it require one. Each person has to become aware within himself of what willing means, through intellectual intuition, and everyone will be able to do so without any difficulty'.[222] I take it that Fichte would agree that willing can be given a nominal definition, because willing can be roughly set apart from other things, or, in Leibnizian terms, there are marks of willing that serve to distinguish it from other things. Presumably, even one such as McTaggart, who denies that causation involves some notion of power or activity, has this intellectual intuition by which one knows about willing.[223]

A key point here for Fichte is that one finds oneself *only* as willing. This is, I submit, a guise of the multi-faceted Fichtean doctrine of the primacy of the practical. I do not find myself as thinking, but as willing. Why should this be so? I think we can motivate the thought thus:

---

[221] Leibniz (1989. 26)

[222] *SE* IV:19 p25-6. Willing is also described as 'something primary, grounded absolutely in itself and in nothing outside of itself'. *SE* IV 24, p30

[223] McTaggart says 'But I do not believe that there is any such activity to be perceived even when our volitions are causes. In my own case I can perceive no such activity. And I *can* perceive something else which could be mistaken for such an activity*. I am conscious of willing*. And then, after an interval of more or less duration, I am conscious that the result which I willed – the movement of my arm, for example – has taken place. In some cases, also, I am conscious of a feeling of tension or strain within myself. But this is all' 'The Meaning of Causality' in (1934[1915]) 164 first emphasis original, other emphasis added. This description of McTaggart's seems strained – the way he describes it, it is as if it is a surprise to me that I find after my willing that a bodily movement is accomplished.

1) To originally find myself is a kind of apprehension, or a thought.[224]
2) To find myself as thinking, I would have to have originally been engaged in a thought or apprehension.
3) Therefore there would have to be a further act of finding beyond the first act of finding, *ad infinitum*.
4) Therefore I do not find myself originally as thinking.

That Fichte has something like this in mind, I think, can be seen with the following passage. He says 'Originally and immediately, the former [thinking] is, for itself, by no means an object of any particular new consciousness, but is simply consciousness itself'.[225] This is an instance of Fichte's so-called 'original insight' – that the ordinary way of thinking about self-consciousness, as a kind of consciousness of consciousness, is untenable. Consciousness ordinarily takes objects, but when consciousness is conscious of thinking, this is not so – there is an undivided subject-object, rather than a subject and object relation. But finding oneself is a kind of subject and object relation, so it has to be consciousness *of* something else – it has to have an intentional object that is distinct from itself. It has to be directed to that which is objective, or directed primarily towards objects. This is the will, or willing.[226] Fichte describes the objective, willing, as existing 'independently of thinking' and 'hence something *real*'.[227] It has to be a real determining of oneself through oneself, which is the character of willing. It seems to me that Fichte argues in the following way (in the 'proof' of the theorem, after he has finished explicating it).

1) The I is that for which the thinking/acting subject and the one thought about/acted upon are the same
2) The object of thought is supposed to be something originally objective, that is, it is found as something which is not itself thought.
3) But given 1), and the character of this object of thought, given that it is self-ascribed, it must be the I, conceived in a certain way.
4) This way would have to be such that the I could be thought of as object.
5) The I can only be thought of as object when it is real acting (otherwise a regress threatens, and the I cannot be properly called an object of thought then).
6) The real acting must be a kind of determining of oneself through oneself, or willing

[224] *SE* IV: 19, p25
[225] *SE* IV:20, 26
[226] Note that Fichte rarely uses the objectifying language of 'the will' and prefers to talk of 'willing'.
[227] *SE* IV: 22, 27-8

Fichte therefore seems to give an account which relies on the idea that since willing is a kind of real determination, rather than an ideal (or subjective) determination of the I, it is the only kind of thing that would enable the I to grasp an object. One might raise a question about how the self-ascription in 3 happens. Fichte does not attempt to broach this question, explicitly stating that he is not here concerned with how it happens, but only that it does.[228] I would suggest that Fichte's motivation here is that we are originally sensitive (but perhaps falling short of representing) to some kind of willing, and that we are also sensitive to the idea that there could be no other candidate for self-ascription. In fact, thinking of it as self-ascription is in some way to do the act a disservice, because it may seem to imply that there are other candidates, such as in a case where a forgetful person doesn't know who the Professor Emeritus is, and looks through the department books and concludes that it must be themselves on the basis of their search. This self-ascription in this case, and others like them,[229] are not primordial enough for Fichte, because the thing that is to be self-ascribed is able to be thought of under a description different to that under which it is self-ascribed. One can think of it as 'that title' or 'the title of the department', and only after some other thought does one think of it as 'my title'. But in this case of willing there is no possibility of thinking of it as 'that willing' distinct from 'my willing'. I take it this is because in Fichte's mind the proto-rational creature would be sensitive to the very personal and wholly unique relation in which they stand to their own will. Indeed, this original *de se* knowledge may be used as a model to explain other, less basic *de se* ascriptions.

The second theorem is 'Willing itself, however, is thinkable only under the presupposition of something different from the I'.[230] We are told that to will something 'means to demand that some determinate object…become an actual object of experience'.[231] This therefore requires that we already have the concept of something that is external to us, which is 'possible only through experience'. I take it that Fichte does not mean that the concept is *learned* through experience, but that the concept's 'activation condition' is experience. That is, in order to possess the concept I must first have had experience, but that does not mean that there is an experience that taught me the concept. This is the same as Kant's famous claim that one cannot infer from the fact that all knowledge begins with experience to the (false) idea that all knowledge arises out of experience.[232] The third theorem is 'In order to find my true essence

[228] *SE* IV:19, p25
[229] The classic case of the sugar in the supermarket is one such case, or any case with the same structure. Perry (1979) is the *locus classicus* for cases like this.
[230] *SE* IV:23, p29
[231] *SE* IV:23, p29
[232] Kant *CPR* B1. Fichte also says something similar to Kant, which is compatible with what he says here on my reading, but is contradictory if we read it in the sense of experience teaching me the concept, in the earlier WL. Here Fichte says that 'If I am to present anything at all, I must oppose it to the presenting self. Now within the object of presentation there can and must be an X of some sort, whereby it discloses itself as something to be presented, and not as that which presents. But *that*

I must therefore think away all that is foreign in willing. What then remains is my pure being'.[233] The explanation of this involves a distinction between things that can be known mediately and immediately. Something can be known mediately if it can be known through knowledge of what grounds it. Fichte's example of is a pushing of a ball – if I knew the conditions which characterise the balls starting point and force, then I could infer to the conditions which characterise its endpoint. I could also know this by perception, but the fact that it would be possible for me to know it mediately shows that the movement of the ball is something dependent. Something that cannot be known in this way, something that could only be known through itself, is therefore something which one cannot be said to know by inferring from some prior conditions of the thing, or from knowledge of what grounds that thing. In this sense, willing cannot be known mediately – one cannot infer from something outside the I that supposedly grounds willing to explain it.

It seems (though it is not stated explicitly) that Fichte is thinking that this follows from what he said above, that willing cannot admit of a real definition. Is there a principle in the background to the effect that if something cannot admit of a real definition then that thing can only be grasped immediately and never mediately, or perhaps the reverse, that if something can only be grasped immediately then it cannot admit of real definition?[234] It is perhaps so. Fichte says in the early *WL* that the logical rule of definition is the following: 'it [the definition] must furnish the generic concept, which contains the ground of conjunction, and the specific difference, which contains the ground of distinction'.[235] So willing can be said to be contained within the concept of 'I-hood' or 'Activities of the I', and its conspecific is thinking.[236] Now willing satisfies the logical rule of definition (and therefore can be nominally defined) because there is something which differentiates it from thinking. The reason that it cannot be given a real definition is because the *differentia* cannot be exhibited in thought, only by intuition. Leibniz says that a real definition is one that establishes that a thing is possible, or a definition from which one can deduce the possibility of something. Fichte's thought here is that such a definition is not possible of willing, because of two reasons. The first reason is that in order to deduce the possibility of something, one would have to resolve it into its constituent

---

everything, wherein this X may be, is not that which presents, but an item to be presented, is something that no object can teach me; for merely in order to set up something as an *object*, I have to know this already; hence it must lie initially in myself, the presenter, in advance of any possible experience' (I, 104-5, p105).

[233] *SE* IV:24, p30

[234] This is possibly Meier's view in the *Excerpt From the Doctrine of Reason* (2016) for he states that 'If one would make a real definition, then one seeks to cognise the essence of the matter explained, either by mediate experience, or by abstraction, or by a proof from reason, or by arbitrary combination' (section 281, 63) and earlier that the only concepts we gain by immediate experience are concepts of 'actual things, insofar as they are present to us' (section 256, 58).

[235] *SK* 1:118, p116. I take it this is just another way of stating the Leibnizian point above.

[236] Fichte talks in a way similar to this at *SE* IV:84-5, 84

components.[237] This cannot be done with willing. The second reason is that, even if, *per impossible*, this could be done, no one would require a real definition of willing, because in order to be able to know that the real definition was a real definition, one would have had to have already had intellectual intuition of willing, rendering the definition superfluous.[238]

This is then why willing cannot be known mediately, because it cannot be picked out by purely discursive features; instead, it needs to be known on the basis of acquaintance with itself, that is to say, immediately. Willing therefore in some sense is to be called 'absolute', or at least that it has the appearance of absoluteness. There is a sense in which 'absolute' might be taken as a phenomenal character claim. This is in line with how Fichte describes willing – as a 'demand', but this paragraph is not about this kind of phenomenal character, rather about the explanation of why willing occurs at all. The absoluteness of the phenomenal character applies at the distributive level – at the level of individual episodes of willing. Fichte tells us as much when he is discussing the example of the steel spring. He says 'Within this spring there is undoubtedly a striving to push back against what presses upon it; hence this striving within the spring is directed outward. This would be an image of actual willing, as the *state* of a rational being; but this is not what we are talking about here'.[239] The absoluteness under discussion here occurs at the collective level – at the level of willing as such. It should be pointed out that Fichte has not yet committed himself to the idea that all episodes of willing have an absoluteness that is inexplicable. That is to say, the question: 'Why are you willing this or that X?' might still have an explanation, but the question 'Why are you willing *at all*?' has no explanation (or at least, no explanation that we could comprehend).[240]

There is a claim that the absoluteness of willing is easily found within oneself, and this knowledge cannot be discursively imparted, which is a recap of what went before. Fichte then claims that just because this is so, it does not follow that no explanation or ground for the appearance of absoluteness could be given. There might be something that explains or grounds the appearance of absoluteness, even though the appearance does not have the right kind of conceptual features to be discussed properly in abstraction from intuition.[241] So

---

[237] As Stang (2016, 23) says.

[238] This is the force, I think, of Fichte's clause at the end of the quoted passage on why willing does not admit of real definition – because it would be unnecessary. See also Beck (1965) who attributes to Kant the view that in real definition we 'make at least a problematical existential judgement and state the conditions under which this judgement could be verified so that the *definiendum* will be seen to have "objective reference"' (65). I take it that Fichte may agree, and point out that the very idea that one would make a problematic judgement regarding the existence of willing that could go any distance toward explicating the concept to a being that did not have such a concept is wrongheaded.

[239] *SE* IV:26-7, p32

[240] This last clause has similar overtones to Kant's *Groundwork* III claim that reason must think itself to be the author of its own principles.

[241] Fichte seems to think that such a grounding explanation of the appearance of absoluteness would render the absoluteness illusory. But he draws a contrast between this case and the case of Critical Idealism on things in space and time, which are further grounded (by the activity of the I) but *not*

one cannot infer from appearance of absoluteness to the fact of the appearance's ungroundability. But even so, 'no one will be able to provide such an explanation of willing from something else nor even to say anything comprehensible in that regard'.[242] Therefore, even though the possibility is not foreclosed by strict inference, the possibility is clouded by incomprehensibility. For the Dogmatist, this would be because ultimately speaking willing needs to be explained away, rather than explained.[243] Fichte then explains that the Idealist takes this as a sign that a philosophical project can be pursued on which we take the appearance at face-value – that willing is really absolute. This is taken as a kind of faith. The reason it is a kind of faith is because there is no possibility of a philosophical proof either way – the willing cannot be demonstrated to be absolute (because the inference from non-discursive grasp to non-groundable does not work) but cannot be demonstrated otherwise (because anything we say in this regard is incomprehensible, that is, cannot be counted as knowledge even if it is a coherent story). This faith is a sign of no theoretical insight, but rather of a practical interest.[244] It is a necessary condition on the possibility of the things we are already committed to (morality, freedom, and so on) that willing has this character. So we can take it that it does. Here again is a sign of the pervasive doctrine of the primacy of the practical. Here it is a kind of methodological primacy – theoretical reason and conceptual grasp are *silent* on the question of whether the appearance is real or illusory, but we have good practical reason to assume that it is not, so we can assume so. This is a reasonable position, even though it is openly faith, not knowledge.[245] One might ask why the absoluteness needs to be preserved. The answer to this question is, I think, to be found in the content of the practical interest. One of these is freedom. Fichte's thought is that freedom would be sacrificed if the absoluteness turned out to be illusory. This would be so because if the ultimate explanation for why rational agents willed anything at all was something that ultimately referred to an object of some kind (as per dogmatism), then that explanation would put rational beings at the mercy of passive or inert being. The explanation for why rational beings have the faculty of willing can only refer to the I, if independence, and therefore freedom, is to be preserved. We should therefore proceed to a discussion of this absoluteness, which is the last part of the first 'Problem' of the work.

---

rendered illusory. It is unclear to me what distinction Fichte is drawing here. The best I can think of is that objects in space and time are not rendered illusory because even common rational cognition knows that there is a distinction between the way things are and the way they appear *when it comes to objects outside the mind*, but not when it comes to those within the mind.

[242] *SE* IV:25, p31

[243] This is also probably a reason for Fichte distinguishing between empirical objects and the absoluteness of willing. An explanation of ordinary empirical objects is not explaining them away. But this seems unsatisfactory. What is it to 'explain away' unless it is in some way to show that the thing in question is illusory? If that is so, then the distinction that Fichte is drawing is no more illuminated.

[244] *SE* IV:25-6 p31

[245] It is the same kind of thing as Kant's arguments for the postulates of practical reason.

The example of the steel spring that Fichte gives is illuminating. The spring, when compressed, 'strives' to decompress. The condition, but not the ground, of this striving is the external object that is compressing the spring. The ground of the striving is the spring itself, or one of its properties, in this case the dispositional property to decompress. This inner ground is the analogue of self-determination. It is self-determination because the ground of the striving is itself. Fichte says 'it is the inner tendency of the spring to determine itself to a counterstriving, understood as the genuine essence of elasticity and as the ultimate ground of all of its appearances, just as soon as the conditions of their manifestation are present, a ground which cannot be explained any further'.[246] The fact that Fichte refers to this inability to explain any further is because the elasticity here is the analogue of willing for a rational being.

One might think that the analogy inapt, for two reasons, one which Fichte acknowledges and one which he does not. The one which he does is that the characteristic of the spring (elasticity) differs essentially from that of the rational being (willing), in the sense that one is necessarily activated upon the presence of a condition and the other is not.[247] This in some ways sounds like the phenomenon, discussed in contemporary philosophy, of a finkish disposition. A disposition is finkish if, given circumstances under which the disposition would normally issue in a response, it does not do so.[248] Firstly, many philosophers think that there cannot be so-called 'intrinsic finks', namely, a disposition that is finkish because of some other property held by the item in question, as opposed to being made finkish by an outside influence.[249] Secondly, that a disposition is finkish always admits of some determinate explanation as to why, in this case, the disposition did not trigger, even though it normally would. For Fichte, the only explanation for why this happens with rational beings is that they did not choose to do so.

The disanalogy that he does not mention is that it seems intuitive to think that the sciences could give an account or explanation of elasticity, so even if we agree that willing has the ungrounded character of absoluteness, elasticity is a poor comparison. It could be that what Fichte has in mind is something like the Putnam comparison between fundamental and ordinary explanation. Putnam gives the example of trying to explain why a round shape would not fit into a square hole using micro-physical properties. It would be, at best, too lengthy and full of irrelevant detail. This is because it would not capture the phenomenon at the level at which explanation was asked for. Similarly, it could be that here Fichte thinks of elasticity as not admitting of a further explanation because if we go to the micro-physical level (as the

[246] *SE* IV:27, p33
[247] Fichte seems to think that this is characteristic of what he calls 'a drive'. See *SE* IV:29, p34
[248] The *locus classicus* for this is Martin (1994).
[249] Clarke (2008) argues that there can be intrinsic finks.

dogmatist would want us to) we would end up very far removed from what we wanted explaining in the first place.[250]

When it comes to thinking through the I, Fichte has this to say – 'the problem is that of thinking the I at the requisite level of abstraction as something *subsisting* and fixed'.[251] But this should not be confused with thinking of the I as it is in itself; we are thinking of the I as it is for the I (as is consistent with the general strategy of transcendental philosophy) but also thinking of how the I necessarily appears to the I. Fichte's view here seems to be that there is a distinction to be drawn between the necessary shape of the I in the way the I conceives itself, and the way the I must be in itself. The latter would be inappropriate, in the sense that the I cannot be said to have a being outside of self-positing.[252] Then we also have to think of the thing that is the ground of an absolute willing. So we have to think of something that characterises the I at some level deeper than its manifest contingent appearances, that grounds absolute willing.

This is, Fichte tells us, the 'absolute tendency toward the absolute' or 'absolute indeterminability through anything outside itself' or 'the tendency to determine itself absolutely, without any external impetus'.[253] We are also told that it is not only a force, faculty, power, or drive. Fichte's thinking here seems to be a version of the thought that since a dispositional property (which is what all of the aforementioned things are variants of) must have a categorical base, and we are seeking here the categorical base in the I, a dispositional property would only be the way in which this categorical base manifests – the thing being sought here would be deeper than the disposition in question. This is what Fichte means, I think, when he says that the tendency in question cannot be these dispositional properties because they are 'nothing actual but only what we think of as preceding actuality, in order to be able to integrate the latter into a *series* of our acts of thinking'.[254] Perhaps there is some influence from Schulze here. What would be added to an account of something by merely saying 'It has a power of X'. The inference from 'S does X' to 'S has a power of X-ing', is trivial

---

[250] Perhaps one way of saying this is that the 'properly explains' relation is intransitive – even if theory X properly explains phenomenon Y, one cannot infer that theory X' which properly explains theory X will properly explain phenomenon Y. This cannot always be true, because Keplar's laws, which explain certain planetary movements, are themselves explained by Newton's theory, which also explains the planetary movements. But all we need for this claim to be plausible is the idea that it is not always transitive.

[251] *SE* IV:28, p33

[252] See, for example, the claim in the First Introduction that 'The nature of intelligence consists in this *immediate* unity of being and seeing. What is in it, and what it is in general, it is *for itself*; and it is that, *qua* intellect, only insofar as it is that for itself'. (I:435, p17). Similarly from the early Jena WL: 'The self's own positing of itself is thus its own pure activity. The *self posits itself*, and by virtue of this mere self-assertion it *exists*; and conversely, the self *exists* and *posits* its own existence by virtue of merely existing' (I:96, p97)

[253] *SE* IV:28, p33. It seems that Fichte thinks these designations are equivalent insofar as they have the same reference.

[254] *SE* IV:28, p33-4

in some sense. Perhaps this is why dispositions of this kind (force, faculty, power) are nothing actual – because the trivial inference holds, there is always a further question of whether the power of X-ing is something to be taken account of in the object, as opposed to just a way of putting some order on the flow of experience. The tendency, being 'not only' a dispositional property, can to some extent be characterised this way, but to do so exclusively would be to leave out the important categorical base.

The important points for the theory of freedom here are the following. The I is fundamentally a practical thing – self-consciousness reveals it as willing, not as thinking. The basis of this willing is in the I's absolute tendency towards the absolute, or self-activity for its own sake. This is based on a practical decision to treat the appearance of absoluteness of willing (that it cannot be further explained) as genuine. The Dogmatist is in the position of being tied to the necessity to give a theory of the grounds of absoluteness in willing, but such an enquiry is beyond the limits of reason. The significance of the theory of willing is that it needs to be clarified not only what it means to be free, but what it means for a self-conscious agent to be an agent at all.

I shall continue with the exegesis of this part of the *System*, because it is here that freedom first makes an appearance. Fichte tells us that we have produced a kind of consciousness that itself tells us that it is the original consciousness. Recall that we were engaged in proving that 'I find myself only as willing', and then given that willing is always objective, that is, outwardly directed, we also needed to investigate what it is that the 'myself' refers to in that theorem. That is what led to the discovery of the I as originally a kind of tendency towards self-activity for its own sake. The philosopher has therefore discovered this, but claims that it is already there anyway – it is implicit knowledge that has been made explicit.[255] The first important point here is that the I possesses an absolute power of intuition. The I possesses intrinsically a power of singular thought – to be a thinker (or a 'finder') is just to have this power. There then follows an important paragraph:

'The intuiting subject (the intellect), which becomes an intellect precisely by means of the postulated act, posits the tendency to absolute activity described above, in accordance with the postulate, as – *itself*: that is, as identical with *itself, the intellect*. The previously mentioned absoluteness of real acting *thereby* becomes the essence of an intellect and is brought under the *sway of the concept*, and this is how the absoluteness of real acting first becomes *freedom* proper: the absoluteness of absoluteness, the absolute power to make itself absolutely. –

[255] *SE* IV:30-1, p35-6. See also Wood (2000, 95f)

Through the consciousness of its own absoluteness the I tears itself away – from itself – and puts itself forward as something self-sufficient.'[256]

It is characteristic of intuition, as Fichte sees it, that it be directed toward something 'that is there independently of it, and that it is there just as it is intuited to be'.[257] So when the I-as-intellect intuits the I-as-tendency, the I-as-tendency is there independently of the I-as-intellect; it is not created or brought into being by the I-as-intellect. This is a result of the fact that the I-as-tendency fundamentally belongs in the domain of the *real* or *objective* – that is, it is action and directed towards things outside it. The I-as-intellect is *ideal* in the sense that its representations are parasitic on the real things. Fichte's talk here of 'the intellect' can be somewhat misleading, insofar as it is tempting to think of the fully formed intellect with discursive thought somehow noticing something. But what Fichte is really pointing to, I think, is the primordial episode of self-ascription, or the non-observational *a priori* character of knowing one's own actions. As is often stated, one doesn't need to observe that one has an intention to raise one's arm and that one's arm is rising; one knows it immediately. This is what is happening here. The tendency of the I is accompanied by an intuition of it. In other words, the I knows that there is some action, but know it comes to know this in a *de se* way.

Fichte tries to explain the above-quoted paragraph again by appeal to the image of the spring. The reason why we do not attribute freedom to the spring is that it is determined by something outside itself. In other words, its ground lies outside itself. There are two possible oppositions to 'that which is grounded by something outside itself'. The first is that which has no ground. There are no such candidates here – the idea of a 'being without any ground…cannot be thought of at all'.[258] Note that Fichte does not say that a being without any ground is impossible, but that it cannot be an object of thought. This leaves open the possibility that the world as such, or perhaps God, is a being without any ground, but that they are beyond the ken of comprehending thought. The I needs to have a ground because it needs to be an object of thought by itself. The second opposition, then, is that which determines itself. How can something determine itself? The only possible way in which this can happen is that thought is present. For Fichte, things have natures – to think of a thing and to think of its nature is to think of the same thing. Fichte also calls the nature the 'total sum of its determinations'[259] But given that thought, or thinking, is no thing, but 'purely…the agility of the intellect'[260], this condition does not apply to it. As Fichte says: 'As an intellect with a concept of its own real

[256] *SE* IV:32, 37
[257] *SE* IV:32, 37
[258] *SE* IV:35, 38
[259] *SE* IV: 36 38
[260] *SE* IV: 36, 38

being, what is free precedes its real being, and the former contains the ground of the latter. The concept of a certain being precedes this being, and the latter depends upon the former'.[261]

It is unclear why the intellect is unique in this regard. Of natures and things, Fichte tells us that the nature of a thing 'lies in its fixed subsistence, lacking any inner movement, passive and dead', and that in grasping a thing 'you have already grasped along with it this passive, unchangeable subsistence; and from this it follows, as something predestined, that under a certain condition a certain change will follow'.[262] The fact that change is included in this later passage is puzzling. It could be that Fichte is thinking of the following model of natures: A things nature just is the combination of all predicates that are correctly applied to it divided into times. Therefore the nature of the thing does not change when a predicate stops applying to an object, because the nature has time-indexed predicates. But this is such an abstract conception that it is unclear why it should not be applicable to the intellect. Evidence that Fichte is not thinking of this conception is that he speaks over these pages of things existing in advance of their nature – that this is not possible in the case of things in general, but is possible in the case of an intellect. If the nature was a set of time-indexed predicates then it would be eternal, but this would mean that all the natures exist independently of and prior to the things themselves. However, Fichte speaks of the things and their natures as coinciding, so this is probably not what he means. Perhaps we can get a better grasp by formulating the argument schematically:

1) Only that which determines itself can be free
2) To determine itself, a thing has to exist in advance of its nature
3) Therefore a free being has to exist in advance of its nature
4) Things cannot exist in advance of their natures
5) So whatever self-determines is not a thing, but thinking
6) Whatever self-determines must be an intellect
7) Therefore whatever is free must be an intellect

This is a good overall account of the argument here, but there is a step missing between 4 and 5. The obvious question to ask is why Fichte assumes that the thing/thinking dichotomy is exhaustive, or, to put it another way, what entitles the move from 'not a thing' to 'thinking'. It appears that this is because when it comes to determining itself, Fichte asks 'What does this "*itself*" mean?' and answers 'Some duality is obviously being thought in this case'.[263] This duality, he tells us, is accounted for by the supposition that the I is an intellect with a concept

[261] *SE* IV:36, 38
[262] *SE* IV:34, 37
[263] *SE* IV:36, 40

of its own real being. Representations of things can precede the things themselves, and indeed how this is possible is the fundamental question of practical philosophy.[264]

This can only work because of the subject-object distinction opened up in the prior part of the text. The I-as-intellect intuits that it is the very same thing as the I-as-willing. From the outside vantage point, it would always seem as though they were one and the same, but the I-as-intellect needs to discover this for itself, and it discovers it by realising that it always has been – that is, it is absolute. The I 'tearing itself away from itself' is the moment of consciousness that is linked to the realisation that the I is identical with itself as willing. But because willing is hereby brought under the 'sway of the concept' it has the potential for freedom.

The I intuits itself as the tendency to absolute activity – it has this tendency as its object, and insofar as it does this it posits itself as free. Freedom is said to be 'possessing the power of causality by means of mere concepts' and Fichte says that Kant gives an 'excellent nominal explanation' of freedom as 'the power to begin a state absolutely'.[265] The question is then how freedom is possible. The reason why Kant's explanation remains nominal is that this 'how-question' has not been answered. It can only be answered, we are told, via a genetic deduction of freedom, which is what we have just been engaged in. The possibility of beginning a state absolutely is possible because the I originally intuits itself as a tendency to absolute activity for the sake of absolute activity, and in a sense exists in advance of its nature. Importantly, Fichte tells us that this was only possible because of the distinctive vantage-point provided by the *Wissenschaftslehre* – that we start not with an investigation of being, but with the activity of the I.[266] Here we have a hint that the partly practical motivations for the *Wissenschaftslehre* may yet yield an entirely non-practical conclusion which is viewed as not conditionally true, that is, not merely true if one wants to hold on to morality and freedom, but one which is just true.

The final section of the first deduction is taken up with trying to see how the I becomes conscious of the tendency towards absolute self-activity. This is different from the previous section for two related reasons. The first is that in the prior section we, as philosophical audience, were 'mere spectators of a self-intuition on the part of the original I. What we established was not something we ourselves had thought, but something the I had thought'.[267] This is different to the current section because here we will not begin with an original reflection. The other difference is that before, the reflection was taken to be possible, but here, we will

[264] *SE* IV:2, 7-8
[265] *SE* IV:37, 41
[266] *SE* IV: 37-8, 41-2
[267] *SE* IV:39, 43

see the proof of its possibility. Fichte's aim here is to show that the I has this intuition of itself as a tendency, and that this intuition of itself as a tendency is a self-positing of itself as free.

There is however a slight peculiarity here. Recall that Fichte said above that the tendency toward absolute self-activity was not a mere force, power, faculty, or drive. Here, there are two things which seem at least somewhat in tension, one more so than the other. The first is the theorem that the I also posits itself only as a power.[268] The second is that the tendency is a drive.[269] I shall deal with the former first, which is the less problematic one. Fichte's reason earlier in the text that the tendency is not only a power is that 'the latter [ein Vermoegen] is nothing actual but only what we think of as preceding actuality, in order to be able to integrate the latter into a *series* of our acts of thinking'.[270] Here the same word is used, and we are told that the I 'posits itself only as a power'. The proof of this proposition is as follows:

1) The intellect is 'absolutely self-determining, nothing but *pure activity*…incapable of any determination through its nature and essence'
2) Inclinations and tendencies and drives are not pure activity and are determinations of natures (hidden premise)
3) So no inclination is possible 'within the active force that stands under the sway of an intellect, to the extent, anyway, that it does stand under the intellect's sway'
4) So the tendency must be something that is compatible with pure activity, or it must be able to be under the sway of an intellect.
5) A pure power 'i.e., only a concept of the sort to which some actuality can be connected by means of thinking' is compatible with pure activity, and is the only concept of its kind which is like this.
6) Therefore the tendency is a power.

The worry is that the characteristic of a pure power in premise 5 is the same as the characteristic of it a few pages earlier, but now Fichte seems to be saying that this characteristic is the reason why the I can intuit itself only as a power, whereas earlier he said that it was the reason why the I could not so intuit itself. So what is going on here? I can only think that there has been a subtle transition in the way the I has been conceived between here and previously. Before, we are told that the concept of a power makes up but does not exhaust the I. Now we are told that it does exhaust the I. The shift is because of the argument that the I must in some way exist in advance of its nature, which is tantamount to saying that it has no nature. Because it has no nature in this technical sense (there is clearly something that it is to

[268] *SE* IV:38, 42
[269] *SE* IV:40, 43-4
[270] *SE* IV:28 33-4

be an I, so it has a nature in at least one sense), it turns out that the aforementioned categorical base on which this dispositional part of the I would have rested is not present. The tendency to absolute activity, which constitutes the essence of the I, is only a power, insofar as it is a determinable without a determinate. It is this kind of structure that allows the I to in some sense exist in advance of its nature, which also allows it to posit itself only as a power. Before, Fichte did not want to prejudge the issue of whether there would turn out to be a categorical base or not.[271]

The second problematic thing about this section is related. Fichte says earlier, just after saying that the absolute tendency could not be merely a faculty, says that it could not be a drive, because a drive is a power that, upon the presence of a certain condition, will necessarily cause an action or reaction of some kind.[272] Again Fichte has an injunction that we must not prejudge the enquiry by claiming this of the tendency to self-activity for its own sake, but here he seems to think that it is acceptable to claim that it is so. This suggests that a similar transition to above has taken place. There are two key phrasings in the statement of the theorem is that the posited tendency 'necessarily manifests itself as' and that it does so 'to the entire I'. The importance of these is roughly this: the tendency is not in itself a drive, construed as an analogue of a force that necessarily entails a happening based on a condition, but it manifests itself as such. The fact that it reveals itself to the 'entire I' as this, rather than merely to the I-as-intellect, is also important. Before, we were dealing with the I-as-intellect, not the entire subject-object.

The stated goal of this section, however, was to show how the I becomes aware of this tendency, which is necessary for my purposes here, because it will also show in a genetic fashion how the I comes to posit itself as free. Fichte's argument for this can be stated as follows:

1) When the I reflects upon itself (making an object of itself in thought), it posits or judges that the things reflected upon are identical to itself.
2) One of these objective features is a drive
3) The I therefore self-ascribes a drive

The argument is meant to establish that the I can self-ascribe the intentional content of its thought about itself. This is the mechanism of intellectual intuition. Because the tendency to self-activity for its own sake can be described in some sense as manifesting as a drive (Fichte's earlier worry about this notwithstanding), it means that when the I intuits the objective

[271] This he says at IV:29, 34
[272] IV:29, 34

to be itself, it can self-ascribe this found drive. The difference between this drive and all others is that the tendency to self-activity cannot drive with necessity or 'mechanical compulsion'.[273] This is because, though acts of intellectual intuition prior in the text, the active force of self-activity for its own sake has been 'brought under the sway of the concept'.

Perhaps what is going on is something like this: Fichte claims that this is not merely a drive, or force, or power, or, as we would say, a disposition. This is because these are nothing 'actual' but 'only what we think of as preceding actuality'. In other words the I cannot be explained as having a mere disposition as its essential character, because we would then be left trying to explain how the disposition is activated. Rather, it has to be that the dispositional character of the I is *always* activated, and thus the I is always active. This thesis itself requires more investigation. I shall therefore devote some space to elucidating this idea.

The Fichtean view that the I is always active was distilled from the discussion over Aenesidemus' response to Reinhold and Kant. On Aenesidemus' Humean view, what Kant and Reinhold are doing looks like dogmatic metaphysics. This is because they survey the mind, find that it has acts (such as sensation, perception, or at the most general level, representation) and then posits a capacity for these acts. While this may sound in some sense innocuous or trivial, in the sense that to say that 'X phi-d' we have to think that 'It is possible that X would phi', it belies a more substantive question. We are then confronted with the following. Aenesidemus raises the challenge: how is this inference different from the dogmatic inference from one's having an idea of an object to there being an object in the world?

In order to avoid this challenge, Fichte conceives a middle path. On the one side, we have the realists, represented by Reinhold, who will claim that the inference from a usage of power to a thing that has the power is legitimate. On the other, we have phenomenalists like Aenesidemus, who claim that a faculty is nothing other than its manifestations. The crucial move that Fichte makes, as I see it, is to agree with the phenomenalist that the faculty is nothing over and above its manifestations, but then argue that the phenomenalists counterfactual – 'were there to be no manifestations, there would be no faculty' is true, but never satisfied, because the antecedent is always false. This gives Fichte a conception which is at once non-reifying but also non-phenomenalist.[274] This is hinted at when Fichte claims that if one thinks that I have to exist in order to do things is 'maintaining that the I exists independently of its actions', which is to treat consciousness as a type of object.[275]

Fichte then argues that this consciousness of the tendency could not be consciousness by way of a feeling, but must be by way of a thought. With an argument by elimination, Fichte argues that the only ways the I could become conscious of the drive would be by feeling or by

[273] *SE* IV:43, 46
[274] See also Breazeale (2013a, 408).
[275] Fichte *FTP,* K29/H29, 112.

thought, but it cannot be by feeling, because feeling requires a dependency of the I on something outside it, whereas in this we are considering the I and its moments as all inner to it.[276] The thought which is consciousness of the tendency must be determinate, in order to be a thought of a particular thing at all. But the ordinary ways in which thought is determinate do not apply here. Thought is ordinarily determinate either by reference to its object (i.e. thought represents a determinate object) or because of another thought (e.g. an inference from one proposition to another). It cannot be the former, because the drive is a determination of the entire I, which is subject-object, or ungraspable. It cannot be the latter, because 'the I thinks itself in this act of thinking'.[277] Fichte seems to be operating with a principle that if an x is ungraspable, and x has a property F, then even though Fx is graspable, F is not itself graspable immediately.[278] An analogy is helpful. It is often suggested that God is ungraspable, but that we might be able to grasp some of his properties analogically or metaphorically, and so we cannot grasp them immediately, as to grasp them immediately would in a sense to grasp God, but we can get some distance toward explicating them. Similarly we cannot grasp the subject-object = X that is the entire I, and so cannot grasp its determinations or properties immediately, but we can get some distance towards comprehending them via philosophy – the intellectual activity of replicating the process of consciousness via genetic deductions.

This thought must therefore be unique in this regard: it is determined by itself – and with this 'thinking as such becomes absolute with respect to its form; we obtain a series that commences purely and simply with a thought that is itself not grounded on anything else and is not connected to anything else'.[279] This thinking must be unique, because otherwise consciousness would be inexplicable. This is similar to what was found earlier in the text. There is something like the self-ascription of a willing going on, but a more primordial version of that activity. The tendency manifests itself as a drive and the I intuits this as itself, and this intuition is intellectual.

As for the content of the thought, Fichte proceeds to consider the two aspects of the I – the subjective determining the objective and the objective determining the subjective. This is due to the limitation of the ungraspability of the entire I. if we were able to grasp the subject-object as such, then we would be God. So the best we can do is see how the subjective and objective features of the I interact, or in Fichte's parlance, reciprocally determine each other. This reciprocal determination appears to be a kind of thinking of the one in terms of the other. So when we think of the subjective aspect of the I as determined by the objective, we think of the subjective, as it were, translated into the vocabulary of the objective. The subjective as

[276] *SE* IV:44, 47
[277] *SE* IV:45-6, 48-9
[278] *SE* IV:48, 50
[279] *SE* IV:46, 49

determined by the objective yields that the intellect 'has to give itself the unbreakable law of absolute self-activity'.[280] The objective as determined by the subjective is then 'determined, produced, and conditioned by this subjective power'.[281] The reciprocal determination of these is 'the legislation in question manifests itself only on the condition that one thinks of oneself as free, but when one thinks of oneself as free, this legislation necessarily manifests itself'.[282] Or, as Fichte says a few sentences later, this is 'the *necessary manner* of thinking our freedom'.[283] We should recap. The tendency to self-activity for its own sake manifests as a drive, which gives rise not to a feeling, as would be the case at the level of natural consciousness, but to a thought. This thought is the I-as-intellect intellectually intuiting itself as having the tendency, or self-ascribing the tendency in a completely original act. This thought has as its content that freedom is a kind of law of self-activity.

This is the goal of the deduction, as Fichte states. What Fichte takes himself to have accomplished is to have shown that if we take ourselves to be the kind of things (rational beings) that we ordinarily do, then we can see how there is a deduction to a conclusion that there is a certain way we should act. The very general claim that there are certain ways in which one should act is all that can be deduced from the *Wissenschaftslehre* proper, without beginning on the project of ethics as such. It might be objected that all Fichte has shown, on the most charitable reading, is that if I have certain practical interests, then I must (in some sense of 'must') think of myself as free. He has not shown that I really am free. Fichte is aware that this is the case. One thing we can rule out is the view that the reason why I am driven to think of myself as free is some sort of empirical psychological claim about learned patterns of thinking. It is, if Fichte is right, deeper than that. Thinking of myself as free is (at this stage, at least) a fundamentally normative necessity. It is in some way guaranteed by the hypothetical imperative. If I want to act, then I have to take myself as free, or give up the project of acting (which I cannot do in general).[284] Again Fichte describes this as an act of faith.

As Kant would have it, we are immediately confronted by the fact of reason, which we can take in two ways. We can take it at face value, and then we take ourselves to be free and morally bound (which are indeed one and the same thought), or we can try to explain it by something else, which in Fichte's mind would turn it into an illusion. As above, I think the best explanation for why Fichte thinks this (which is by no means obvious) is that the fact of reason carries with it a claim of absoluteness, and this claim is an essential part of the fact of reason. The fact of reason is not merely 'You are free/morally bound', but 'You are free/morally bound *absolutely*'. Therefore, if we tried to explain the fact of reason, we would be engaged in

[280] *SE* IV:48, 51
[281] *SE* IV:48, 51
[282] *SE* IV:48, 51
[283] *SE* IV:48, 51
[284] Thus Korsgaard (2011: 3) 'acting is our plight'

explaining it away, which would be tantamount to trying to show that it is illusory. Another way of saying this is that, for Fichte, the fact of reason is two claims. One claim is a claim to the existence of moral bindingness. The other is a claim to moral authority. For Fichte, if we attempt to explain the fact of reason, we undermine the claim to moral authority, but moral authority is so central to the fact of reason that we end up explaining it away as illusion, and if moral authority is dethroned then moral bindingness is weakened. This is all well and good, but still remains conditional. What is needed for Fichte's aims is a real deduction that free action occurs.

3.2) Freedom in the second Deduction of the System

The stated task of the second deduction is to deduce the reality and applicability of the principle of morality and freedom with which Fichte concluded the first Deduction. I take it that Fichte is thinking that there is a possible view according to which the moral law is known but is in some way inapplicable to the world. Perhaps such a view would have moral realism for its metaphysics (that there are moral facts) but error theory for an epistemology (that there is no way of applying these facts in judgement). Such a view separates the possession of moral concepts from their use.[285] The important points for my purposes though are the parts of the deduction in which Fichte claims that freedom is a theoretical principle for the determination of our world, and the second theorem of this deduction, which states that the self-ascription of the power of freedom is not possible without an actual exercise of freedom. Contrary to the order in which these appear in the text, I shall take the latter first.

3.2.1) The proof of free will

The second theorem of the second deduction is as follows: 'A rational being is equally unable to ascribe to itself a power of freedom without finding in itself an actual exercise of this power, that is, an actual act of free willing'.[286] The claim is that a 'merely ideal representation' of a free

[285] Historically speaking, the view that comes closest to approximating this description is Luther's. Luther thinks that there are moral facts (grounded in the will of God) that we know – they are 'written on our hearts' but our wills are so corrupt that the moral facts can become somewhat distorted. Luther would not want to say that they are inapplicable (for right judgement is possible on Luther's view), but it is his stated view that merely knowing the moral truths is insufficient for applying them, because of sin. This is bound up with Luther's subtle discussion of the so-called theological use of the law. See for example the 1517 *Disputation Concerning Scholastic Theology* (1957) and the *Bondage of the Will*.

[286] *SE* IV:83, 82

act of willing is not possible without the perception (and therefore the actuality) of an act of willing.[287] The argument is obscure, and there is a lot to unpack. The first major claim is that the 'universal form of all free willing' is 'an absolute transition (accomplished by means of absolute self-activity) of what is subjective into what is objective'.[288] So a representation of a willing is a representation of an absolute transition from what is subjective (which is the concept of an end) to what is objective (the knowledge that I will that-p). Fichte thinks that it is important that, unlike in most other instances of basic perception, the perception does not arise from a feeling. He says, 'I cannot say that I feel my willing…for I can feel only a limitation of my activity, whereas my willing is the activity itself'.[289] Again, intellectual intuition is our access to willing.

The proof itself, as distinct from these preliminary remarks, is again obscure, but I think it roughly is the following:

1) In the I, the subjective and objective cannot be present apart from one another
2) Consciousness arises necessarily from the conjunction of these two elements
3) The representation of an act of willing is a merely subjective one
4) Given (1), there must be an objective counterpart
5) This can only be willing
6) So there is a perception of willing, which means that there is an episode of free willing

This does not appear to be quite right, however. It misses out Fichte's important clarification or precisification that the 'pure form of what is objective' is produced 'only by an act of willing'.[290] The pure form of what is objective for the ideal representation of willing is something characteristic of willing, namely, the knowledge (via intellectual intuition) that I am willing. Willing is originally objective, or 'the distinctive character of willing lies in its pure objectivity'.[291] Less formally, Fichte's idea is that I could not have the thought that I am willing unless I was able to intuit that a willing was taking place. The merely ideal representation, or concept, of willing (and freedom) has a possession-condition which is the intuition (perception) of an act of willing, that is, a transition from a conceptual construction to a willing, which is accompanied by immediate knowing. So we can restate the proof taking this into account as follows:

1) The merely ideal representation or concept of willing is, *qua* representation, a conforming of the subjective to the objective

[287] *SE* IV:84, 83
[288] *SE* IV:86, 85
[289] *SE* IV:86, 85
[290] *SE* IV:87, 86
[291] *SE* IV:85, 84

2) This conforming can only happen if the pure form of the objective is intuited, which must originally happen within the I
3) The pure form of what is objective is provided by willing
4) Therefore willing must be intuited, and there must be an actual instance of willing

This, I think, is a better approximation of Fichte here. Premise 3 is supported by earlier comments that Fichte has made about the nature of willing and its originally objective character. It is originally objective because it is accompanied by a knowing of it, which is a determinate knowledge of something distinct from the intellect. It might appear that this argument in a sense proves too much. One could substitute any representation for the representation of the concept of willing in premise 1, and the argument would still go through. Here I think this is a result that Fichte desires – that we could have no representation apart from willing. This is a position that Fichte has argued for in the 1794/5 *Wissenschaftslehre*, in the so-called striving argument. Ultimately speaking, I think that for Fichte everything rests on the intuiting of our own willing. For example, Fichte says that willing 'is something with which all human beings are very well acquainted; moreover, as the philosopher demonstrates, this is also the starting point of all consciousness and is that through which alone all consciousness is mediated'.[292]

Fichte rephrases the argument in a summary paragraph. Here the argument runs as follows:

1) I originally intuit my activity as an object, and as determinate
2) To intuit my activity as a determinate object is to intuit a willing
3) There is no possibility of a seems-is distinction at the level of intellectual intuition (hidden premise)
4) So there is an intuition of an actual willing, which is free because it has its ground in thinking

For this argument, the hidden premise is absolutely key. I think it is clear from Fichte that he cannot accept there to be a gap between seems and is for intellectual intuition. Intellectual intuition is the unity of the subject and the object, or the knowledge that when I think of myself, I think of the thinker, and the thought is identical with the thinking. If this could be illusory, then there would need to be a further act of unity, and so on.

If this argument (in any of its forms) can work, then Fichte has given us the proof that we needed to secure the representation of ourselves as free at the end of the first deduction. That is, if it is true that I could not have the representation of willing in general without there being

[292] *SE* IV:88, 87

a genuinely free willing, then it is not merely true that I could not be self-conscious without representing myself as free and morally bound, but that I could not be self-conscious without actually being free. However, it is important to see the modesty of Fichte's claim here. He has disproven one version of determinism, if this argument is successful. The determinist thesis that thought is epiphenomenal, or always secondary to what actually happens in the will, is the thesis that he has proven false with this argument. Why is this so? It is because if Fichte is right, then what is going on is that willings are accompanied by free constructions in thought (via reflective judgement, perhaps) of the concept of an end, which then causes the willing, or the transition between the subjective to the objective. But it could still be true that in another sense (the sense that really matters to us) we are all unfree, because it could still be true that the state of the universe is such that I could not have acted otherwise, even if from my point of view there was no encroaching features that determined my choice. We shall have to wait for Fichte's account of freedom as a theoretical principle for determining the world, or the account of freedom and nature, to see any response to that type of determinism. But first, we should look at the third theorem, which is relevant to the second, and continues the work of the proof of the second.

The third theorem is important for my purposes because it could be imagined that, for example, an evil demon intervened at all times and made my willing (which I have an intellectual intuition of) non-efficacious, and so the willing would not result in action. The worry is about the causal story of action; one might grant that we have intuition of an actual episode of our free willing, but think that if we had no direct causal impact on the world then free will would be only an honorific title with no meaning. It is this threat that Fichte wants to dispel with this theorem, which runs: 'A rational being cannot find in itself any application of its freedom or its willing without at the same time ascribing to itself an actual causality outside of itself'.[293] Another way he puts this is to say that it is claimed that 'I cannot find myself to be actually willing, without finding something else within me…consciousness originally arises just as little from the representation of a sheer, impotent act of willing as from the representation of our power of willing as such'.[294] In other words, just as the actual act of willing was needed, in addition to the conceptual representation of this power, we also need that act to be efficacious in the world.

The activity of willing is a determinate one – we do not will in general, but will this or that. The question is – how does this determinacy arise? Activity in general cannot determine itself, so it must be determined through its opposite, something outside it that limits the activity. I cannot intellectually intuit the manner in which I am limited – this would be to be able to intellectually

[293] *SE* IV:89, 87
[294] *SE* IV:89, 87

intuit things outside the I – so I must feel them in experience. This is the "check" [*Anstoss*]. There is a something which blocks my activity, which is also a condition of there being determinate activity. Another way of saying this is that the I, in finding itself to be causally efficacious, finds itself to be finite. This must be so because if activity were always unrestricted, then there would be no limit, and without the limit, there could be no check – but then without the check there is no reflection or positing, which means that the I would not be self-conscious. So willing, if it is to be determinate, needs to be resisted by something outside of it (which is felt), and then overcomes this. That just is to be efficacious in the sensible world. So the threat of being locked in, as it were – being free but inefficacious, has been dealt with. We shall come back to this in more detail later when we look at Fichte's theory of causation, with a view to established whether or not he signs up to agent-causation. For now I will look at the account of freedom and nature.

3.2.2) Freedom and Nature

It is sometimes claimed that the fundamental problem of the philosophy of mind is the placement problem: How does mind fit into the world? Fichte reverses this problem. We can instead see him as asking: How is it that the world fits around the I? In other words, Fichte does not take nature as basic and then ask how it is that minds could arise out of nature. He takes the mind as basic and then asks what kind of world there would have to be. In doing so, Fichte gestures at an account which can give an answer to how it is that we are natural beings as well as rational beings, and so can himself give the answer of how the mind fits into the world.

The first thing that is necessary is that the I have a physical aspect, or physical power. This we have already seen some of in the above section. For self-consciousness to exist, there must be things which resist activity and which can be overcome by activity (because only activity can resist activity). This means that there is a general transcendental condition on the nature of the world – the world must be such that action is possible. Interestingly, Fichte also argues in the following way:

1) A free being acts as an intellect – that is, its actions proceed from concepts of ends (or concept of an effect)
2) Anything in the concept of an end is intelligible to an intellect

3) Therefore the concepts of ends must in general be such that the ends (or effects) are thinkable, or 'What is to be brought about must therefore be so constituted that it can at least be thought of by an intellect'.[295]

This needs further amending, because Fichte makes it clear that it is not merely the capacity to be grasped by an intellect which is important here, but the capacity to be grasped as contingent. The event or object in question (which is the content of the concept of the end) must be thought of as being possibly existing or possibly not existing.[296] One thing that is not an end that we can have, and thus cannot fall under the domain of the moral law, is moving something such that it is not in space. This is not even thinkable. Fichte, as a good Kantian here, says 'I cannot, for example, will to posit something outside of space, for I cannot think of anything outside of all space'.[297] From the preceding we can infer that 'if something is a product of freedom, then it is contingent'. The conjecture is then this – is it possible that the reverse is true; that 'if something is contingent, then it is a product of freedom'. There are two senses in which this can be taken. One is the theoretical sense; one is the practical sense. In the theoretical sense, it means that 'if something is contingent, then it is a result of the productive imagination or the activity of the intellect'. This is true, but not what is meant here – *everything* is a result of the activity or synthesis of the productive imagination[298], and so taking it in this sense would be too weak. The practical sense is to think that what it would mean is that everything is already a result of our real physical activity. But this is just plain false. What we need is some principle that navigates the middle ground here.

This principle is provided by this: 'that our freedom itself *is a theoretical principle for the determination of our world*'.[299] The result of the search for the meaning of how we find out what we ought to do has resulted in this. There are again two ways to read this principle, a stronger and a weaker. The weaker is suggested by what Fichte says immediately afterwards. He says 'Our world is absolutely nothing other than the Not-I; it is posited only in order to explain the limitedness of the I, and hence it receives all its determinations only through opposition to the I. among other predicates, however, or rather, more than any other predicate, that of "freedom" is supposed to pertain to the I'.[300] This suggests the weaker version of the principle: that (given the principle of reflective opposition, or *omnis determinatio est negatio*)[301] if the I is free then the Not-I is not free. This means that the world is 'determined' insofar as

[295] *SE* IV:66, 68
[296] *SE* 66-7, 68
[297] *SE* 67, 69
[298] This is of course a transcendental faculty, named for its analogue (the reproductive imagination). Fichte clearly distinguishes the two, like Kant, but does not follow Kant in the detail of how they work.
[299] *SE* 68, 70
[300] *SE* 68n, 70
[301] Fichte reads this principle in line with a traditional rationalist: that all predication is exclusion of some other predicate(s), with the caveat that this is primarily a principle of thought.

we have excluded one disjunct of a contradictory pair of predicates from application to it. But this is a very minimal sense of determination.

Whilst it is certainly true that Fichte wants this conclusion, he also, I think, wants the stronger conclusion. This is the conclusion that if freedom determines the world, then we can be sure in advance that the world is in some sense friendly to our moral purposes. In other words, if the moral law demands such-and-such, then we can be sure that the world is such that such-and-such is at least possible.[302] Fichte's language here could have been taken from the *Critique of Judgement* (which he knew well, insofar as he began a commentary upon it).[303] In the published introduction, Kant draws some technical distinctions between domain and territory. A concept's territory is the area in which it has correct application. A concept's domain is that over which it can legislate.[304] One way of viewing the stronger reading of Fichte here is that it is not merely that the moral law or freedom has the sensible world as its territory, but the sensible world is also the domain of the moral law, or freedom. This can only be true insofar as the moral law's domain-relation is different from that of theoretical reason, otherwise there would be no gap between ought and is, and the moral law would 'legislate' for empirical data, which of course it does not.

Part of Kant's problem is that the faculties of reason and understanding have different legislation over the same domain. The understanding has the sensible world as its domain insofar as it is theoretically legislative (that is, insofar as the pure concepts of the understanding operate in a synthesis with given intuition and the productive imagination). Reason has the sensible world as its domain insofar as it is practically legislative. Kant is concerned because the two domains cannot be unified. They cannot be unified because the concept of nature only allows us to represent nature 'as mere appearances rather than as things in themselves, whereas the concept of freedom does indeed allow us to present its object as a thing in itself, but not in intuition'.[305] This is the source of the 'immense gulf' fixed between the sensible and the supersensible 'so that no transition from the sensible to the supersensible (and hence by means of the theoretical use of reason) is possible, just as if they were two different worlds'.[306] The problem is compounded insofar as the first of these 'worlds' (the sensible/the domain of the understanding) cannot have an influence on the second (the supersensible/the domain of reason) but the reverse is supposed to be possible. Hence, Kant

[302] Whether this means that Fichte signs up to 'ought-implies-can' is unclear. He clearly does agree with it in some sense, but famously argues that we have infinite obligations, so he cannot sign up to it fully.
[303] GA II/1. Fichte also probably refers to the introduction of CJ when he says that Kant 'maintains that there is no bridge leading from the sensible to the supersensible world' and that the WL 'has no trouble in constructing such a bridge – 'the intelligible world is the condition for the possibility of the world of appearances; the latter is constructed on the basis of the former' (K124/H115, 260)
[304] Kant *CJ* 5:174
[305] Kant *CJ* 5:175
[306] Kant *CJ* 5:175-6

says ‘it must be possible to think of nature as being such that the lawfulness in its form will harmonise with at least the possibility of [achieving] the purposes that we are to achieve in nature according to laws of freedom’.[307] Thus Kant’s project in the work is to find a way of justifying this thought, which of course finds its terminus in the rational hope that the moral project is possible, a hope justified by the experience of beauty and the sublime as well as rational purposiveness in nature.[308]

Fichte’s setup allows him to jettison Kant’s attempt at solving this issue. This is because in showing that freedom is at once a theoretical as well as a practical principle for the determination of the world, he has pointed the way to a solution to Kant’s problem which does not involve the metaphysics of Kantian transcendental idealism (chiefly things in themselves) nor reliance on rational hope given a justification by regulative ideas of beauty and purposiveness. Recall that the threat of determinism which remains for Fichte is the determinism which says: “Given that agents have no control over the laws of nature nor over the beginning conditions of the universe, every event necessarily follows from these, and so all freedom is illusory, even if I must think myself to be free”. We can view Fichte’s aim, then, in the freedom and nature sections to be an attempt to defeat this type of determinism by undercutting its assumption – that there could be a world which simultaneously contains agents like us and laws which necessarily determine future events in the manner required. This is the fulfilment of what I said above – that Fichte reverses the placement problem.

That freedom is a theoretical principle for the determination of the world is only a schematic and highly abstract principle, however. What we need is some more determinate way of fleshing out the account in a more perspicuous manner. In order to see this more fully, we should turn to other texts from the same period – the infamous *On the Basis of our Belief in a Divine Governance of the World Order* as well as the *Wissenschaftslehre nova methodo*. I shall begin with the latter.

In the final section of the WLnm, Fichte tries to complete his synthesis of the aspects of consciousness.[309] It is a hallmark of this later work that it is oriented toward showing the fundamental interconnectedness of consciousness and its conditions, rather than presenting them as a series.[310] One of these interconnected parts is the consciousness of myself as a body. We have already seen some of the build-up to this from the System of Ethics. Here we confront this very intriguing claim:

---

[307] Kant *CJ* 5:176

[308] The details of Kant’s project in the third Critique will need to be overlooked here. For an overview, as well as some reflections on its immediate reception, see Gardner (2016)

[309] Discussions of Fichte on naturalism and nature already in the literature include Girndt (1992), Lütterfelds (1992), Breazeale (2014).

[310] E.g. *FTP* K227/H247 ‘This whole synthesis looks like a framework or a series; but consciousness resembles a circle’.

'I and my body are absolutely one, simply looked at in different ways. I as "pure I", in its supreme purity, and I as "body" are entirely the same. The distinction that appears to us is based entirely upon the difference between these ways of looking at [the same thing].'[311]

Whilst it is probably not correct to attribute to Fichte here a very strong thesis of identity – insofar as the I and the body could differ in modal properties[312], for example – he clearly thinks that there is a very close relationship between the I considered as subject and the I considered as object, that is, as body. For my purposes, the important claims that Fichte makes are regarding what he calls 'articulation'. Articulation is a property of bodies. We already know (from the *SE* and *FTP*) that in order to be self-conscious a thinker must perceive an act of willing and that this act of willing must be directed outside itself, and that I must be able to see that my activity is efficacious in a world. There are of course conditions on how this must be so. The condition on this activity directed outside itself is that it must come from an articulated body. An articulated body is one that has part-whole relations, such that the parts are movable or modifiable relatively independently of the wholes. Fichte says:

'I move my entire body: taken by itself, my body is a whole; in relation to nature [as a whole], however, my body is only a part. I move my arm: this too, taken by itself, is a whole; at the same time, however, my arm is also a part of a larger whole, namely, my body, etc'[313]

A more concrete example will help. It is characteristic of part-whole relations that parts can have properties that the whole does not. I can tap my finger or raise my arm without moving other parts of my body. These basic actions do not require any other apparatus (aside from the necessary biological features such as nerves and so forth, which I do not have willing control over). What is characteristic of articulation in bodies is that 'it is up to freedom to decide what shall be treated as a part and what shall be treated as a whole'.[314] What Fichte means by this is that if I decide to move my arm, I treat my arm as the locus of the action – it is the whole that I willingly move. It has parts, which I willingly move by virtue of willingly moving the whole, but do not willingly move the parts *per se*. I could, of course, willingly move them independently of raising my arm. One worry we may have here is that Fichte's account seems

[311] *FTP* K234/H256
[312] However it would also be incorrect to attribute to Fichte some kind of "neutral monism" later held by James and Russell, according to which the body and the I would be both modifications of some third thing, neither mental nor physical. This would, in Fichte's eyes, be a kind of dogmatism, because it would be a version of insisting that the I must be a kind of active thing, rather than activity itself.
[313] *FTP* K235/H257
[314] *FTP* K235/H257

to make it so that basic actions are individuated by the part-whole relationship. That is to say, if I raise my arm, and thereby treat my arm as the locus of this basic action, and raise a finger on that hand simultaneously with the arm, then I appear to have done one thing, but Fichte seems to need to say that I have done two things simultaneously. This worry can be somewhat dissipated by the fact that these putatively different acts are indeed independent of one another – I could raise my arm and raise my finger without doing the other (apart from the sense in which I must raise my finger along with my arm, but willingly raising my finger would be to basically act, and the rising of my finger in virtue of the rising of my arm is not a basic action).

The fact that the I is an articulated body is important in two related respects. The first is that it means that the body has to be a part of nature. The second, and more important, is that a free being is in some sense created by nature. These are both to be expected. Recall that Fichte's guiding principle here is that 'freedom is a theoretical principle for the determination of the world', or, as I have put the point, the world (nature) must be such as to allow for the emergence of free beings. Now, if free beings are embodied (in whatever sense) that body must be such as to be created by nature, which means that the free being must be such as to be created by nature.[315]

Linked to the concept of articulation is the concept of organisation. Whilst articulation is a feature of a body that it has in virtue of part-whole relationships, organisation (which a body must have if it is to have articulation) is a label that refers to the fact that the body as a whole is a 'real whole'.[316] Organisation is said to 'follow from' articulation, and by this I take Fichte to mean that one can infer from the presence of articulation to organisation, but not the reverse, i.e. organisation is a necessary condition for articulation and articulation is a sufficient condition of organisation.[317] The body, and everything else in nature, is a real whole because the boundaries of the things are 'also nature and [are] posited by nature'.[318] Nature itself is therefore '[not only] an organising power, it is [also] organised'.[319] Fichte also says:

'{Nature as a whole} must necessarily be an organised whole, because individual organised wholes are possible within nature, and these are made possible only by means of the entire force of nature. Individual organised wholes are simply products of the organisation of the whole universe'.[320]

---

[315] *FTP* K236/H258 '...this body is a product of nature. I.e., nature produces itself, in conformity with mechanical laws….since this body is something merely discovered, i.e., is [part of] nature, the articulation [of the same] cannot be anything but the product of a purely natural law'
[316] *FTP* K236/H258
[317] *FTP* K236/H258
[318] *FTP* K236/H258
[319] *FTP* K237/H259
[320] *FTP* K238/H259

In other words, we can say that for Fichte, in these very general reflections on what nature would have to be like given freedom, nature contains within itself the possibility of its immanent mechanical laws producing something with freedom. Fichte will not want to go any further in establishing what nature must be like (the issue of consternation between Fichte and the early Schelling). Indeed, Fichte even says at the very end of section 19 (after reviewing the results) that:

'One must grasp reason as a whole, and then one will find that no conflict is present and that nature is quite absolutely posited by itself as absolute being, opposed to nothing but the absolutely posited *I*. this is the perspective that has to be adopted by natural science'.[321]

The philosopher can only progress so far with investigations into nature. The *Wissenschaftslehre*, being the science of science as such, gives a place to all other sciences and their first principles. It does not replace these sciences. Natural science in general then receives its first principle from the *WL* – that nature is absolute being and should be investigated as such – but no determinate result is prefigured by the *WL* itself, apart from the general stipulation that the world must be such as to allow freedom. But even this is not a burden on the scientist. To see why, we should consider that Fichte is very clear that in his mind natural laws are necessary or immutable statements. One job of the scientist is to discover these natural laws. The fact that they have some necessity about them is not something that is going to worry Fichte. As a speculative suggestion, it could be that what Fichte has in mind is something like the following: the statement "Natural laws are necessary" is one that is distributive – it applies to each natural law individually, but not the set of natural laws. The statement "Natural laws make room for freedom" applies to collectively. One suggestion that could be used to further determine this is that individually the laws feature in some efficient-causal explanation, but collectively they are teleologically directed. This can be true even if no one particular law is itself teleologically directed. Our suggestion above then can be reworded thus: Nature contains within itself the possibility of its immanent mechanical laws, which are jointly teleologically directed, producing something with freedom.[322] There is a question here about whether the possibility just indicated requires some kind of guarantee or fact to enable it to obtain, or whether it could be a mere brute fact about the world that it contains the possibility (or actuality) of rational agents.

---

[321] *FTP* K240/H261. When Fichte says 'opposed to nothing' in this passage I take him to mean that the WL by itself cannot tell us what kind of predicates must truly apply to nature, in the spheres within nature itself, or those that are investigated by natural science.

[322] This is in some ways similar to Guyer's suggestion for Kant's views on the antinomy of teleological judgement (2005, 336)

Recall that we were led here by investigating Fichte's idea that freedom could be a theoretical principle for the determination of the world. If this is so, and it is true that articulation and organisation are 1) products of freedom and 2) necessary conditions on self-consciousness and 3) in some sense identical with (or at least very close to) the I, then, I think, we are landed with the problem of saying how it is that this can come about. The first answer is that it is a mere brute fact, that there is no explanation of how it is that nature comes to produce free agents.[323] The worry I am trying to articulate here is the following. If 1), 2), and 3) are all true, then it could be cosmic accident that rational agents exist at all. This is worrying not just because of possible tensions with the general program of idealism, but because of the worry that this might itself make morality a kind of cosmic accident or contingency. Therefore there is a question of what underwrites or guarantees the possibility or probability of the existence of organised beings (thence rational agents) in the first place. This question needs to be disambiguated from a related but distinction question. This second question is the question of what made the first rational being(s) come into existence, as distinct from the question of the probability of organised beings. Fichte's answer to this will be that there is no explanation, because in some sense the rational being's coming into existence is an act of that rational being (given that the I is activity).

This, together with his account of why actions cannot be explained at all to the extent that they are free (which I will elucidate in chapter 5) provides an answer to this question and undercuts a main assumption behind part of it. The main question I want to answer here is captured in the following: How is it that the possibility of nature producing beings with freedom by its purely natural laws comes to obtain? Answers to this question could be 1) that it is a brute fact 2) that Nature explains it 3) that Freedom or Morality explain it. Fichte's answer must be, I think, the third. This is because it needs to show us how it can be that Freedom and Nature can be united.

Following Kant, Fichte thinks that Freedom and Nature cannot be united by conceiving of Freedom as like Nature, but only by making Freedom the primary aspect. The sense in which Freedom is primary is that there always remains the hard core of subjectivity, the absolute I, which is not enveloped by nature. This puts Fichte in contrast to Schelling's Identity Philosophy and *Naturphilosophie*. For Schelling, Freedom and Nature must have both originated from an indifference point which is their identity. For Fichte, it must be that there is Freedom 'in the beginning' as it were.[324] There are different ways he could approach an answer. I will pursue

[323] I use the phrase 'produce free agents' as shorthand. Strictly speaking, nature does not produce free agents, because the activity that is the I is itself a product of the I, by a free act of positing and intellectual intuition. But nature does produce the body (which have features of organisation and articulation that are themselves products of a purely natural law.

[324] This debate is brought out in the letters collected by Vater and Wood (2012). Especially important is Fichte's letter to Schelling reproduced on 48f.

what I see to be the most promising strategy, which involves bringing in Fichte's views on God. We shall therefore conclude this section by investigating some aspects of Fichte's conception of God insofar as it bears on this question of how freedom arises within nature.[325]

3.2.3) The role of God

Our main text now is the late Jena essay *On the Basis of our Belief in a Divine Governance of the World-Order* (DG). This essay can be viewed as a Kantian moral argument for the existence of God. However, it also precipitated an upheaval in Fichte's professional life – the *Atheismusstreit*. I shall run through what I take to be the main thrust of the argument, before turning to the importance of the text for the question of freedom and the guarantee that nature is purposive in some way.

The argument is roughly as follows.

1) I find myself morally constrained
2) I could not distance myself from this moral constraint without alienating myself from myself
3) Therefore, I must posit the end of this moral constraint (call this the Highest Good) as binding on me
4) Positing the end as binding on me means that I must try to bring it about
5) Trying to bring it about relies on faith that it will be brought about
6) This faith is in the moral world-order itself (the Provident Order)
7) This Provident Order is itself God
8) Therefore, in order to posit that the moral end is binding on me and to try to achieve such an end, I am committed to belief in God.

There are several controversial steps in this reconstruction, but the most important is the identification of the Provident Order with God. Fichte has a couple of things to say that will alleviate some of this peculiarity. The first is a general principle on the nature of concepts. For Fichte there is a principle to the effect that the possession conditions on concepts constrain the conditions of acceptable use. This is important for the Atheism controversy, because Fichte denies that concepts such as 'person', 'substance' can apply to God, at least in the

[325] In terms of contemporary positions, I place Fichte as a near-naturalist, akin to Baker (2017). A near-naturalist is one who denies the metaphilosophical positions of the naturalist (that science is the arbiter of what exists and what doesn't) but is happy to yield to the scientist about the description of the natural world alone.

same way that we use them in the ordinary (or even philosophical) sphere. Notoriously, he says that 'the concept of God as a particular substance is impossible and contradictory'.[326] What motivates Fichte to say this, I think, is the thought that because the genesis of our concepts lies within certain boundaries, it cannot be the case that a concept can legitimately be applied beyond these boundaries. For example, in the Nova Methodo, Fichte says that because God alone is 'holy', by which he understands that God alone does not have the desire which, if fulfilled, would contravene the law (a necessary part of the genetic story). He then says that therefore 'no consciousness can be ascribed to God, or at least this is incomprehensible to us'.[327] Another similar passage from the *Grundlage* is 'For the deity, that is, for a consciousness in which everything would be posited by the mere fact of the self having been posited (though for us the concept of such a consciousness is unthinkable), our Science of Knowledge would have no content, since in such a consciousness there could be no other positing whatever...'.[328] In this passage it seems that Fichte is saying that finite consciousness is distinguished from God's consciousness because finite consciousness needs a process to draw it out, it is not present all at once, just given with the original positing of the self. For finite selves, there needs to be further acts of positing, which are free (in the *Grundlage* this is of course the second act of positing (the Not-I) and the positing of mutual determination). Therefore Fichte might have in mind some thesis according to which how we come to grasp a concept, or how we think a concept, constrains how we can apply that concept.

I do not want to attribute a very strong version of this thesis, which would be most implausible. A very strong version would be "The possession conditions of the concept always determine the whole scope of the concept" – which would end up with us having to say that we cannot apply the concept of 'tree' to trees that have not yet been directly experienced. But Fichte does want some general condition on concepts such that the way we come to know their rules of application constrain the circumstances in which those rules themselves apply. The most obvious example is the summons and the concept of a person. Fichte's argument in the first part of the *Foundations of Natural Right* is that we come to know that there are people outside of us through actual interaction with those people. Now, because this is the case we can then genuinely apply the concept to things that are such that, were we to have stood in that relation, then we would have been summoned by them. Here is then another reason why we could not apply the concept of person to God directly, as the concept of person is tied up with the fact of being-summoned, which is never true of God.

---

[326] *DG* 188, 152. This may strike one as odd (especially in light of Spinoza) but it is a common claim amongst Church Fathers and scholastic philosophers as well as reformers. For example, St. Augustine *Confessions* III.6.42 & IV.16.29, St. Thomas Aquinas *Summa Theologica* Pt.1, Q.3. Art.5.
[327] K145/H137, 295
[328] SK I:253, 224

If this goes for 'person', then it goes for every concept, and thus we cannot make good theoretical application of concepts to God. But we can have this moral argument, which gets us some of the attributes that the theoretician will want. For Fichte, the most important are probably Providence, Life, and Incarnate.[329] A general way to sum up this is that Fichte does not subscribe to 'Perfect Being Theology'. PBT takes its cue from a way of thinking about God such that God is thought to be like us but maximised. We know a few things, God knows all things, we have some power, God has all power, and so on. Fichte might think that this is how we come to acquire a concept of God in the first place, but it has no place in the philosophical account of what God is like. I now need to turn to the notion of the Provident Order and the sensible world.

Fichte says:

'My entire existence, the existence of all moral beings, and the sensible world, as the common theatre of our actions, thereby obtain a relation to morality. There thus opens before us an entirely new order, of which the sensible world, with all of its immanent laws, is merely the passive foundation'.[330]

This might strike us as somewhat odd. Fichte speaks here of the sensible world being the 'passive foundation' of the moral. He also says that the sensible world proceeds along its own path 'in order to constitute a sphere for freedom'.[331] We have met this kind of talk before. I suggest that these are to be read as only one side of the story. Fichte tells us this much himself when, in the next paragraph, says very different sounding things, 'from the transcendental point of view'. That seems to imply that the passage above is to be taken from the ordinary view. If so, we have a solution to our issue. From the ordinary viewpoint, it does indeed seem that nature is primary and morality secondary. But from the transcendental viewpoint, the viewpoint of the *Wissenschaftslehre*, we can see that the situation is reversed, hence the 'truly real element in things, the true basic stuff of all appearance' being 'the material of our duty made sensible'.

[329] These come out more clearly in Fichte's developed philosophy of religion, the lectures *The Way Toward the Blessed Life*. I will only mention these insofar as they bear on the problem of free will, which they will do when it comes to a possible problem of divine foreknowledge. For a speculative reconstruction of Fichte's philosophy of religion, see Kinlaw (1992) where he argues that Fichte's conception of God is "kenotic" (a theological term referring to St. Paul's use of the word 'kenosis' in Philippians; usually thought to mean something like 'self-emptying'). For example, Kinlaw says What Fichte is claiming is not only that the only thing that can arise outside God is *Dasein* or kenosis – since God's *Dasein* is God's kenosis, but that only in God's self-othering can God exist or have actuality' (45-6).

[330] Fichte *DG* 353/184, 146

[331] Fichte *DG* 353/184, 146

However, this would be too simple. It cannot be, for Fichte, that the transcendental view is straightforwardly in tension with the ordinary view. Rather, the transcendental view must *vindicate* (to some extent, at least) the ordinary. There could be two ways in which common sense could be said to be vindicated by the transcendental view. One way would be to say that common sense is vindicated just in case all or most of the specific claims (the letter) of common sense are preserved by transcendental philosophy. Another would be to say that common sense is vindicated just in case the worldview, or the mode of being (the spirit) of common sense is preserved by transcendental philosophy. I take it here that Fichte would agree to the latter – that is, that transcendental philosophy is at liberty to diverge from common sense on individual claims and to trump those common-sense claims with claims of its own, so long as these latter claims do not depart from the spirit of common-sense (whatever that might be).[332]

There is, on Fichte's view, some conflict between the spirit and the letter. He says that 'the surest evidence' that we accept the letter of our philosophy but 'it remains a foreign addition which does not belong to us' (that we reject the spirit) is that 'we accept errors in life which we have refuted in our study'.[333] So there can be tenets of common sense which are trumped by philosophical reflection. Spirit is also a kind of ability or capacity. A philosopher who has spirit in this sense has the ability to raise to consciousness 'deeper feelings underlying those other feelings which relate to the physical world'.[334] In other words, natural consciousness comprises a set of feelings, some of which are about the physical world, and some are not. The ones that are not (e.g. feelings of my own freedom or moral feelings of a certain kind) tend to be less visible to the thinker in question than those that do relate to the physical world. Why is this? It is because the I is action, and so to become acquainted with it 'means to become acquainted with its acts',[335] but to become acquainted with its acts is to be at first acquainted with the objects of those acts.[336] Connected with this is Fichte's thesis that in intuition or thinking about an object the I is transparent to itself, or that the I 'disappears into the object'[337] The philosopher who has spirit in this sense is the one who is able to move from the standpoint of ordinary life (thinking about objects) to the standpoint of philosophy (thinking about the thinker). At the level of philosophy, for example, it will turn out that for Fichte the world is very much unlike what we take it to be from the standpoint of life. I take it that, like

[332] Fichte *DG* 348/178, 143. This is also a distinction present in Kant, especially the *Groundwork* and other texts which rely heavily on common rational moral cognition.
[333] Fichte *Concerning the Difference between the Spirit and the Letter in Philosophy* (*CDSL*) (1988[1794]) 27, 214-5
[334] Fichte *CDSL* 9, 199. In *Appeal* this is said to be within reach in some way of everyone. The thought will be something like "Such a life cannot possibly be my true vocation, there must, oh, there must be another entirely different lot for me!" (424, 100).
[335] Fichte *CDSL* 10, 200
[336] Fichte *CDSL* 11, 201
[337] Fichte, *WLnm* K29/H29, 110-1

Kant before him, Fichte is an empirical realist and a transcendental idealist, and thus accepts the claims of life (that there exist objects outside of me, for example) so long as these claims are not taken unrestrictedly.[338] So Fichte can have it that both the claims of common sense, or life, and the claims of philosophy, or the *Wissenschaftslehre*, are acceptable, and that the latter vindicates in spirit the former, even if the claims of philosophy look like they undermine the claims of common sense.

We need to look at this passage as giving us a multiple relation, rather than two separate relations of foundation or grounding. It clearly is the case that nature can 'ground' freedom by being the sphere in which free action and morality can be realised. But there is another sense of grounding at play. This can be found if we focusing on the 'passive' rather than the 'foundation'. The 'goal of reason' Fichte tells us, can only be actualised by the efficacious acting of rational beings. It therefore looks as though the sensible world can be the passive foundation of the moral in virtue of being the only possible candidate for the sphere of moral action, but the moral world provides the basis of the sensible by being the very reason why the sensible exists in the first place.[339]

If this is right, then we can see why God can be said to guarantee the teleological proposal above. Recall that my final gloss on Fichte's view about the relation between freedom and nature was this: "Nature contains within itself the possibility of its immanent mechanical laws, which are jointly teleologically directed, producing something with freedom." Now we can include God and God's providence in this formula, which could read: "Nature contains within itself the inevitability of its immanent mechanical laws, which are jointly teleologically directed, being part of the Provident Order which is itself God, producing beings with freedom". I use the word 'inevitable' here, because I think Fichte wants to get at the thought that it in some sense necessary that free beings (those that contribute to the moral world order by living in it and through it) exist, even though there was a time at which free beings did not exist (before the emergence of humankind).[340] In other words, at the root of things there exists the normative order, which implies at some stage the existence of rational creatures like us. I therefore reject the identification of the absolute I with God, as they play fundamentally different roles for Fichte. God, as the moral world order, is part of the explanation of the finitude of free rational beings whose bodies are the products of natural laws. The absolute I is the most general features of subjectivity insofar as there exists such a thing as subjectivity at all.

[338] For example, see *WLnm* (K24/H27, 106) for the claim that 'When the idealist says "outside of me" he means "outside of reason"; when the individual says the same thing, he means "outside of my person"'.
[339] This is a key claim of axiarchism. See Leslie (1979, 6).
[340] I used the word 'emergence' here in a neutral sense. I do not want to ascribe to Fichte 'emergentist' views such as those of Hasker (1999), though Fichte may be sympathetic to some of what Hasker says.

Another way of putting this is to say that whilst it is an aim of Fichte's idealism to explain all reality in terms of the I, it is a mistake to think that this means that the I must *precede* everything. There can be backward-looking explanations (such as the genetic deductions of consciousness) and forward-looking explanations (such as the existence of the I, and its connection to morality, explaining why anything exists at all), both of which centre around the I.[341] None of this, however, crosses the boundary into either pre-critical metaphysics or absolute idealism. Fichte can have the views that I have described and still hold that this is all 'for the I' – that is, the I is still primary in philosophical explanation, rather than, as the absolute idealists hold, an appearance of some underlying unity. The interpretation and explanation of nature is therefore still fundamentally different to Schelling's project of *Naturphilosophie*.

As a final note, we should see that this answer to the Kantian problem of the unity of nature and freedom leaves open the task of bringing nature in accordance with freedom, which is part of the rational and substantive end of agency.[342] This is the task which was expressed in the *Grundlage* as bringing the not-I in agreement with the I, or responding to the demand that they be absolutely alike, which is why all consciousness has striving as a pre-theoretical condition. The account I have given of the moral world order guaranteeing that natural beings will arise which will become rational beings does nothing to answer the issue of how freedom and nature are to be reconciled in this practical sense, which is of course the task of real ethics and the practical projects of law and politics.

This concludes my account of Freedom as it has its place in the system of the *WL*. We have seen that Fichte argues not only that I could not be self-conscious unless I were conscious of the moral law and freedom, but that I could not be self-conscious unless I were actually free. This freedom is the absolute spontaneity of a causal series, underdetermined by prior events in time. In a sense, freedom remains inexplicable. Ultimately it relies on intellectual intuition, and the immediacy of that kind of intuition. We have also seen that Fichte has an ambitious account which tries to solve Kant's problem of Freedom and Nature, as well as an account which brings a Provident God in to underwrite that solution. I will therefore turn to the second part of the thesis, in which I look at particular topics adjacent to freedom but not directly

[341] I therefore think it is apt to think of Fichte as close to the axiarchism of Leslie (1979), as one who thinks that the reason why anything exists at all is because there is a root 'oughtness' in things.
[342] Michelle Kosch has (2014, 2015) defended a view according to which self-sufficiency is to be spelled out wholly (or at least mainly) in terms of the domination of nature. But this, to my mind, is incomplete – there are also other kinds of theoretical ends as well as spiritual ends which make up what it is to be self-sufficient, even if the domination of nature is a key part of freeing the rational agent. The same mistake, I think, is made in an essay critical of Fichte by Dryden (2013). Ware (2018) has recently argued that Kosch's picture is incomplete, and Langlois (2017) presents a fuller picture of Fichte to which I am sympathetic. Kosch's more recent view (2018) seems to be more general, and instead of domination of nature, Kosch emphasises the independence from interference of nature, especially in her reconstruction of Fichte's arguments from the *Foundations of Natural Right* (158-9) and *Lectures on the Vocation of the Scholar* (164).

addressed by the system. These are agent causation (chapter 4), evil (chapter 5) and divine foreknowledge (chapter 6).

4) Freedom and causation

According to Kant, freedom is a form of causation. We see this in both the Groundwork and the Critique of Pure Reason. Fichte signs up to this view. Immediately then we can reject as an interpretation of Fichte a view which we can call simple indeterminism. This would be the view that free choices are uncaused in some way. It is commonly said by defenders of determinism and compatibilism[343] that if this is what libertarianism requires then that means that libertarianism has to be some sort of supernatural doctrine, and thus can be rejected as requiring too much. But Fichte would agree with this. Given that freedom is a form of causation then we cannot be simple indeterminists. But then what kind of causation is it? Fichte agrees with Kant that it is a causation of the will. Questions arising here might be: 1) How does Fichte think about the will (and is it different from the Kantian picture, i.e. that the will is practical reason)? 2) Does this require a commitment to agent-causation? We should look at these in turn, but first we should look at what Fichte might say against a Newtonian or some similar believer in universal determinism, and see what Fichte thinks about the relationship between causation and time, before we answer these questions.

Fichte thinks there are at least three types of causation in the world. The first and most general of these is mechanism. This is conceived as the motions of bodies imparting motions to other bodies. The second and more restricted is biological causation. This belongs properly only to those beings that are organic. The third is rational causation, or causation by the will, had only by those organic beings that are agents. Fichte says of the first two that each member of the series is determined in advance, but this fails to obtain with the last.[344] As Kosch notes, this is not to say that Fichte thinks that organisms are determined by mechanical laws alone (as a Cartesian would have it). The organisms are determined, but by biological laws, including the *Bildungstrieb*, or drive toward organisation. The final product, he says, of nature is a drive, which is a mere power, not yet a causal force. One question that arises here which needs to be dealt with is how Fichte can conceive of agents as free and not determined in advance by other things (unlike biological organisms) but also as a particular kind of product of nature (like biological organisms). Fichte ultimately explains this by his views on intersubjectivity and the theory of the summons. But this puts the problem back one step. There is still the question of how it is that there are rational natural beings at all. Part of the answer can be given by saying

[343] Hobart (1934), Ayer (1982)
[344] *SE IV:* 128.

that the I reflects on its drive, and in doing so 'tears itself away' from its prior state and is then free. This happens through the summons. I am not aware of this tearing away, as it is a condition for consciousness in the first place.

This could also be explained by thinking about the transitivity of causation. For causation to be universally transitive, it would have to hold that for any causal relation, if X causes Y and Y causes Z, then X causes Z. This is commonly appealed to when an explanation of action is required. In a model where reasons for actions are thought of as causes, we can say that the action was caused by a reason which was caused by some event in the world, thus, we might want to say that the action was caused by the event in the world. Fichte would deny this transitivity, and expresses it by the notion of a leap in consciousness. It should be noted that universal transitivity gives very counter-intuitive answers – namely that for anything that ever happened, it was true that the big bang (or God, or whatever) caused it. So we might want to have some restricted version of transitivity, and that is what Fichte has – of course universal non-transitivity is a very peculiar idea, and Fichte does not want to sign up for that either. We shall see in the next section, however, that thinking of causation as non-transitive is not really accurate to Fichte's views. It would also not be what Fichte needs for freedom. It cannot be just that some event X (my having a desire) causes me to perform an action, and my having a desire is caused by the object of my desire, but the object of my desire cannot be said to cause my action. That would have to be true, but it would be insufficient. It would also need to be that my having a desire does not cause my action. My action, in order to be free, must be uncaused, or spontaneous. So whilst it is true that Fichte sometimes characterises freedom as a break from the past, he does so in such a way that he means something different to non-transitivity.

We should now turn to looking at causation and its relation to time.

4.1) Causation and time

The main passage that needs taking account of in any overview of Fichte on causation is his perplexing statement which seems to say that causes and effects are, as a rule, simultaneous. The passage is as follows:

'It is, for example, certainly true that the events in the world are connected with one another as causes and effects. But no time whatsoever is contained within the concept of causality, for the effect is absolutely simultaneous with the cause.'[345]

[345] WLnm K186/H188, p368

‘Everything that exists is an effect of some cause and is simultaneous with this cause. But what about this cause itself? It is, in turn, an effect of some other cause, and so on, ad infinitum. No time comes into being in this way; everything is present in a single stroke.’[346]

The first passage comes when Fichte argues that it is just not possible to understand time and causation unless one conceives of time as a mere form of intuition in the Kantian sense. In a sense, we can see Fichte running a similar argument to Kant regarding geometry or mathematics to prove the synthetic a priori nature of temporal intuition or temporal judgement. Fichte says that the concept of causation contains no predicate which has any temporal bearing.[347] This means that time and causation must be synthesised to give us causation across time. But if this is true, then it means that all perceptions of temporal succession as a key aspect of causation are due to the imposition of form of the subject, rather than in the things themselves.[348] So time must be a form of intuition, because the concept of causation and our experience of causation do not match up unless this is assumed. Fichte says that the origin of time as a part of experience arises through ‘the original act of intuiting an act of thinking’ and this act of thinking is the positing of the cause, and then the effect, and then proceeding from one to the other.[349] In Fichte’s view, however, it cannot be merely any cause and effect which is noticed in this way, but it has to be the concept of an end and the willing of that end, or the deliberation and the willing of an end. To the transcendental viewpoint, this all happens within a single stroke, but ordinary consciousness views it as a determinable and a determinate, which is the most basic way of describing the temporal relationship.

But then in the latter passage Fichte seems to shift from talking about conceptual relations to talking about real objects. Of course, he could just be spelling out a direct implication of the concept of causation. He might be saying that if we just consider the causal relations on their own, then it is impossible to see how temporal succession comes into play, because, as stated above, the concept of causation contains nothing of time itself.

The trouble for free will should be clear. If all causes and effects are simultaneous *in re* then there is an illusory aspect of time, and we have a position not unlike the dogmatist, who thinks there is a kind of rational grounding world and then a world which we perceive. But this cannot be Fichte’s view. It also is in tension with his other views concerning the nature of character

[346] WLnm K186/H188, p368

[347] If this is what Fichte means, then he is following Kant, who thinks that the concept of causation can be spoken about independently of the schematism, which is the synthesis of the pure category of ground and consequent with the intuition of time. Kant says in a practical context, at least, we can think of causation without temporal predicates, *Metaphysics of Morals* 6:280n

[348] In McTaggart’s phrasing, it is compatible with thinking this that there could be a C-series of purely rational ground and consequent relations, even if there is no B-series that is instantiated in time by this rational ordering.

[349] WLnm K186/188, p368

and the role of God as I have reconstructed it. Therefore for textual coherence reasons I suggest that Fichte is concerned solely with the concept of causation and what the concept would entail, as opposed to real relations between existing objects. The fact, then, that we perceive temporal succession is as a result of the fact that time is a form of intuition, and that this meshes with our concept of causation. Fichte has to reject Kant's account of the schematism if this is to be so, but this is also to be expected. The schematism in Kant is designed to solve the problem of how our pure concepts and our intuitions could have anything in common when one is entirely pure a priori and one is empirical and a posteriori. This is not a problem that would arise for Fichte, because it is entailed by a radical separation of form and content, which is foreign to Fichte.

A question which is relevant to the question here is how I come to have the concept of causation at all. This is important for Fichte's genetic account, because the centrality of willing needs there to be an account of how it is that I become conscious that I make things happen in the world. Of course for Fichte, the story of the possession-conditions on the concept of causation has to ultimately show how it is that we construct this concept ourselves.

But the main worry of this section has been averted – that is, that Fichte might commit himself to some sort of McTaggartian C-series view, which seems to be in tension with both his wider view and the reality of free will or the open future. We have seen how we can read the text so as to not commit Fichte to such a view. Now we should move on to discuss how Fichte thinks about causation and the will, with a view to answering the question of whether he has a Kantian view and whether he subscribes to agent causation. It is my contention that, insofar as Kant's view is that the will is practical reason, Fichte actually separates himself from Kant on this issue. For Fichte, the will needs to possess a causal power, which appears at the transcendental level of consciousness as well as the empirical level.

4.2) The Causal Power of the Will

Fichte states that willing is a causal power, or that causal power attaches to willing. But the question of what it is exactly that we discover when we find ourselves to be willing or exercising our causal power is one that, as we discovered before (in chapter 3) cannot be one that we know discursively, but is founded on intuition. Moreover, it is 'impossible to image a will without at the same time imagining some impulse, some application of power. Willing is a genuine inner efficacy, an act of affecting oneself'.[350] The thought that I will something is equivalent to thinking that 'the present state of my feelings, or the object that is presently in a certain

[350] FTP K126/H118, 264

condition, ought to become other than it is'.[351] The puzzle is to say how the will can bring about this change in an object, or in my condition (which for Fichte would be two ways of asking the same question), without interrupting the unity of consciousness. The unity of consciousness would be interrupted if the thing or state in question immediately just responded to the wills demand that it be other than it is. If it were not a continuous or gradual change then the I would be completely changed, insofar as its acts of positing at the first moment would have no connection with those of the second moment. So Fichte needs a way of spelling out how willing can occur without jeopardising this unity. The answer lies in the causal nature of the will and its connection with temporal succession.

He claims that the unity of the manifold of feeling requires the causal power of the will. But the claim is really stronger than this – he also claims '*only* if we think of the manifold as unified in this way are we able to think of the will as exercising causality'.[352] The unification of the manifold requires a series of counterfactual dependence of feelings upon other feelings. He says 'Every possible *B* one grasps must be viewed as conditioned by some *A*. (Conversely, *A* might well exist even if *B* did not…'.[353] From this arises temporal succession of feelings. The line of argument here might be the following: The object of a willing is always a determinate series of acting and sensing, or to change the present state of my feelings. This means that there must be a gradual movement from this state of feeling (call it A) to the willed state (B). But a gradual movement from A to B requires a manifold of feeling. A manifold of feeling requires that there be no two contiguous feelings that are opposed to one another.[354] So a manifold of feeling requires a relationship of dependence of the state of my feeling at one moment, B, to another moment, C, and so on, right back to the starting point, A. This requires that our ideal activity produce something objective and persisting. But this is to posit temporal succession. So the knowledge that we have of all persisting and actually existing objects is on the basis of feeling, but we only have connected feelings insofar as we have a will. After this genesis of the concepts within the I, there arises an objective and sensible world. Fichte says:

'The concept of force is {the mediating concept,} the bridge between the intelligible world and the sensible world, and it is by means of this concept that the I goes outside of itself and makes the transition to a sensible world. By means of this concept, the I represents itself to itself as an object and connects its own consciousness to an objective world. In this way, I become an

[351] FTP 126/H117, 264
[352] *FTP* K128/H119, 267 emphasis added
[353] *FTP* K128/H119, 267
[354] Recall that when Fichte says 'opposed' he most often means 'of a different kind' or 'different in nature'.

object for myself, an object of perception, and a sensible world is connected for me with this object that I become.'[355]

Note that the argument is not supposed to establish that all feelings we ever have are conditioned in this way. That would be to assign too much to the domain of the willable. What Fichte takes himself to have shown is that the way experience and consciousness of a sensible world arises is by the will extending over time and the ideal activity bringing feelings of that duration into a unity – recall that the stated goal of the chapter was to show how the concept of force originates in the mind, not the concept's application conditions. This is the argument that Fichte alludes to in section 13 as the argument which shows how feeling is dependent upon willing. After this excursus, Fichte returns to the stated problem of 12, which is to show how it is possible to give measurements to exertions of force or energy. I will turn now to discussing another argument, which appears in the System of Ethics. This argument is an argument to prove the theorem that 'A rational being cannot find in itself any application of its freedom or its willing without at the same time ascribing to itself an actual causality outside of itself'.[356] The argument begins with the idea of activity. Activity is the 'simplest intuition; it is sheer inner agility and absolutely nothing more'.[357] Activity is to be determinate, if consciousness is to be possible, but it can only be determined by means of its opposite – limitation. We can formalise the rest of the argument like this:

1) One can only think of a manifold of activity (several particular actions) if one thinks of determinate, or limited, activity
2) One can only think of determinate, or limited, activity if it is felt
3) The activity can only be felt if it is actual, and is a certain quantity (or quantum) of activity
4) A quantum is anything that is sensible that fills a moment of time.
5) Anything that fills a moment of time is an infinitely divisible manifold, and thus the perceived limitation must be a manifold
6) So the I must be posited as related to this manifold as eliminating and breaking though boundaries and resistance
7) This is what it means to ascribe to the I causality in the sensible world.

The key moves in this argument are between premise three and four, and premise five and six. The thought is that I can only be self-conscious if my activity is determinate, and the only way it is determinate (because I cannot have intellectual intuition into how I am limited) is

[355] *FTP*, K131/H121, 271
[356] *SE* IV:89, 87
[357] *SE* IV:90, 88

through the world. But this means it is determinate through sensible intuition, which means it must actually occur, because sensible intuition is based on feelings – simple sensations or apprehensions, the appearance and the reality of which are together; by which I mean that there is no sense to be made of appearing to feel X but not actually feeling X.[358] Anything that can be posited through the senses, or anything which is encountered in sensible intuition can be described as a quantum or some kind of quantity. This intuition of something through the senses fills a moment or unit of time. But if it fills a moment of time, given that time is an infinitely divisible manifold, or a manifold that has no basic or fundamental unit, then the thing that fills a moment of time (the sensible intuition) must itself consist of an infinitely divisible manifold. This is intuitive, insofar as an event takes up a passage of time, if the time can be divided infinitely into segments, then this goes for the event that fills that time as well. This explains the transition between premise three and four. The transition between premise five and six is more difficult. This is because it just is not clear why it is inferable from the thought that the perceived limitation must be an infinitely divisible manifold, to the thought that the I must be posited in relation to this as eliminating or breaking through boundaries. There are two distinct ways to read this inference. The first way is that Fichte is giving a sort of reductio argument, showing that any thought about how I am sensibly limited turns out to show that I already, in this most basic act, exercise a kind of causality on the world. The second way is that Fichte is giving a more straightforward modus ponens, on which limitation and activity are in a way reciprocal concepts, and given that we already know that the will is an activity, we can know that it is limited, but the way in which it is limited, or the way in which it is known to be limited, turns out to reveal that it must already, in being limited, possess a causal power.

I think the key is that we cannot intellectually intuit the ways in which our activity is limited. Because we cannot do this, any thought of my activity being limited needs to be in some sense discovered. But it can only be discovered if I already have exercised my causal power to act in that way. In other words, I think the second reading of the argument is the more plausible one. The manifold of boundaries or resistance is because in each moment of the action that we can think, there is both a limitation (it takes effort to move my hands or legs in a certain way) but there is also an overcoming of this limitation by sheer effort or will. The causal power of the will is evident in actions that we perform because once we see that the I is a kind of activity, which must be limited in some ways, which can only be then known through investigation, we can also see that I would only know that my will has a causal power if I would feel it to be limited, because it is only through overcoming those limitations or boundaries that I know myself to be willing. This may appear to be in tension with Fichte's first theorem – that

[358] This is similar to the by-now familiar Kripkean (1980) point that there is no sense to be made of the idea that I can judge that I am in pain and not be in pain. To judge that I am in pain just is to be in pain.

I find myself as myself only as willing. But these are in agreement, insofar as Fichte subscribes to a view of consciousness on which it is a circle, or that each condition on self-consciousness is linked synthetically to every other. This could also be phrased as Fichte's view that the conditions of self-consciousness are simultaneously conditioned by the first principle or starting point and that they in turn condition it. In other words, the starting point (the awareness of myself as willing, or the bare fact of self-positing) gets reinterpreted in the light of the later conditions and deductions, until the genetic story of self-consciousness has closed the gap between what consciousness must transcendentally be and how it appears in ordinary consciousness.

4.3) The Account of the Drives

This has all been an account of what Fichte takes to be the reasons to think that one could not have a will or be self-conscious unless one had a causal power to affect the world. Indeed, part of what it means to have a will is to have such a power. But this is distinct from the question of how rational causation works, which is the question I now turn to.

Despite his disagreement with Kant on the nature of causation as a category and its relation to time, Fichte does agree with Kant that the will is a kind of causation by concepts. For example, he says that because I have to posit myself absolutely as active, 'I cannot describe this activity otherwise than as the causality of a concept'.[359] Absolute activity can only be presented in the form of 'causality by means of a concept'.[360] This kind of concept must be the concept of an end. So the activity that characterises rational beings must be a causality by means of the thought or grasping of some end to be effected.

Fichte's view involves a theory of the drives. There are three relevant drives here – the natural, the pure, and the ethical. The natural drive is the most important for questions of causality or causal power, because, on Fichte's model, it is the source of all my causal power. The pure drive has no causal power of its own, but appropriates the causal power of the natural drive, and the mixture of these two – the causal power of the one, and the constitutive goal of the other – are brought together to form the ethical drive.[361] We should go into more detail here to see how exactly it is that the I has causal power.

Fichte says that the final product of my nature (or my natural being) is a drive.[362] When the intellect 'brings this under the sway of the concept' as he says, the I is not determined in

[359] *SE* IV:9, 14
[360] *SE* IV:9, 14
[361] A formally similar, but substantially very different, drive theory is present in Schiller, who has it that the form-drive and the matter-drive a synthesised in the play-drive.
[362] *SE* IV:132, 126

advance by this tendency, but is 'purely and simply indeterminable'.[363] Drives are characterised as having 'no causality in a spiritual being', by which Fichte means no causality *without also* the endorsement of reason, but then of course reason itself is the difference-maker.[364] What Fichte calls 'formal freedom' is the knowledge that whenever I act I do so in accordance with a concept.[365] The natural drive is said to be the source of my causal power, as 'I cannot will anything other than what nature would also will, if it could will', and 'I never actually do anything nor can I ever do anything that is not demanded by the natural drive, because the latter exhausts the entire sphere of my possible acting'.[366] The model seems to be that in non-rational beings, the natural drive is sufficient to move them to act, and they act solely on the basis of what the natural drive provides for them. But in rational beings, the natural drive having been brought under the sway of the concept or appropriated by reason, this is never sufficient, because there is always a further practical question of whether I should act this way or whether it would be best to act this way.[367] To an intellect, that something is desired is not sufficient grounds for acting, or, to put it another way, that something is desired can be sufficient grounds for acting only if the intellect takes it to be sufficient grounds.[368] This can give rise to what might be called the proleptic effect of one's beliefs.[369] As we have seen, Fichte thinks that how one conceives of oneself has an effect on the things one is able to do. If I truly believe myself to be determined, for Fichte, then that will tend toward making me more passive and not able to do things that others who are convinced of their freedom may be able to do.[370] If I am a firm believer in my freedom, then I will tend to be more morally motivated, or even just more able to engage in rational deliberation, as it is entailed by my belief in free will that I am free to choose to do the right action and that I can meaningfully deliberate about what to do.[371]

---

[363] *SE* IV:134, 128
[364] *SE* IV:135, 128
[365] *SE* IV:135, 129
[366] *SE* IV:148f, 141
[367] I may be, but may not be, in conflict here with Breazeale (2016, 26). Breazeale states that after the I has reflected on its natural drives and material freedom, the I is 'capable of accomplishing what nature alone could never accomplish'. If this is read as a claim that the I's reflection on its freedom generates a new kind of power, then my reading conflicts with Breazeale's. But if Breazeale is simply saying that nature could never bring about free actions by free beings, but it is only free beings that could bring about their free actions, then there is no conflict.
[368] Or, as Fichte puts it, the source of the object of the will may be the natural drive, but 'it is given by the natural drive as an object of longing, of desiring, but by no means as an object of the will, that is, as the object of a determinate decision to realise such an object' *SE* IV:159, p150.
[369] I take the term 'proleptic effect' from Herman (2009, 154). It is possible that Fichte might say that this proleptic effect extends to the publication of philosophy, especially moral philosophy, as this is part of 'universal moral cultivation', and this is a duty of all *SE* IV:235, 224. But Fichte also later expresses the view that the belief that one's publishing efforts will go some way to bringing about such culture is mistaken in *Characteristics of the Present Age* (CPA) (1889a, 12).
[370] This seems to be a thought in the First Introduction,
[371] That there is a proleptic effect in belief and the will seems to be supported by some research in social psychology, as detailed in Holton (2009). Pearce and Pickard's (2010) discussion of the 'sick

So we have seen how the natural drive contains all our empirical causal power, or delineates the sphere of possible actions. In order for it to be ethically or morally directed, it needs to be mixed with the pure drive. The pure drive is a drive for self-sufficiency, or absolute independence from nature.[372] Fichte sees a contradiction between what has been revealed about consciousness and morality and the reflections on the natural drive. How can it be true, he in effect says, that a) I have to perform an action which is supposed to spring from rational agency alone, and b) that all actions have their roots or sources in the natural drive? They are united in the 'actuality of acting', and by this Fichte means that the nature of action provides the solution.[373] Action always aims at liberation from nature (as per the pure drive) but we can never gain full independence from nature, so it is not as if we enlist the natural drive to perform an action that it cannot, but we enlist the natural drive to perform actions in service of a task the culmination of which would annul the natural drive. Fichte seems quite clear about this, insofar as the final end of a rational being lies 'in infinity' and the 'ultimate goal of finite reason' is the 'complete annihilation of the individual and the fusion of the latter into the absolutely pure form of reason or into God'.[374]

4.4) That Free Actions Cannot be Explained

How does rational causation work, then? It is the contemplation or deliberation over a certain manifold of actions, all of which, in order to be serious candidates, are possible for my natural drive (which means that they must be possible for my body).[375] Willing then selects one of these possibilities, closing off the rest. Willing is in this context 'an absolutely free transition from indeterminacy to determinacy, accompanied by a consciousness of this transition'.[376] This transition takes place in accordance with concepts (of ends) and in accordance with evaluations. But ultimately, Fichte thinks, rational causation cannot be explained. It is causation because I make things happen, or I cause things, but it is unlike other kinds of causation insofar as it does not admit of nomological form. There are 'canons' of practical reasoning, such as the rule that when choosing C out of A, B, or C, there must be a characteristic (X) by which C is preferred or chosen, and then the reason why this X decided the choice 'can lie only in a universal rule that the rational being already possesses. This rule

role' and why it is not always appropriate is also relevant, as is Pickard (2015) on the ability to do otherwise in the context of addiction.

[372] *SE* IV:149, 142

[373] *SE* IV:149, 142

[374] *SE* IV:151, 143

[375] Of course we could imagine being able to perform actions that we cannot perform, but these are never 'serious candidates' insofar as intending implies believing possible.

[376] *SE* IV:158, 149

must be the major premise of a syllogism that would go as follows: whatever is of such and such a kind (=X) must be preferred to everything else; now C is of this kind; hence, etc'.[377] This rule is a rule or canon of reasoning, insofar as it is a rule that consistent practical reasoning must follow (perhaps with additional clauses about how some other characteristic Y may be present which pulls in a different direction from the action chosen on the basis of X). But on the question of whether free actions can be explained, or comprehended, Fichte says:

'"To comprehend" means to connect one act of thinking to another act of thinking, to think the former by means of the latter. Wherever such mediation is possible there is no freedom but only mechanism. It is therefore absolutely contradictory to want to comprehend an act of freedom. Were we able to comprehend it, then – precisely for this reason – it would not be freedom'.[378]

This can be restated as saying that some agent acts 'for absolutely no other reason than because he wills to do it'.[379] The thought seems to be the following. To comprehend an action is to connect it to a principle or prior condition by which the action would necessarily follow. This might be able to be done a few times (given that we might know that someone would act in such-and-such a manner if confronted with some kind of situation), but cannot be in general true, and the explanation will terminate before a point at which it becomes a bona fide one. Here we might draw a distinction that Descartes draws in the *Meditations*. Descartes says that understanding is different from comprehension – to comprehend is to be able to put one's arms around something, so to speak, in thought.[380] But to understand is to reach out and touch the thing. We can therefore, on Descartes' view, be said to understand many things which we do not, and indeed cannot, comprehend – the obvious case being God. Fichte might have a similar thought in mind – to comprehend, to connect one thought with another one so thoroughly that the former can be thought by means of the latter – is something more than mere understanding. We might then understand actions, or sympathise with them, but not comprehend them. So one might explain why I performed an action because I have a certain end in view, but then equally the question of why I have that end in view is open to question. Even if this end is the moral end of self-sufficiency, the characteristic of the end or action is never sufficient to explain why I chose it, only my choice is sufficient. But then this explains

[377] *SE* IV:179, 170
[378] *SE* IV:182, 173
[379] *SE* IV:186, 177
[380] Third Meditation. Hence also John 1:5, according to the King James version, 'And the light shineth in darkness; and the darkness comprehended it not'.

nothing – this is just to say that I chose it because I did. The explanatory buck stops with me alone.[381]

4.5) Agent Causation?

Under the label of Agent-causalism, we might comprehend a number of positions. There are two which are particularly important, and these are ones which Fichte would object to. The first is the sense that we get from Samuel Clarke and analytic philosophers like Chisholm and O'Connor.[382] This kind of agent-causation requires there to be an abiding self or agent, which is a substance that has the power of acting. Fichte would unequivocally reject this position, because it requires a reified conception of the self, which is itself a kind of activity or agility. Agent-causalism then purports to lead us out of a picture according to which we are subject to the laws of nature and are unfree, but ends up retaining the very root of that picture – that agents are a type of thing.

A second kind of agent-causalism results in the problem that confronted Chisholm – namely, that there is an infinite regress of causation, because on this variety of agent-causalism we are said to cause our own actions. But if to act is to cause an action, then in order to cause our own action we would have to cause the causing of our own action. This is Chisholm's regress problem. This problem is independent of thinking that the self is an abiding thing, which Chisholm also accepts. The question is then whether Fichte is committed to thinking that we cause our actions or whether we cause only the results of our actions. It is clear that he accepts that we cause the results of our actions, because that is a necessary condition of being self-conscious at all – that we exercise a causal power on the world. Fichte would say that my act of will causes something to change. It would however be inapt to say that I cause my act of will. Whilst I have control over my will, and it is able to be directed, my willing to do something is, in the paradigm case, in Fichte's mind, identical with my body moving in such a way. This is compatible with thinking that the movement of my body is in some sense dependent on the movement of my will. Even necessarily identical properties can have dependence relations like this (for example, that water has the surface properties it does and

[381] Compare Lucas (1970) who also thinks that ultimately these questions resolve into assertions that I saw fit to take X as a reason to perform some action. Relatedly, see Temple (1917), who argues that the only way in which explanation could ever stop is if we sympathise with an action, or more precisely, the intention which that action was meant to carry out. Temple applies this primarily to the question of why God created anything at all, but he is clear that it can be used in ordinary situations as well.

[382] Clarke (1717), Chisholm (1976, 1982, 1995). As far as Chisholm goes, the 1995 paper represents a change in view from the earlier book and paper. O'Connor (1995)

that it has a certain chemical makeup are necessarily identical properties but the chemical makeup is explanatorily prior).[383] We can see this at work when Fichte says:

'The driving of nature is concentrated and contained in our material body. This driving possesses no causality in itself but obtains its causality immediately as a result of our will. For the reason stated above our will becomes an immediate cause in our body. We only have to will, and what we have willed ensues in our body'.[384]

That our will is an immediate cause in our body is Fichte's way of expressing the point we would now express by the terminology of basic actions, or acting basically. A basic action is one which I can perform without performing any other action. When I will, I will that something happen, and my body immediately moves to make something happen. The immediacy here is, I think, Fichte's way of framing the point that there is no temporal or causal gap ordinarily between willing that my arm move and my arm's moving. There are two further points to make. The first is that the body is the necessary instrument of causality because only a material body could causally affect material things. It is Fichte's stated view that the idea that thought alone could cause matter to change is incoherent – only matter can change matter. This means that I must be matter, in a certain way – that I must be identical in a way with my body (as we have seen in chapter 3). It would then be odd of Fichte to say that we cause our own actions, because the will is not something material in this way. The will is the aspect of thinking about our bodies, and the body is the correlate of the will.

Secondly, that this relation is immediate is spelled out by Fichte in terms of the body being subject to the power of the will. So the reason that I do not cause my own actions is because in order to be said to cause my actions I would have to stand in a relation to my body in which I do not actually stand. I cause change in the world because there is a gap or divide between me and the world, such that what I immediately will does not as a rule immediately come about. But this is different with the body – if I will to move my body in a certain way then I ordinarily do so, without even the need to form a prior intention to do so – the intention and the doing come about simultaneously, even though we might say that there is a kind of dependence relation between the one and the other. I cause changes in the physical world as a result of the fact that the physical world is not something that is under my will.

So Fichte may be more appropriately labelled an event-causal libertarian. That is, he would agree with AC that we can point to the agent as the source of their action. But how he

[383] For this example, and the broader point, I am indebted to William Mann (2015, 72)
[384] *SE* IV:214, 204

understands agency is in terms of actions.[385] Given that to be an agent is to be a kind of action, and actions are a species of events, then to be an agent is to be a kind of event. So Fichte's view might then be that all causation is event causation, and agent-causation is a subspecies of that. But agent causation is a subspecies of event causation just as physical-force causation or psychological-drive causation are subspecies of event causation. We might think then that all causation is governed by some lawlike connections. In the case of agent causation this would just be the laws of rationality, and more specifically, morality. The difference would be that these laws do not, as natural laws do, explain their instances, and neither do they compel their instances. It is Fichte's avowed position, as we have seen, that free actions are in a sense inexplicable – they cannot be explained.

A further position that we could consider as a candidate for something like Fichte's view is volitionism. On a volitionist account, such as Lowe's, episodes of willing are the most basic kind of thing a free agent can do, and they are uncaused, or spontaneous.[386] Importantly, Lowe says that 'Rather than *cause* his or her own volitions, an agent simply *performs*, or *executes*, or *enacts* them – and this kind of relationship between agent and volition is not a causal one, because it is an "internal" one'.[387] I think this is in line with Fichte's view – that willing and bodily movement is internally related because ordinarily there is no gap between the willing and the bodily movement. In Lowe's phrasing, it can be that the 'mental cause is a cause of the physical events that constitute the wholly physical sufficient cause of e [the event]'.[388] This is also acceptable, insofar as Fichte can say that even though my willing (the mental event) and the bodily movement (the physical event) are the same, there is still a qusai-causal relation between them, because the latter depends on the former. Again Fichte can agree with Lowe when he says that the relationship that I stand in to my body is a *sui generis* one, being 'neither identity nor constitution and yet implying, as those relations also do, spatial coincidence'.[389]

One of the most attractive aspects of Lowe's volitionist account is that it does not need to think of events as causally potent. In fact, Lowe insists throughout that talk of one events causing another is parasitic on talk of one thing's possessing a causal power and acting in a certain way to bring about a change in another thing.[390] This is again congenial to Fichte – although Fichte does not have much to say on this particular issue, it being a mainstay of

[385] One could understand agency in terms of actions if one gave a definition of agency such that X possesses agency just in case X is the kind of thing that can act, as well as be acted upon. This is not Fichte's way of understanding it. One could give this definition and be a substance dualist, for example.
[386] Lowe (2008, 7)
[387] (2008, 7f)
[388] Lowe (2008, 70)
[389] (2008, 168)
[390] E.g. (2008, 4, 7, 122, 138)

analytic approaches to causation, much of what he says on causation suggests that he would find it strange to talk of events themselves having causal power, as opposed to the things acting.

Having seen two ways in which Fichte is not an agent causalist (he denies both a substantial self and that I cause my own actions), we can see one way in which Fichte may be appropriately labelled agent-causalist. This is a sense in which I as an agent, as a whole, bring about changes in the world. For Davidson's (1963) causal view, for example, it is the belief-desire combination appropriately responding to the world which causes my action. But Fichte would say that this separation of a belief and a desire from the agent is artificial – I take my beliefs and desires to be reasons to support my action, but these determinations of me do not by themselves have causal impetus. So if agent causation is just the thesis that agents are the locus of their actions, as opposed to some sub-agential aspect of them, then Fichte is indeed an agent-causalist, but not in any stronger sense.

Therefore, I conclude that Fichte is not an agent-causalist in the sense in which that term is usually used, but is closer to a kind of volitionist view, according to which the will has causal power, but I do not cause my willings or my actions, only the results of those willings or actions.

5) Evil

In this chapter I am going to investigate Fichte's account of evil. The core question here is "How is it possible that we act freely and wrongly?" It is well-known (as I touched on in chapter 1) that Kant's account of freedom and morality in the Groundwork seem to lead to the idea that freedom and morality and unfreedom and immorality line up so tightly that one cannot be acting both freely and wrongly.[391] Clearly if this is the implication then more work needs to be done. I shall focus on Fichte's account of evil as *vis inertia* as presented in the System of Ethics. Firstly I will outline the thesis and Fichte's argument, before turning to defence of it.

Fichte has a particular burden explaining how evil is possible because he thinks that at least some moral knowledge is a necessary condition on self-consciousness. One might then be puzzled, insofar as it would seem to be the most readily available knowledge. In other words, Fichte's story needs to involve some kind of self-deception, which is exactly what we will find. But it is worth pointing out that insofar as the psychologist or sociologist tries to explain present actions with reference to previous patterns of life Fichte will ultimately have to think that their

[391] At least, this is the way Reinhold presents things, but Reinhold is targeting Schmid, who himself was attempting to reject Ulrich's position, which was itself an attempt to make sense of Kant by arguing that the intelligible character of an agent is original and unalterable. See Guyer (2017) for this history.

explanations are wrong, partly because they enter into the (mistaken, Fichte thinks) game of providing explanations of moral and immoral actions.

Fichte first points to a seeming necessity or fatalism. He says that it is 'absolutely impossible and contradictory that anyone with a clear consciousness of his duty at the moment he acts could, in good consciousness, *decide not to do his duty*, that he should rebel against the law, refusing to obey it and making it his maxim not to do his duty, because it is his duty'.[392] The second half of this sentence recalls Kant's own conception of what diabolical evil would be – the aiming at evil for its own sake.[393] I take it that by putting in the phrase 'in good consciousness' here Fichte is suggesting that whenever something like this happens it is always a fault of self-deception, because it would be impossible for it to happen in good consciousness. Indeed this is suggested by the next paragraph – here we are told that it is possible to '*render obscure* within oneself the clear consciousness of what duty demands'.[394] The supposed necessitarianism is then: 1) If the moral law is kept in mind, then one necessarily does as duty demands 2) If the moral law is not kept in mind, then one necessarily does not do as duty demands. This would be a problem (intelligible fatalism, as Fichte calls it) unless the question of whether the moral law were kept in mind is in some sense up to the agent in question.[395]

Here though is a problem – Fichte says that the choice to keep the moral law in mind or not cannot be one that is made with the maxim in mind. It seems that he means that one cannot consciously think something along the lines of "In order that my life go better, I will attempt to forget my duty". Clearly this seems an odd thing to think, insofar as consciously attempting to forget one's duty requires one focus one's attention on forgetting something. Fichte claims that such a thing amounts to the diabolical choosing not to do one's duty in the moment of choice. But it seems that they are not exactly analogous. The first case (the diabolical case) is one in which 1) I absolutely demand of myself that I X and 2) I absolutely demand of myself that I not-X.[396] The second case is made self-contradictory because of two distinct and

[392] *SE* IV:191, 181-2
[393] This is the standard way of parsing Kant's conception of 'diabolical evil', but Wood (2014, 35-7) argues that Kant actually leaves this open, and what diabolical evil really would be is the will having 'an original rational incentive to act contrary to moral reason'.
[394] *SE* IV:192, 182
[395] The intelligible fatalist's error here is clearly explained by Goh (2012)
[396] This is how Fichte explains it. But this seems odd too. Surely this would just amount to, in the moment of choice, thinking that I was previously mistaken about my duty. The oddness arises because Fichte characterises deciding not to do one's duty in the same terms as deciding to do one's duty (i.e. demands to oneself). But the characteristic of deciding not to do one's duty is surely the deliberate relaxing of the demand, or the bargaining down of a demand, rather than the acknowledgement of a separate demand. Some of the oddness is lifted if we pay attention to the phrasing 'that he should rebel against the law, refusing to obey it, and making it his maxim not to do his duty, because it is his duty'. This shows us that Fichte has a very particular case in mind (reminiscent of Lucifer's fall, which is presumably why it is called diabolical). For a good account of what the Lucifer case in Milton shows us about evil, see Midgley (1984) who argues that Satan's

opposed demands to the self, but also because of the fact that I cannot think about thinking about p without thinking about p.[397] In the maxim that I forget my duty, I demand of myself that I forget that which I otherwise demand of myself. Either way, it clearly happens that the moral law is not always within the mind of an agent, and that there are many other things which occupy the mind. Fichte says then that it 'simply happens, just because it happens, absolutely without any higher reason'. The reason why this is so mysterious is because it is something that we do, something that we choose, but not something we do with the maxim in mind or with consciousness. It is an abstraction from the moral law which arises as a sort of parasitic phenomenon alongside our general indeterminate thinking-about-nothing-in-particular and inattentiveness.[398] The question is now how to explain this tendency to obscure the moral law to ourselves.

This is where the *vis inertia* makes its appearance. The question is: how is it possible that finite rational agents remain at a certain level of obscurity with respect to the moral law? The answer to this question would not be a merely negative proposal (e.g. that most people merely lack the knowledge or will to do so) but requires a positive proposal too. This positive proposal is the 'original laziness or inertia'.[399] This pertains to human beings insofar as they are products of nature. Nature itself (and therefore every natural being) has a force of inertia, insofar as it is relatively static and remains the way it is. Given that human beings are products of nature and the body is produced by a 'merely natural law', it follows that finite rational agents will be subject to inertia.[400]

It is worth noting that Fichte appears to reject the thesis of original sin, at least as that is usually interpreted. A classic way of describing the doctrine would be that it is the view that humans are by nature evil, or that the human will is by nature evil or misdirected. Fichte thinks this is wrong.[401] It is not that we are evil by nature – that would be to put too much emphasis on nature and not enough on freedom. It is that we allow ourselves, by choice, to remain largely as we are and not try to effect moral change in our lives. There is also the detrimental

---

purposes are 'parasitical on God's' and that he does not know what he wants to do 'till they find that it will compete with him and displease him'. (135).

[397] This does not mean that thinking the thought 'that p', and thinking the thought 'that I am thinking about p' are not different thoughts – they clearly have different truth conditions, for one.

[398] Many in the Christian tradition also have similar worries to Fichte – how is it that God's law becomes so obscured within our hearts, given that a) at least part of it is 'written on our hearts' and b) the other is revealed in God's Word? The answer for the traditional theologians is, of course, original sin, and more particularly, the noetic effects of sin.

[399] *SE* IV:199, p189

[400] A very similar view to Fichte's was held by one of the Late 19th Century Fichte 'school', Charles Carroll Everett in *Theism and The Christian Faith* (1909). This is detailed by W. W. Fenn (1910, 15), when he says that for Everett, 'since the Spirit is onward-pressing, sin consists in failure to rise to ever higher levels of life. In a word, sin is inertia. Man's real being and destiny is to live and grow in obedience to the immanent Spirit, if he fail to respond to this inner impulse he is in a state of sin'

[401] In the *Addresses to the German Nation*, Fichte explicitly rejects such a view at e.g. (2008, 130)

influence on us of society. Whilst Fichte is critical of Rousseau, he still thinks that Rousseau has got something right – namely, that society introduces a whole host of evils which are not present without society.[402] However, the right answer to society is not, as Fichte reads Rousseau, to return to a more primitive state, or at least to try to get ourselves into situations such that we can more easily avoid temptation (such as the moral advice given by Doris)[403], but to deliberately and wilfully transform these dispositions or liabilities to make us better agents who inhabit a better world.[404]

A final aspect of Fichte's view should be considered. This is how Fichte thinks we should overcome evil. Ultimately, and in keeping with his view that free actions cannot be explained or comprehended, there is not a great deal philosophically that he can say, although he may give edifying or exhortatory remarks.[405] Indeed, his preferred mechanism for coming to self-consciousness that one succumbs to temptation or inertia is to attend religious services which will (presumably) have sermons and readings from Scripture dedicated to just this kind of exhortatory project.[406] Indeed, as Kosch says, the church is the primary instrument of overcoming evil because it is a) the forum in which moral discussion is had with a view to moral consensus, and b) because it is the particular duty of religious leaders to promote others' exercise of their rational capacities for the sake of the good by strengthening moral feeling.[407] But it seems that any venue where the moral cases/virtues and vices and so on are presented concretely will work. Moral conversion, then, is not something that Fichte generally thinks can be achieved by oneself, though of course the effort is primarily directed towards the character of one's own will. This also can go some way to explaining the existence of people who ordinarily are somewhat attuned to the good but also sometimes fall into inertia. This is because we can only ever intuitively grasp the moral law in part, never as a whole, so whilst intellectually we might agree that we ought to perform certain actions, our intuitive grasp over this is lacking.

---

[402] Fichte's criticism of Rousseau is largely in the Lectures on the Vocation of the Scholar. For interpretation, see Clarke (2013).
[403] Doris (2002)
[404] It might then be true that Fichte criticises the doctrine of original sin on the basis that it leads to apathy or laziness with respect to improving one's moral character. But this is based on a more substantive criticism that it presupposes that humans have a determinate "nature" that can be discussed, as well as thinking that something can be good or bad independently of it being a result of a free action. Of course, the traditional theologian, depending on their view, might say either that we are responsible (for seeking grace and absolution) even though we are not causally responsible, or that in some sense all original sin requires is some kind of choice which shows that we are depraved in some way.
[405] Indeed, this is why Fichte often sounds as if he is giving a sermon, especially in the final section ("Faith") of the *Vocation of Man*.
[406] For a striking example, see Romans 1:27-32
[407] Kosch (2018, 72-3). Fichte's claims on the role of religious leaders are at IV:350ff.

I shall now turn to some objections to Fichte's view. For the first, I shall consider a rival view, namely Schelling's. The second will be looking at issues surrounding motivational vs cognitive interpretations of Fichte's view. The third will look at whether Fichte's view can account for the commission of evil, rather than just the omission of good.

5.1) Schellingian objections

There are at least three issues with Fichte's account against which it needs to be defended. The first is a Schellingian response that Fichte is making the origin of evil outside the will, so it ends up being dogmatism. The second is a problem about the normative and motivational aspects of how Fichte is thinking about the will and the moral law. The third, most significant problem, is how to account for the commission of evil actions, as opposed to merely the omission of good ones. I shall take these in turn.

The first problem was raised by Schelling, in the *Philosophical Investigations into the Essence of Human Freedom*. For Schelling, the reality of evil is necessary for the reality of freedom. But this seems to mean more to him than merely the common thought that either a) one cannot have knowledge of good without knowledge of evil, or that somehow good could not exist without evil or b) that part of freedom is the ability to do wrong. Schelling sets himself in opposition to any view according to which evil is a mere privation of the good. He says, for example, that '…only man, the most complete of all visible creatures, is capable of evil, shows already that the ground of evil could not in any way lie in lack or deprivation'.[408] In other words, if evil lay in privation, then it would show up in places where in does not actually show up – in the non-rational creatures, which lack reason, and are therefore 'incomplete'. This in itself would not trouble Fichte, as he is also keen to emphasise that mere privations cannot explain evil. It is not simply that I lack awareness of the moral law that is the source of evil, but that I lack awareness of it because I have endorsed a natural drive which results in me obscuring my consciousness of it.

Schelling also signals distance from Fichte with the following passage:

'But, on the supposition that sensuality or a passive attitude to external impressions may bring forth evil actions with a sort of necessity, then man himself would surely only be passive in these actions; that is, evil viewed in relation to his own actions, thus subjectively, would have

[408] Schelling (2006, hereafter *HF*) 36

no meaning; and since that which follows from a determination of nature also cannot be objectively evil, evil would have no meaning at all.'[409]

The 'supposition' in the first sentence reminds us of the 'intelligible fatalist' view – that if we had a clear consciousness of the moral law then necessarily we would do our duty, and if we had no clear consciousness then necessarily we would not do our duty. The passivity here which is present for the intelligible fatalist is not present for Fichte because he has argued that whether we have such a clear consciousness of duty is at least partly up to us. Schelling's charge is that the intelligible fatalist makes evil impossible, because for the agent, from the agent's point of view, they aren't doing anything, but only nature is acting through them, we might say, but nature itself considered by itself cannot be said to be evil (perhaps because it admits of no properly moral evaluation at all).[410]

The explicit break from Fichte however comes later – when Schelling settles on the meaning of freedom. Something is free, Schelling says, if it acts 'only in accord with the laws of its own being and is determined by nothing else either in or outside itself'.[411] If we confine our attention to the sphere of rational beings, then Fichte might not have an issue with this – because the only 'laws of its own being' are strictly moral laws, the laws of freedom. But Schelling seems to want this definition to have a wider scope than merely rational beings, and if that is the case then from a Fichtean point of view, Schelling's proposed definition has too many false positives – it endows mere things with freedom. Schelling says that the 'inner necessity' of acting only in accord with the laws of one's own being 'is itself freedom', and that 'the essence of man is fundamentally *his own act*; necessity and freedom are in one another as one being'.[412] Here we see a Spinozist influence on Schelling which is not present (or at least, not *as* present) in Fichte. For Spinoza, something is free just if that thing acts in accordance with laws of its own being, just as Schelling describes. For Fichte, as well as the false positives he would claim, there is another issue with this definition. It tells us that things freely act in accordance with their natures, but I (as a rational being) have no nature other than that I give myself. This then yields that this definition of freedom is circular – that to be free is to act in accordance with that which I freely give to myself. Whilst this is true, it is trivially so, and so cannot be a real definition of freedom. Whilst Fichte agrees that autonomous self-legislation is an important part of the

[409] Schelling *HF* 39
[410] The other route that could be taken, instead of saying that 'the natural' considered by itself has no properly moral evaluation, is the Augustinian route of claiming that 'to be' is already 'to be good', and so any evil must be insubstantial – only ever at the level of predicate or property.
[411] Schelling *HF* 50. This is present in Schelling's earlier (1800) essay on transcendental idealism. But here he notes that for God, who is the freest in this respect, is also the most necessitated.
[412] Schelling *HF* 50

moral life, it cannot fully capture the absolute transition from not-willing-that-P to willing-that-P which is essential to freedom.

Schelling's criticism of Fichte, then, is that Fichte who grasped 'speculatively the concept of such an act [the act that is one's own being], fell prey once again to the philanthropism prevalent in his moral theory and wanted to find this evil that precedes all empirical action in the lethargy of human nature'.[413] In other words, Fichte could have achieved a properly philosophical or systematic understanding of evil, but instead opted for a theory more readily understandable or applicable to life, which ends up misunderstanding the nature of evil.[414] But I think this gets Fichte wrong – it confuses two distinct questions which he is dealing with. The first question could be called the 'source question', the second, the 'explanation question'. The source question is roughly: What is the source of evil, or where does evil originally come from? The explanation question is roughly: How does evil action happen, or, given the source of evil, how is this applied to explain evil action? Fichte's answer to the source question is the inertia of nature, which afflicts rational beings insofar as they are also natural beings. But it would be a mistake to think that therefore inertia alone explains why individual agents act badly. That can only be explained because the agent in question *endorses* or *accepts* that condition, that is, that evil actions happen because they are freely willed by the agents involved.[415] In other words, Fichte locates the source of evil in nature, but the explanation for why evil actions occur is because of human actions with respect to what nature gives us.

This would not satisfy Schelling, however. Schelling writes '…if freedom is to be saved by nothing other than the complete contingency of actions, then it is not to be saved at all'.[416] The thought appears to be that if our only ground for thinking of rational agents as in general free were the fact that free choice (or *Willkuer)* sometimes issues in this, and sometimes in that, then there would be no way of distinguishing this situation from one in which actions are groundless or random. This is why, following Kant, Schelling posits an intelligible character or being 'outside all causal connectedness as it is outside or above all time'.[417] The same worries,

---

[413] Schelling *HF* 53

[414] Another way to put it would be that what Fichte has done is let his own particular 'desires-for philosophy' (to use Double's 1996 phrasing) get in the way of the truth of things. So we could see Schelling reading this as seeing Fichte think that philosophy, because it is to be always in the service of life, thinking that it would be better if we had a certain view of evil rather than another. This would be so if Fichte signed up to what Double (25) calls 'Philosophy as Praxis' – the view that the primary, goal of philosophy is to make us morally better people. Whilst there definitely are passages that could support such a reading of Fichte, I think it would be one-sided at best. It also must be said that Schelling also has some passages which could support a 'Philosophy as Praxis' view – for example 'A system that contradicts the most holy feelings, character and moral consciousness, can never be called, at least in this respect, a system of reason, but rather only one of non-reason' *HF* 74

[415] This is to the extent that free actions can be explained, and it is Fichte's avowed view that free actions cannot, *qua* free, be explained.

[416] Schelling *HF* 49

[417] Schelling *HF* 49

then, arise with Schelling's views, as they did for Kant – what would it mean to say that my intelligible character or being has chosen something timelessly, and how is this choice related to the here and now? Schelling's innovation is a complex story about how the duality in human nature eventually reaches down to the nature of being itself. This is why, given God's nature, being 'essentially love and goodness' what is morally necessary in him 'also follows with a truly metaphysical necessity' and 'Activated selfhood is necessary for the rigor of life; without it there would be sheer death, a falling asleep of the good; for, where there is no struggle, there is no life'.[418] Evil must exist, then, insofar as good exists, and this 'must' carries with it some metaphysical or conceptual necessity, and this necessity goes right to the root of being. As Gardner says 'man and the world result not from a separate act of creation but come to be in the course of God's self-genesis – their existence is implicated in God's become, and God does not come to completion without man and the world' which implies that the same 'duality of principles' which makes up God's being also makes up man's being.[419] In the terms above, we can therefore say that Schelling's real worry about Fichte is that Fichte's answer to the source question does not take evil seriously enough, and does not go far enough. It cannot be just that nature has principles which make it true that any natural being suffers from a kind of force, which becomes evil if related to choice in a certain way. It has to be that being itself has a split, or that evil and good are in some way 'dialectically' related – 'good and evil are the same thing only seen from different sides' or that the ground of being is the 'absolute *indifference*' of good and evil.[420]

At as far as Fichte goes, however, this would not be very impactful. Schelling's story relies on intelligible character (or something very much like it) which Fichte cannot make sense of, and also is more straightforwardly a standard metaphysical account of concepts of being, self, and good, which, to Fichte's ears, sounds very much like transcendental realism or dogmatism. In conclusion, we can see that Fichte does not need to be worried about Schelling's criticism, because Fichte's own theory is robust enough to account for how it is that rational agents can choose the bad, and because Schelling's own account cannot give Fichte an internal criticism to worry about, but tells a competing narrative, key parts of which Fichte thinks he has already shown to be either unnecessary or incoherent.

5.2) Cognitive and Motivational factors

The second problem is this. Fichte agrees with the intelligible fatalist that were we to have clear consciousness of the moral law (fixing it before our minds) then we would necessarily do our duty (or at least try to). The difference is that for Fichte, whether the moral law is fixed

[418] Schelling *HF* 61.
[419] Gardner (2017, 148, 151f)
[420] Schelling *HF* 68

before our minds is itself an act of will. But thinking that fixing the moral law in attention would in some way necessarily lead to performance of good actions seems to suggest that Fichte needs to think that the normative reason "that it is my duty" is the same as the motivational reason "that it is my duty". This might strike us as odd, considering that we often think we should distinguish between reasons for an agent to act and reasons that motivate them to act (and both of these are distinguished from explanatory reasons which state why they acted). I think Fichte's answer here has to be something along the following lines: thinking that there is such a thing as grasping the moral law in a mere cognitive or merely intellectual act is mistaken. There is no such thing as grasping the moral law (or what the moral law demands of one) without also being motivated to so act. Here is a sketch of a possible position.

It is tempting to see Fichte as a moral realist in some sense – that is, someone who believes that there are independent or objective facts about morality. This captures a lot of what Fichte does say, but he also claims that the moral law does not extend beyond the intellect or does not subsist outside of the I. What I take these two seemingly-in-tension views to amount to is this: that there are moral duties which attach to individuals in virtue of those being rational agents, but there are no free-floating facts about morality. So all propositional contents of which it is true to say that it is an imperatival moral statement (that I ought to X, that you ought to Y, that everyone ought to Z) are indexed or attached to subjectivity in some way. This could be the I as such – and therefore everyone is bound by this duty (these are Fichte's universal duties), or it could be some proper subset of rational creatures (say, scholars, which have particular duties). It could even be just one person, who would have particular duties with respect to their spouse, for example.

But if by 'moral realism' we mean that there need to be some mind-independent or thought-independent abstract entities which would exist whether rational beings did or not, then Fichte is not a moral realist. This is shown more clearly when we consider the various ways we can consider the realism debate. We can think of the debate semantically, epistemologically, or metaphysically. A semantic way of framing the debate is the question "Are there true statements of the form: I/You/We ought to (ought not to) X (where X is an action of some generic kind)?" To this question, Fichte would clearly answer yes, even though, of course, Fichte does not consider the question in such terms. The epistemological way of thinking about the question can be seen as "Can we have knowledge (as opposed to mere subjective opinion) of moral truth/facts?" To this, again Fichte answers in the affirmative.[421] The Metaphysical way of framing the question would be: "Are there mind-independent moral entities/properties?" To

[421] One might think that, given factivity of knowledge, an affirmative answer to the Knowledge question implies an affirmative answer to the Semantic question. But it is not clear whether this can be so, given one might have an expressivist view according to which I can know various moral claims to be true or false, but that is because the moral claims are actually carefully construed attitudinal expressions, rather than cognitive truth-apt claims.

this, it is unclear what Fichte would say, because of the ambiguity in the concept of mind-dependence. Roughly, we can think of mind-dependence as a weak claim (depends on mentality or thought or features of rational agency in general) or a strong claim (depends on the thoughts of determinate individuals). Morality is not, I think, going to be strongly mind-dependent, but only weakly so, which puts it on the same footing as the empirical objects of ordinary experience.[422]

Insofar as morality or moral truths are then given transcendental justifications, the question of the realist (or indeed the sceptic) about whether they *really* exist ceases to have meaning. It becomes asking whether there are things in themselves which somehow lie behind ordinary moral experience, which is tantamount to asking whether it is possible to represent un-representable things. If, in order to be a realist, one needs to have a view according to which objects would exist or truths would be true even without reason, then Fichte is no realist. But there is also no reason for Fichte to concede the terminological ground here – the dogmatist realist is gerrymandering the word so it applies only to their view, and it would be begging the question for them to say that one couldn't be a realist unless one accepted transcendental realism about moral properties or truths. So Fichte can be called a realist about morality, just as he is a realist about empirical objects.

Because of this form of realism (the moral duties are as real as empirical objects – that is, they are necessarily posited by an I who judges) there arises the question of motivation. There are two possible routes that Fichte can now take. One route is through the Primacy of Practical Reason. The other can be called the "Indexical route". This is where I think the primacy of practical reason shows itself. Recall that PPR means that there is some transcendental or necessary dependence of theoretical reasoning (or our ability to cognise things) upon practical reasoning (our ability to do things, or more primitively, our ability to strive or try). We then see that grasping practical realities is more fundamental than grasping empirical or theoretical realities, and so the question of how we can know about a moral demand and yet not come to be motivated to fulfil that moral demand rests on a mistake. The indexical route is different. This is to take seriously what I ascribed to Fichte above – that there are no duties that are not attached or indexed to agents. In other words, an agent would always in a position to ask whether or not they themselves are bound by the moral law in this regard (which would be to ask whether it is a part of their vocation). The identification of any duty would then be: "I am required or demanded to X or Y" and the "I am" part functions as the motivational key. In other

[422] A further ambiguity – some duties are clearly strongly mind-dependent insofar as if I did not choose to get married or enter into a loving relationship (which in strictly moral terms for Fichte would mean the same thing), then I would not have certain duties – so the fact I have these duties is caused by my choices (call this Causal Mind-Dependence). But this mind-dependence is benign, and different from the conceptual mind dependence according to which there would be no sense to be made of there being objective facts about whether I have duties or not.

words, it is all well and good to be told that this or that is a moral fact, but to be motivated, one has to understand that it involves oneself.

Fichte seems to suggest by his remarks on the deep connection between knowing what the moral law requires of one and then performing the actions that there is a deep connection between knowledge and motivation. We can then say that for Fichte, something like the following principle holds: 'Whenever S knows that they are bound by a duty, then S is motivated to some degree to perform the required action'.[423] This seems plausible, insofar as it is difficult to imagine a case of an agent who recognises that something is their duty but does not even recognise any *prima facie* motivation (they may of course come to think that on balance they are not motivated). Evil happens, then, when one's knowledge of one's duty does not provide sufficient motivation to one to counterbalance the endorsement of the passivity engendered by the *vis inertia*. That is the generic description of reasons why evil happens, and we can outline some specific ways in which this might be realised.

A first way is the failure of an agent to sufficiently realise that the duty has a 'me-ness' – that it is not just a duty, but is *my* duty. This can be seen in cases where there is a diffusion of responsibility – agents claim that they did not/do not need to act thus and so (even though it is required by duty) because someone else will, or because perhaps no one else is. Of course, this is only applicable to some duties – it would be very odd indeed if someone claimed that they didn't have the motivation to exercise because someone else would do it. A second way this generic reason would be realised is a failure to recognise the unconditionality of the duty. This can happen in cases of diffusion as well, but also happens when an agent lists a morally good action in a 'to-do' list, as it were – slotting it in among other things to do, rather than seeing that the demanding character of the moral action should take precedence (within the bounds of conscience, of course). This explains sufficiently why it is that there is a failure of motivation and cognition, and how these two things are bound together, I think.[424]

---

[423] This is also near to Kosch's reading of Fichte (2018, 5).

[424] Fichte might have a position according to which our cognitive and motivational capacities are themselves split by evil – the allegorical eating of the fruit of the tree means that the life according to instinctual reason is no longer possible, but this opens up the space for the life according to self-conscious reason, as Fichte's philosophy of history has it. *CPA* 7ff. Thus, when God says to Adam that if he eats from the tree 'that day he will surely die' (Gen 2:15), it signals that spiritual death when we are estranged from our own natures (a theologian who provides a philosophically rich account of this 'death' is provided by Bonhoeffer 1959). But when the serpent says that Adam and Eve 'shall be as Gods' (Gen 3:4) they also speak truly, because the law of morality being the sole norm of the entire being is at least part of what God is. Note that Fichte occasionally speaks of the project of morality to make us like God or to fuse us with God, e.g. 'The complete annihilation of the individual and the fusion of the latter into the absolutely pure form of reason or into God is indeed the ultimate goal of finite reason; but this is not possible in any time' *SE* IV:151, 143, and Fichte's early lecture 'Concerning Human Dignity' (in Fichte 1988) 'Earth, heaven, time, space, the limitations of sensuousness: they all vanish from me at this thought. And should not individuality also vanish? I will not lead you back to individuality…*All individuals are included in the one great unity of pure spirit*' (I:416, 86) and this is said to be 'an *unreachable ideal*, an ultimate goal' (I:416n, 86n)

5.3) How is great evil possible?

The third problem above is the most important objection to Fichte here. This is the problem of how to explain the commission of evil as well as the omission of good. It is relatively easy, on Fichte's account, to explain how it is that someone can fail to live up to the moral law and why they do not perform actions that duty demands. It is harder to explain why someone should perform evil actions. Clearly, Fichte cannot think that the reason why a murder or theft occurred is because the murderer/thief was lazy. The absurdity is taken to the limit when we consider evil on a massive scale such as war crimes like the holocaust. In these events, there was clearly a great amount of thought and planning that went into making such efficient killing machines. Straightforward laziness is not an option.[425]

Clearly Fichte's account has to be more nuanced. I think it is. We can see that Fichte is convinced that the explanation for a lack of activity (omissions) is explained by laziness. But this also goes some way toward explaining active evil, or commission of evil. This is because the agent who does so exhibits a kind of laziness of thought which is slightly different to the above laziness.[426] The first laziness, or omission-laziness, is laziness with respect to actions or particular duties. But this laziness is explained with reference to a prior inertia regarding what it means to be an agent. The person remains at a certain level of self-conception – the level of self-interestedness. This level explains why they do not perform certain good actions, but also explains why the agent performs bad ones. An agent who is self-interested would pursue things that are easier than otherwise, and this may involve contravention of negative duties, e.g. stealing and so on. So Fichte's theory of inertia can explain the commission of evil as well as the omission of duty. But it is not clear that the objector would be satisfied – given that the objector was worried about the commission of evil on a mass scale. Just as it would be unsatisfactory to say that a failure of motivation could explain these, it seems equally unsatisfactory to claim that these are explained through self-interest, although self-interest surely plays a role.

So how is it that people can commit such evils?

---

[425] On Kant's view, whilst the good will cannot be augmented or made better by external circumstances or conditions, the evil will can be made worse. So the people behind the holocaust with planning expertise were made all the more villainous by the fact that they had such skills. Fichte can agree with this – he would say that the fact that people so technically skilled committed such atrocities makes it worse, because they are using their skills in service of evil.

[426] My account here owes much to the recent work of Allen Wood (2016) and Owen Ware (2015)

There is also a third kind of evil, which is arguably a more insidious kind. This is the kind when our constitutive end of self-sufficiency is interpreted through the lens of inertia. It is a core part of Fichte's moral philosophy that to be a rational agent is to have a certain end. This end is agency for its own sake, freedom for its own sake, self-sufficiency, and so on. We might, as some philosophers do, worry whether this constitutive end is 'material' enough – that is, whether it is purely formal or not.[427] This is relevant to my concerns in the following way. The fact that the end is (necessarily, Fichte thinks) under-described and so ends up being somewhat formal is important to the explanation of evil. This is because, if one is not at the level of proper reflection on what it means to be an agent and to have moral duties, then one will be bound to misinterpret the end, and what the end entails. Just as we bargain down the moral law until we are satisfied that what it asks of us is not as demanding as it is objectively, we distort our end until it becomes something more manageable. This is, for example, how Fichte would have viewed the utilitarian version of this end. For Fichte, the utilitarian is right to say that human (or rational, or conscious) agents have an end, just in virtue of being conscious or rational, but wrong to specify this end as happiness.[428] Happiness may well be a part of the final end as it really is, because happiness may be, as it is for Kant, finally proportioned to virtue. But it is just a mistake to think that happiness is the whole or complete end. Fichte will view the utilitarian (in moral terms) as thinking of themselves in such a way that the only end that they could reasonably have is happiness. In this way, their view of the final end and their consequentialism run together. Because they are at such a level of self-conception according to which their final end could not be anything other than happiness, they also are unable to conceive that there could be such a thing as actions done for their own sake. On Fichte's view, by contrast, because the final end is conceived in a formal way, there is no problem in conceiving of actions as having to-be-done-ness for their own sake rather than having it conferred on them by some further substantive end. In sum, the utilitarian, on Fichte's view, cannot conceive of anything apart from in terms of pain or pleasure, and this is because of their standpoint on themselves. They remain at a standpoint where they are unable to conceive of themselves in any other way – and unable to see that it can be true that happiness is caused by what is good, rather than happiness conferring goodness on action.

Above, when I said that our constitutive end is viewed through the lens of inertia, this is what I was suggesting. Fichte's idea is that remaining at the level of self-interest actually in some

[427] For the recent debate see Kosch (2014, 2015, 2018), who thinks the principle is material and substantive, and thereby reads Fichte as a non-welfarist consequentialist, Wood (2016) who argues that the principle is still formal, and reads Fichte as a deontologist, Ware (2018), who adopts an intermediary perfectionist view.

[428] He would also take issue with the utilitarian's belief (evident in both Bentham and recent utilitarians, most famously Singer) that the criteria of good and bad actions can be meaningfully extended to non-rational creatures.

ways determines how we conceive of ourselves as agents in such a deep way that even our deepest or most distant ends or vocation are conceived through that standpoint. If this is the case, we can see how self-sufficiency gets re-interpreted in terms of self-interest. Self-sufficiency in the lens of self-interest yields domination. This is Fichte's answer to how great evil is possible – it has its roots in the desire for domination, which is itself a cognitive failing (insofar as more honest and open self-reflection would have changed it) and a motivational failing (insofar as an agent at this level conceives of everything from the view at this level – they cannot see why they should abandon the standpoint in favour of a moral one, insofar as abandoning the standpoint does not reward them in any material way). So the commission of great evil has its roots in self-interest, which has its roots in inertia, or the endorsement of inertia. Great evil is possible because the standpoint from which agents conceive of themselves self-interestedly is also the standpoint from which our final end is seen in terms of everything serving us, and us serving nothing, in other words, domination of nature and others.[429] This is a perversion or corruption of Fichte's real view, which is that the moral, religious, and other principles can be posited as a unity or whole, which is infinitely far away from our current situation, but still can be recognised as our final end. In sum, because the person at the level of self-interest only sees things in terms of self-interest, they interpret the final end in terms of self-interest, which comes out as domination; but the real final end is a more harmonious whole than this.[430]

Fichte appears to hold this view insofar as he thinks the maxim of the 'blind drive' to self-sufficiency is 'unrestricted and lawless dominion over everything outside us'.[431] Fichte also holds the view that this domination is often opaque to those who pursue it. He says 'It is not that the human being intends to bring everything outside of him under the absolute sway of his will – he does not intend anything at all, but is only driven blindly -, but he acts as though he had this intention, and he does so for absolutely no other reason than because he wills to

[429] Some take this to be Fichte's real view. For example, Dryden (2013) seems to think it is Fichte's view that the way to resolve issues (such as the body's inherent vulnerabilities) is to have increased control over them. This does capture a part of Fichte's view, but it is incomplete. Fichte thinks of the final end as a kind of wholeness or unity, rather than the domination of the non-rational parts by the rational.

[430] Something like this means that Fichte could be sympathetic to some of the ideas of the Frankfurt School, particularly Adorno and Horkheimer in *Dialectic of Enlightenment* (2000). Though of course it should be emphasised that Adorno and Horkheimer have a complex narrative about how it is that science and technology have led society down a path towards barbarism, there is however perhaps commonality insofar as their critique essentially relies on ideas about the dangers of instrumental reasoning if it is unrestricted. The harmony that I think Fichte has in mind, which always needs to be tempered by the fact of finitude (Breazeale 2013c), puts Fichte in interesting relations to Schiller and others (Gardner 2018). Breazeale (2013c, 150) is worth quoting here: 'The practical lesson of philosophy is that there is nothing *tragic* or *absurd* about having "two souls" and no reason to lament this discovery. In other words, it addresses itself not so much to Faust's discovery that he possesses *zwei Seelen* as to the *ach!* with which he greeted the discovery'.

[431] *SE* IV:186, 177

do it'.[432] The drive to self-sufficiency for its own sake, then, is 'blind' if it is not interpreted in the light of our moral nature, and guided by concepts. That is to say, that to the agent who acts with the blind drive, they cannot see that they do so, but they can be reasonably interpreted as acting in accordance with such a maxim. The dominion is 'lawless' because there is no check or balance for it – it is wholly up to the self-interested agent.[433]

The dominating type or the self-interested individual who sees no further goal than happiness might also be led to commit suicide. Here Fichte argues that the 'decision to die' because it is wholly contrary to every natural impulse,[434] is the 'purest presentation of the supremacy of the concept of nature'.[435] If the person who wishes to commit suicide does so with a 'cool and thoughtful self-awareness' then we can be sure that their action has issued from the 'blind drive to self-sufficiency'.[436] The idea seems to be that the person who commits suicide in this way (as opposed to in the grip of passions or in a deep depression) shows their determination to master nature because they contravene natural impulse with an action issuing from a kind of cost-benefit analysis of remaining alive.[437] Of course, Fichte argues that suicide is immoral, and that the person who decides to die after this cost-benefit analysis exhibits far less 'strength of soul' than the person who endures life. They represent the 'triumph of the law of thought' whereas the former represent only 'the triumph of thought'.[438] In other words, if we consider the matter aright, we should not be misled into thinking that, if life is going to be on balance bad, or contain more suffering than not, that this is a cause for drastic action. The mere drive for mastery over nature does not exhaust what it means to be a moral or rational agent.

An issue which seems to be important to discuss as well is the thought that will occur to Aristotelians. This is roughly the following. On Fichte's view, the moral law requires that we be always active and progressing and so on. In other words, Fichte is committed, both by his view of what morality requires and what evil consists in, that it always requires effort to be moral, or at least requires more effort to be moral than not be. But what of the Aristotelian virtuous agent, who acts for the sake of the good and gains pleasure from so acting? This person

---

[432] *SE* IV:186, 177

[433] It may be that what Fichte means is something similar to Anscombe's (1958) famous objection to Kant's idea of self-legislation – that I cannot be said to bind myself because I always maintain the absolute right to opt-out of any obligation I give to myself. This may be right for Fichte, but only in a restricted sense, only in the sense that the arbitrary will that deviates from the moral law maintains a normative opt-out clause (that they always have the right to change what they will if their desires change) whereas the moral person recognises no such normative right, because their universal or fundamental duties never change.

[434] Here we also get a hint that it is not the natural impulse itself that is bad, but the usage of this impulse for substantively evil ends.

[435] *SE* IV:267, 256

[436] *SE* IV:267, 256

[437] Perhaps another point of connection with Adorno and Horkheimer – the person Fichte is thinking of here is taking merely instrumental rationality to be the only canon of reason, ignoring the substantive end-setting forms of rationality that Fichte insists on.

[438] *SE* IV:268, 156

seems like it might be actually *more effort* for them to contravene the moral law in some sense. They are so accustomed to acting well that they approximate as far as is possible a holy will – a will on which there is no gap between ought and is. So the challenge is that Fichte's theory of evil does not account for the agent who, in remaining the same, may succumb to inertia of the will in some sense, but does so because they are already a moral agent.

Fichte's response to this would be the following. Whilst it is true that the moral agent would remain still in their moral nature insofar as they comprehend their vocation and what it requires of them at every moment, they are still being wholly active for each individual action. They still need to bring themselves out of their current state, which requires a kind of effort, even if it is not a total moral conversion. Fichte just flatly denies one of Aristotle's key thoughts, namely, that virtue can become a habit. For example, he says: 'Practice and vigilance, standing guard over oneself must be continual: no one is certain of his morality even for a moment without continual strenuous effort'.[439]

6) Free Will and Time

In this section, I address a few questions regarding the relationship between time and free will. One of these is the problem of divine foreknowledge. Whilst this is not dealt with explicitly by Fichte, it is worth looking into insofar as on my reading, Fichte's final account of freedom and nature requires God. The problem of divine foreknowledge is usually stated in the following form: 1) God is omniscient, that is, God knows all that there is to know. God's knowledge includes perfect knowledge of the past, present, and future. 2) Humans are free, that is, a human future is undetermined by the past and the present. These two propositions, both of which seem acceptable *prima facie*, seem to lead to a contradiction – that my future is at once determined (because God knows my future perfectly) and undetermined (because I am free). It is clear that God as the moral world order is not going to be like God as normally understood, and so when we speak of God's knowing, Fichte is going to understand this to be a different kind of thing entirely to a person's knowing act, if indeed any statement that 'God knows…' is even meaningful.

6.1) Theological Issues

A way of dealing with this which we can immediately rule out on Fichtean grounds is the 'simple foreknowledge' view, associated with Dominicans following Aquinas. This is the view

[439] *SE* IV:193

that actually we are not free in the sense necessary to generate the problem. Two other options popular in contemporary philosophy are Molinism and Open Theism. I shall discuss these in turn.

A Molinist is someone who, following Luis de Molina SJ[440], claims that God's foreknowledge and human freedom can be reconciled by virtue of the device of distinguishing in God various kinds of knowledge. Roughly, there are three types. Firstly, he has natural knowledge, the knowledge of things that he possesses in virtue of being God – such as knowledge of necessary truths. Secondly, he has free knowledge – the knowledge he possesses in virtue of being God who has (freely) created the actual world, such as knowledge of the past. Thirdly, there is middle knowledge. This is knowledge that God has pre-volitionally (before the act of willing creation) of all true counterfactuals of creaturely freedom. That is, for any proposition "Were N to be in situation S, they would perform action A", if that proposition is true, then it is known to God. This knowledge is pre-volitional but is after the moment of God knowing himself. It is the moment at which God studies the *possibilia* and concepts to decide which world to make actual. Of course, for Molina, this all happens at once, but there is a logical order to these moments. Or, as Molina frequently says the moments are before or after one another 'to our way of conceiving, but with a basis in reality.' For example, Molina says ''It is not simply because things exist outside their causes in eternity that God knows future contingents with certainty; rather, before (in our way of conceiving it, but with a basis in reality) He creates anything at all, He comprehends in Himself – because of the depth of His knowledge – all the things which, as a result of all the secondary causes possible by virtue of His omnipotence, would contingently or simply freely come to be…'[441] The key for Molina is that whilst things come to exist because God wills them, God wills them based on a prior apprehension of their properties, which include the faculties of will and choice. God thereby, on the Molinist view, actualises the world, which is done on the basis of his knowledge of these conditionals. The key to middle knowledge is that the counterfactuals are not *made true* by God, but are true because of the characters or essences of individuals. A Molinist then can confidently say that in any given situation, I could have acted otherwise – that is, I retain the power to act otherwise, but God knows which way I would in fact act, thus appearing to preserve creaturely freedom and God's foreknowledge and providence.[442]

---

[440] Molina's key ideas are in Part IV of his *Concordia*. This is in English as Molina (1988).

[441] (1988, 115-6). See also 179: God knows 'knows 'in the very determination of His will, as in the cause of their coming to be, that those things will come to be. But it is not the case, conversely, that those things will come to be because He knows that they will come to be, since the fact that they will come to be by virtue of the free determination of the divine will is *prior* (in our way of conceiving it, with a basis in reality) to God's knowing this fact on the basis of that very same determination'

[442] Stated like this, Molinism has affinities with Leibniz' views in *Theodicy*, insofar as for Leibniz as for Molina, God actualises the world based on a (logically) prior apprehension of the concepts of *possibilia*, which exist in the mind of God, but about which there can be stated true statements which do not rely on God's action.

However, this is not acceptable to Fichte either. The chief reason is because Molinism relies on there being something that I am independently of and before my existence. But this cannot be – I am only what I make myself to be. In other words, the Molinist needs there to be determinate facts about what I am like that obtain independent of my existence. This is a form of dogmatism – there just are no facts about what a person is like independently of and before the existence of that person, on the idealist view. Fichte needs it to be the case that God can be provident in some other way that does not need God to know 'truths' about individuals before they exist. One might say then that Fichte could have a model in which God denies himself omniscience because he wishes to respect creaturely freedom – so he could know the propositions but chooses not to for our sake. But this of course admits that which Fichte denies – that there are such propositions. To put it another way, what seems to trouble the Fichtean most on this score is not going to be that God knows the propositions, but that there are propositions, or facts, about what I am like independently of my existence and action. At this stage, Fichte is in the stronger dialectical position because he can deny the view that if something is true, then it has always been true, and supplant this with his account of free action and the self that I have shown in section 3. This mirrors a classical objection to Molinism – the so-called grounding objection. This is that in order for there to be a counterfactual of freedom, there must be some categorical statement which grounds it or makes it true.

But, Fichte will say, given that we are essentially self-conscious (that is, the way we are for-ourselves can change the way we are in-ourselves)[443] there is (for at least a proper subset of these counterfactuals of freedom) no fact of the matter which could ground them. This is because in order for there to genuinely be counterfactuals of freedom – "Were I to be in situation S, then I would perform action A" – there would have to be something that I am like which grounds this counterfactual. I would have to have a nature. But as a rational being my nature is indeterminate – I make it so, by conceiving it a certain way. This is the prerogative of essentially self-conscious beings –that we can consider ourselves to be a certain way, or have a certain goal, and then practice until we really are that way.[444]

So Fichte cannot be a Molinist. I will below pick up the thread of propositions and truth and what that could tell us about free will. But I will now turn to Open Theism, the third of the contemporary views on divine foreknowledge.

---

[443] This is Brandom's (2007, 127) phrasing. He goes on to say 'To say of an essentially self-conscious being that what it is for itself is an *essential* element of what it is in itself entails that an alteration in self-*conception* carries with it an alteration in the self f which it is a conception. Essentially self-conscious creatures accordingly enjoy the possibility of a distinctive kind of self-*transformation: making* themselves be different by *taking* themselves to be different' (128)

[444] Rather irreverently, we might say that this is Fichte's explanation of the common psychological advice that one can "fake it 'til you make it" – that merely acting as if one was brave makes one think that one is brave and so ends up actually being brave, for example. This is also a result of the proleptic effect of free will, which I discussed in chapter 4.

Open Theism is, roughly, the denial that the future really exists, and therefore the denial that God knows the future (because there is no future to know). This has some appeal, if one is willing to make what looks like a sacrifice of omniscience, though of course for the open theist, this is no sacrifice, because the future is not there to be known anyway. The problem is that God, being in full knowledge of the past and the present, and being a perfect reasoner, will be able to deduce with as much accuracy as possible the future anyway. One might want to draw a distinction between God's perfect knowledge and God's perfect prediction, but it may turn out that God's perfect predictions end up counting as knowledge anyway.

A second problem with this view is that it is unclear how God is supposed to come to know all the events that unfold throughout the steady passage of present into future. The most obvious way to account for this is God's general concurrence. But invoking this leads to a problem. One of the well-known responses to Anscombian accounts of action as including the knowledge that one is acting thus and so when one is acting thus and so is Davidson's challenge. This challenge supposes that it is incoherent to say that I can know that I am (say) securing my children's future before the event has happened, even though I may act under the description of writing a will so as to provide for my children's future. The challenge is then about how the self-knowledge of action cannot extend into the future in the way required to say that we always have the requisite self-knowledge. If God's general concurrence is granted, then we seem to have to say that God also acts so as to bring these things about, but then Davidson's challenge works against God's action too. This seems an unappealing conclusion. But if we deny that Davidson's challenge works against God's action, then we are in a position where we have to give up Open Theism (because God works throughout the entire world at all times in order to bring states of affairs about and to have the self-knowledge of action required he would have to know the future, *pace* Open Theism).

So it may be that we can avoid Open Theism, for these philosophical issues as well as exegetical worries we may have. Fichte consistently talks about God as provident, and the classical thought of foreknowledge is that perfect foreknowledge is required in order to maintain providence. If God's knowledge was, like ours, in some sense a reaction to the things unfolding independently of that knowledge, then God's providence would not be guaranteed. But it is clear, I think, that for Fichte the providence of God needs to be in some sense believed in. For example, when speaking from the perspective of the religious person,[445] Fichte says

[445] Contrary to tradition, which identifies the contemplative as the higher or superior kind of life, Fichte identifies the active life as the superior. In theological contexts, this is usually used to explain why (at Luke 10:38-42) Christ tells Mary that she does the right thing but chides Martha, even though Martha is the active worker, and why Christ rebukes Judas who claims that Mary could have used the spices with which she washed Christ's feet to give money to the poor. For me, this raises questions about the suitability of thinking about the *Way Toward the Blessed Life* in a mystical context. For the mystical reading, see Perovich (1994), and for criticism of that reading, Zanelotti (2010).

that: 'Whatever may come to pass around him, nothing appears to him strange or unaccountable; he knows assuredly, whether he understand it or not, that it is in God's World, and that there nothing can be that does not tend to Good'.[446] This tendency to the good is the way the believer makes sense of the world as being God's world – and tendency to the good sounds like a way of suggesting providence; literally, that God has provided. This also sounds like an acceptance of God's general concurrence – everything happens either with God's will or with God's permission. Note that Fichte is not saying necessarily that this is how things are, but just that this is how things are perceived by the religious person. It is characteristic, perhaps even constitutive, of the religious point of view to see the world as providentially ordered. This can be seen as an exercise of the reflective power of judgement, rather than the determinative power.

The problem is actually deeper than foreknowledge and is an issue for accounts of Christianity. This deeper problem is the problem of inevitability. The Christian faith appears to accept both of the following statements: 1) God's victory is assured or inevitable 2) We are called to participate in God's victory. 1) has strong biblical and theological grounds. The most obvious thing to say is that if God's victory were independent of God's nature and therefore depended on circumstances external to God, it would jeopardise his omnipotence or omniscience or both. (2) also has biblical and theological grounds, insofar as a person's life is meant to be one of repentance, worship, and good works in the light of faith. But what would be the point of (2), if (1) were true independently of any effort that (2) calls on us to make? This is the problem of inevitability – that it seems to be both true that God's victory is assured and that we are needed for the victory. It is this (or an analogous) problem that I think Fichte is struggling with. A possible solution for Fichte is to say that God's victory is one which inherently requires our participation in the divine life. That is, God's victory is assured only in the conditional sense that it is assured if we participate. This participation is also helped by God, because the moral world order is 'living and actually efficacious' and we know that 'in Him we live, move, and have our being'. God's – the moral world order's - victory becomes actualised in the life of believers that come to be in some way united with Him through Christ. How does this help Fichte on divine foreknowledge? I suggest we pay attention to the way in which God's knowledge functions.

Given God's general concurrence, it could be said that God's knowledge is conceptually a function of his will (traditionally, God's knowledge that p and God's willing that p are thought to be in a real unity, even if conceptually distinct). Recall above our objection to Open Theism – that God's general concurrence and a plausible Anscombian thesis about action means that God can be said to know the future because he knows on the basis of his actions. If we take

[446] Fichte (1889b) *WBL* 476

our theory about the problem of inevitability then we can see that God's "knowledge" of the future is based fundamentally on the fact that he acts through us. But this is not a problem any longer, because it is based on the fact that the moral person knows what the moral person would do. There is still the possibility of falling into evil (by way of inertia), and so there is nothing which here threatens free will. Fichte clearly thinks that the religious person is in some sense unified with God, and is the channel for God in the world. It would then not be surprising if Fichte was also committed to the idea that God acts through us, and in some sense this has bearing on the problem of foreknowledge. This thought is illuminated by Farrer, who works through what it would be like to have a union of a fallible human will and an infallible divine will. He says that grace (or God working in a person) is seen 'in the influx of the divine will into the human' and that the 'grace of God' is not 'to specify or trace the line of causality by which they descend to us from their divine source'.[447] This bears on Fichte because it could not be that God working through us could be expressed as a kind of zero-sum game where more of God equals less of us and more of us equals less of God. That picture would mean that the moral person is actually the least free, because they are permanently being driven by external forces. It has to be, instead, that our will and God's will merge, such that the moral world order (in the guise of the moral law) helps us, but that the action is ours. We can make this a little clearer with the language of vocation, which Fichte employs throughout the moral discussions in the *System of Ethics*.

The individual – all individuals – have a calling or vocation. This vocation is not set by them *per se*, though it is of course influenced by the choices that a person has made up until now. It is both independent of me and dependent on me. It is independent of me because I did not fundamentally elect to be the kind of person I am, because I was called into being by persons outside of me when I was not a person, and from then on there have been many external circumstances, both physical and moral, which have influenced me. It is dependent on me because the response to these circumstances is not itself determined by them. The vocation is God's will for me in the world. To be free is to respond to that vocation positively and affirmatively, overcoming the *vis inertiae* inherent in natural beings at every turn and embracing the destiny of humanity – becoming self-sufficient, or one with God. To be sure, as Farrer says, I can reject my vocation, and hide from it, but this is a kind of suicide. This is what Fichte would say too – I cannot hide from my vocation unless I am willing to perversely hide myself from myself. (First Introduction, Second Introduction, DGW, BL). God knows what I will do because he himself has called me to this vocation. Responding to the vocation just is knowing that God has made it so.

[447] Farrer (1972, 198)

The fact that we each have a vocation, or a 'determinate place in the moral world order' does not mean that we are bound to respond to it, in any sense other than the moral one. It may be God's action that calls me to the vocation, but this is a result of my cooperation with the moral in the first place and as a mixed result of my action and the way the world is. It may be God's action that sustains me in the vocation, but it is still *me* whom God sustains. God's providence is shown, for Fichte, in the fact that in properly responding to our vocations, we and God can be said to be in harmony, in a unity, working together.

There is an issue lurking around here. This is the relationship between rationality, freedom, and evil. One might be troubled by the answer to the question of whether Fichte ought to be troubled by the problem of divine foreknowledge, insofar as the answer I have given relies so heavily on Fichte's account of calling or vocation. It might seem as though this somewhat undercuts Fichte's claims that the I has radical freedom and makes itself what it is. Firstly, it requires stressing that the I is free in this regard, but it has, *qua* rational activity, a constitutive end, which Fichte thinks he has uncovered in the first deduction of the System of Ethics – the end of self-sufficiency, or activity for its own sake.

More fully, however, there is in Fichte, as there is in Kant, the thought that in some sense acting morally and freely are united with necessarily so acting. Of course there are many senses in which this is not true – necessity cannot here mean compulsion or alien causes. Rather, it strikes at some peculiarity about the moral agent, and what it is like to be a moral agent. I call this the Anselmian thesis: that the ability to do evil or sin is no part of what it means to be free. Kant of course agrees – freedom is the acceptance of autonomous as opposed to heteronomous principles. Anselm has different motivations – he does not want to accept an equivocal definition of freedom; God is free but lacks the ability to sin, so it would mean that there would be one definition for God and one for finite rational agents if it were part of creaturely freedom that there is the ability to sin.[448] Anselm goes on to argue that someone 'who so possesses what is fitting and advantageous that he cannot lose it is more free than someone else who possesses the same thing in such a way that he can lose it and can be induced to what is unfitting and disadvantageous' and given that sinning is always unfitting and harmful, it follows that a will that is unable to turn away from not sinning is more free than the will which is able to.[449] Anselm himself defines freedom as 'the ability to keep uprightness-of-will for the sake of this uprightness itself'.[450]

So the Anselmian point is this: that whilst of course we have the ability to perform evil actions, the ability to so perform is not actually a part of what it means to be free. We can adjoin this to the Fichte/Farrer point. This was that, whilst I have the ability to abandon my vocation, I am

[448] Anselm (2000, 1, 192)
[449] (2000, I, 193)
[450] (2000, III, 197)

in doing so doing something perverse – I am running away from myself, not living up to that which I am called to be. The synthesis of these points leaves us here: that as a rational and free being I am most rational and most free when I find that which I must (morally) be. The ability, as Fichte would put it, to be overcome with inertia is not freedom, even though I must freely choose to be so. It is at most a formally free choice which leads to the opposite of material freedom. This is but a sketch of a possible solution to the problem of inevitability, which is the main worry for Fichte here. In the next section, I will discuss why Fichte might not really be troubled by divine foreknowledge at all, because of his views on truth.

6.2) Temporalism vs Eternalism

Here I want to pick up the thread of thinking about eternalism and temporalism about truth. This section, as with much of this chapter, is about trying to get to grips with Fichtean answers that could be given to problems surrounding free will and time. Therefore I do not in general ascribe these views to Fichte, only say that these views seem to be to be amenable to Fichte's perspective. First we should look at Fichte's stated views on truth, in order to see if there is any textual support we might get.

Fichte does not say a great deal about the nature of truth *per se*, but there are scattered remarks throughout, especially in the WLnm. For example, when discussing his theory of perception, and the role of feeling, representation, and so on, Fichte says that 'the truly characteristic feature of an object (or of "reality") is that it is something that is posited in consequence of a feeling'.[451] Briefly, feelings are always a form of limitation – a feeling is the mental act of intuiting that something limits my activity. Fichte also says that an 'external object is an interpretation of our own feeling'.[452] Feelings, being the felt limitations of the I, produce in the reflective I an intuition or a representation, and truth is the name for the harmonious relationship between the feeling and this representation – 'Truth is agreement with ourselves, harmony'.[453] Fichte explains that in imaginative cases, like fiction, truth is inapplicable as a predicate or property, because the intuiting or representing faculty and the faculty of feeling are separated in the mental act of imagining. But in representing real things, or in representations accompanied by a feeling of necessity, the faculty of feeling and the representing faculty are united, and this is why truth can be found.

Fichte also claims that 'Representations of the world are determined by all the laws of reason, {which express themselves in the world,}…'[454] In another passage, he says 'some of them

[451] K105/H96
[452] K105/H96
[453] K106/H96
[454] K107/H97

[other philosophers] have charged that one of the major errors committed by transcendental idealism is that it proceeds in the following circle: "It is supposed to provide an explanation and derivation of the laws of reason; yet, in order to do this, I have to proceed in accordance with these very laws." There is no way to avoid this circle of reason; indeed, this is precisely what shows that there is no "truth in itself"'.[455] These two passages concern what Fichte calls the laws of reason. The first says that there is some relation between the laws of reason and the 'expression' of these in the world. The second says that the fact that we have to use reason in order to explain reason tells us something important about truth. I think that what Fichte is getting at is this: there is no sense to be made of the dogmatist claim that there is a truth in-itself outside of the scope of reason. The dogmatist, in taking the position that there is no necessary or *a priori* connection between the objects of the world and the faculty of reason, has to think it a contingent state of affairs if it turns out that there exist no fundamentally un-representable objects. In other words, I think Fichte is claiming that the only sense to be made of the idea that there be truth 'in-itself' or independent of reason is if there were in principle objects that could be un-representable. But there is no sense to be made of an object that is un-representable.[456] Because the dogmatist will have to say that at best it is a happy coincidence that there are no un-representable objects, or objects that the laws of reason (which include but are not limited to the laws of logic) apply to, then they have to admit that there could be sense made of such a proposal. In summary, in these passages I think Fichte is arguing that there is no sense to be made of the notion of truth independently of the notions of reason and representation.[457] This had been Fichte's position since at least the *Review of Aenesidemus*, as he states there that 'what is logically true for any intellect…is at the same time true in reality and there is no other truth than this'[458]. Beiser reads this, plausibly, as Fichte claiming that it is nonsense to conceive of truth as conforming to a thing-in-itself that exists apart from how it must be conceived by any rational being.[459]

Because of these passages, I believe Fichte would have been hostile to eternalism, which is the idea that there are propositions which are true or false timelessly.

Above, I argued that Fichte is not going to be worried about divine foreknowledge, because he does not think that God can be meaningfully said to know things in the same way that we

---

[455] K167/H163

[456] This is the so-called 'short argument for idealism' that one finds critically discussed in Ameriks. Breazeale replies to Ameriks in defence of such arguments. This also has connections with recent discussions in idealism, especially in Hofweber (2017) who argues that there could be no ineffable truths.

[457] Compare Davidson, who begins his posthumous Truth and Predication by saying 'Nothing would be true or false if there were not rational beings to think so' (2005, 7). The thinking I attribute to Fichte here is similar to the position taken in Rödl (2018, esp.147f), but Rödl confines his attention to the idea that the laws of logic (being the laws of thought) have to apply to everything in principle.

[458] Fichte *Review* I:20

[459] Beiser, (2002, 244)

do. But the problem of foreknowledge is only superficially a problem about God's knowing the future. It really is a problem about the truth about the future. An eternalist says something like: "If P is true at t, then P is true, for all t, and if P is false at t, then P is false, for all t".[460] In other words, a proposition never changes its truth value. Temporalism is the denial of this: the temporalist holds that there are at least some propositions that change in truth value.[461] If eternalism is true, then statements about the future have truth values now, and the failure to grasp those truth-values is a mere epistemic failure. There is nothing intrinsically different about future truth and past truth or present truth. For a temporalist, propositions about the future are intrinsically different – they have no truth values yet; any statement about the future is neither true nor false. Note that if temporalism is true, then there is nothing to be made of the idea of knowing the future, assuming knowledge is factive. We can run an argument like the following:

1) Factivity: If S knows that P, then P
2) So if S knows that P, where P is a future-tense sentence, then P
3) Temporalism: There are no true or false future-tense sentences
4) If there are no true or false future-tense sentences, then one cannot know the sentence to be true (reverse of factivity – If P is not true, then S cannot know that P)
5) So, for any future-tense proposition P, S cannot know that P

This strengthens a case that Fichte might have been sympathetic to. But why think temporalism is true on Fichtean grounds? We should look briefly at some characterisations of eternalism to see why. In Frege, for example, eternalism follows from what 'complete thoughts' are. That the leaf is green is not a complete thought; that the leaf is green at 10am Central European Time is a complete thought, because it specifies its time-reference. Now the specification of the time reference is important because then it is said to be timelessly true that the leaf is green at 10am Central European Time. This is because, in Frege, a proposition's truth and meaning are deeply tied together, and because a proposition's truth is tied to its reference (the True or the False), it cannot change this without changing the proposition that it is. It is clear why Fichte would deny this: it is dogmatism; the Fregean claims that the ultimate philosophical explanation terminates in a realm of Platonic abstract and eternal (or at least everlasting) objects, rather than subjects or the I.[462]

[460] Richard (1981) and Aronszajn (1994) are expositions of an eternalist and a temporalist view, respectively.
[461] Therefore the temporalism under consideration here differs from the temporalism in Schellenberg (2013) which is an epistemic position: that we ought to take into account our place in evolutionary history when considering knowledge.
[462] See Prior (1959) for a classic argument that some expressions, such as "Thank goodness that's over!" cannot be translated into statements with explicit time-references. Another such expression would be something like 'I knew this day would come'. It is implausible in the extreme to think that

We should also look at an argument that we can run in favour of temporalism against eternalism, given what Fichte would take himself to have shown. We can schematise it along the following lines:

1) If a proposition is true at T, then it is always true (eternalist assumption)
2) So that I acted Xly at T was always true, even before T, and before I was born
3) If a proposition is true at T, where T is before my birth, then I cannot change P, nor can I be responsible for P.
4) If I cannot change or be responsible for P, then I am not free with respect to P.
5) On eternalism, I am never free with respect to P, for all P
6) But I am free, so eternalism is false

There are a number of elements in this argument.[463] The key move is 3 and 4. 3 seems clearly true – if something was true even before I was born and could never not be true, because propositions if true at one time are true at all times, then it is difficult to see what kind of control I could be said to have over the truth of that proposition. But the exercise of freedom is the exercise of control over contingent facts. So this seems straightforwardly in conflict with eternalism. In other words, the charge is that eternalism leads to necessitarianism about truth. This is charge 1. The second charge is under the surface, but can be framed like this. If eternalism is true, then propositions that concern me are true without me (because true before I was born). But then the only difference between us and the future is epistemic. The passage of time is merely a manner of finding out which propositions have always been true. But the future of rational agents and the world we live in is supposed to be at least partly up to us. It cannot be merely a matter of finding out that at T I move my arm in such and such a manner. It must be that I have made it true that I move my arm, and thereby know it non-observationally.

Kosch notes that these views find a place in Fichte's substantive ethics. In a discussion of Fichte's views of the morality of suicide and murder, Kosch argues that because Fichte thinks that there can be no purely moral grounds for rejecting a belief, but must always be some epistemic grounds (as we saw earlier in chapter 2 – the intellect is materially independent of the will), Fichte must have some epistemic grounds for thinking that, for example, "I am morally hopeless" or "It would be better if N did not exist" is morally wrong. These epistemic grounds

whatever the person means when uttering this is 'I knew that the 10th March 2000 would come'. Rather, they mean something like, "I knew that you would find me one day". But this is not transferable into a sentence with an explicit time reference which adequately captures the meaning. Prior's conclusion is that the content of thought should not be taken to be always eternalist in nature, which is a direct attack on Frege's position that it would be impossible to make sense of thought *unless* it was eternalist.

[463] This argument shares similarities with Van Inwagen's (1983) famous 'Consequence Argument'.

are exactly what I have been outlining here – that the future, at least as far as it concerns free action, is unknowable, because there are no facts to be known. Kosch says:

'That someone will continue along his path of vice is…something that it is not possible to know, because there is (as yet) no fact of the matter to be known. Indeed Fichte seems to think that we are, strictly, entitled to no beliefs at all about how others will conduct themselves in the future. Our own future actions are similarly indeterminate and therefore unpredictable; and though we can determine these by our intentions, we are obligated to intend the good'[464]

Fichte's practical conviction, or the 'requirement of presumptive optimism', then is implied by his view that the future is radically open.

The eternalist is going to respond by arguing that there is a difference between that such-and-such is the case and the reason why such-and-such is the case. For example, the eternalist may well say that the proposition "N has toast at 10:00am GMT 1st October 2018" has always been true, but the reason why it has always been true is because N used their volition or will to make it so. This is similar to the various kinds of traditional answers to the divine foreknowledge problem. Those answers – Ockhamist or Molinist – generally suggest that the counterfactuals of freedom are true because of the kind of beings that we are, so the fact that God knows that I would act thus and so under certain circumstances is made true because of my will. Therefore God can know it from eternity (and therefore the proposition is true from eternity) even though I make it true.

But this kind of case is strictly disanalogous to the one currently under consideration. The Ockhamist or Molinist answer to divine foreknowledge only gets off the ground because there is a thinker and agent that can be said to know the propositions to be true, namely, God. But in this eternalist case we are considering, there is no such being (at least, there doesn't have to be for an eternalist, and many eternalists have been atheists). In other words, the eternalist is confronted with the problem of how something can be said to be true completely independently of that thing currently existing.

I therefore conclude that if eternalism requires anything like the Fregean picture of propositions, or any merely epistemic difference between past and future, then Fichte would wholeheartedly reject eternalism in favour of temporalism, the view that there are no facts about the future, and that propositions can change their truth values (from neither true nor false to true or to false).

As a summary so far, we have the results that God does not need to know the future, and Fichte has good reasons for thinking that there would be no future for God to know

[464] Kosch (2018, 57). Kosch thinks our future actions are 'determined' by our intentions, because she reads Fichte as thinking that spontaneity only enters at the level of deliberation. (2018, 15).

independently of it actually being so anyway. Thus Fichte has no need to worry about divine foreknowledge, or inevitability, and has reasons to think eternalism false. He also has a way of conserving divine providence. All of this leaves us in the position of being at once free but morally bound. I will finally discuss one passage of the *System of Ethics* in which Fichte seems to avow a kind of predestination, and claim that this is compatible with freedom.

One final point needs to be discussed. In the *System of Ethics*, there is a peculiar section where Fichte remarks about predestination and freedom. Fichte claims that they can be unified or 'perfectly united'.[465] The first thing to note here is that this discussion uses the word 'predestination' because of its associations with the eternal, rather than associations with God's will. Fichte says that the problem arises because some action (the action(s) involved in the summons) are supposed to happen in some sense in time, but in another sense outside of time – so it is a problem about the relation between time and eternity, in a sense. The temporal part is the actual summoning of the individual I which is to-be-summoned. The eternal part is the summoning for the summoned, that is, the *a priori* condition with is a part of the general series of conditions on self-consciousness. But because this is a condition on self-consciousness, and time only appears for self-consciousness, it cannot happen within time. There are actually, as I see it, three separate issues here. The first is the relationship between the temporal and eternal conditions in the summons. The second is the generalised account of the temporal/eternal relationship in action in general. The third is the relationship between the individual person I am and the way I act. I shall take them in turn.

Fichte's first solution is to point out that my actions may be determined from the point of view of another, but for myself they are not. This observation cannot carry the weight it needs to, as Fichte acknowledges, because it cuts both ways, as it were – just as one might conclude that after all I am right and my actions are not determined, it might be that after all everyone else is right and they are. So Fichte's solution to the summons-problem is to focus on the *a priori* nature of self-consciousness conditions – they involve, as he says, 'no time nor temporal sequence, no *one-after-another*, but everything *at once*'. This is a familiar point from the genetic deductions – the deduction is supposed to show the acts of the I, or better, the moments of the one act of self-positing. From the point of view of the summoned, then, it is of course necessary that such a summoning happens (because it is a necessary condition on self-consciousness) but not necessary that such-and-such an agent summoned me. As Fichte puts it 'What is determined is *what* I will experience, but not *from whom*'.[466] Insofar as summoning is a necessary condition, there must be parts of it which are universally shared across all rational agents. Fichte's use of the word 'experience' may be misleading – he does

[465] SE IV:228
[466] SE VI:227

not need to think that being-summoned has a particular phenomenal character that everyone will experience, just that there is such a thing as 'being-summoned' and 'having-been-summoned' which is universally applicable.

The second of the issues in this passage is the generalised account. Fichte claims that 'From this moment onward there again lies ahead of me an infinite number of *predestined* actions, among which I can choose. Both the possibility and the actuality of all these actions are predestined; but it is by no means predestined that the precise action I choose is supposed to be attached to the entire series (that is, to actions A, B, C, etc.) that constitutes my individuality up to this point, and so on *ad infinitum*'. This is a very odd claim. There are three broad ways to interpret it. We could either say 1) that everything about the action including the act-type is fixed, but the act-token is not fixed, 2) that only the time of the action is fixed, 3) that everything is fixed apart from the agent (and perhaps the patient). On the one hand, we could say that what Fichte means is that in our future there are many actions which we will perform, and at each moment it is determined which action-type we will perform, but not which token. This, however, is far too deterministic to really be Fichte's view, or be called a unification of freedom and predestination. It would mean that at some time in the future I will perform a certain action-type (raising my arm) and this is fixed and determined, but I have control over how fast I raise it or whether I raise it so that my elbow is at an acute or obtuse angle. This cannot be what Fichte means. The second way of reading the suggestion is that in the future there are determinate (real) actions, which are determined as to time, but perhaps not any other predicate. To take a step back for a moment, actions (*qua* species of event) can be determined in various ways: time, place, intention, motive, agent, patient, bodily movements and so on. The suggestion here then seems to be that time is fixed but the others are open – so that there is an action which happens at a time, but the agent who performs this action is left open. This is consistent insofar as 'It is necessary that some agent will perform action X' does not imply 'Necessarily, agent A will perform action X' even if 'Agent A performed action X' is true.[467]

The problem with the account is rather that it does not seem meaningful to say that there is an action which is fixed as to time but not as to any other predicate. Surely we would not say that in two possible worlds, one in which at the same time actions took place, but differed in every other respect, that the same actions took place. Consider an act of voting. Now we might say that the possible worlds are the same with respect to time of the vote. So the act 'Voting for X in election E (which takes place at T)' is set. But clearly one can vote out of different motives (self-interest, community interest, moral fervour, etc) and for some ends which may differ. One can also vote in various manners. In other words, it seems artificial at best for

[467] Strictly speaking, Fichte would have to say that the action has its end fixed as well, because no action is without an end, and so to meaningfully speak of the same action as possibly performed in different places by different agents the agents would have to share ends.

Fichte to separate in this way the action from its adverbial descriptions. In more technical language, the identity conditions for actions seem to include some sameness of adverbial description. Without the relevant descriptions in place, it just seems that 'action' would not refer to anything.

One could also read this passage in a slightly modified way from the second proposal, as saying that everything except the agent who acts (and possibly the patient who is acted upon) is the same. So the action is fixed with respect to time, place, motive, intention, bodily movement, and whatever else. But it is not fixed with respect to which causal history it should be joined to. That is another way of saying it is not fixed with respect to who performs the action. Would this account fare any better than the weaker reading above? This has more intuitive plausibility – as it seems more likely that we could compare two possible worlds and hold everything fixed except the agent and patient in a particular action. I think this is probably the best rendering of Fichte's thinking here. But it seems to conflict with his official line – that there exist determinate actions somewhere *out there* in the future, as objects of thought, without any thinker to posit them. To this extent, then, it is perhaps good that he later rejected these thoughts on predestination.

We still need to come to the third part of this. This is the issue of the relationship between the individual person I am and the way I act. Fichte says that I can only act in the way that I actually do, if I am to be the person that I am. This again is ambiguous. One way to take it is a moral sense – which is in line with Fichte's thinking on vocation – that if I were to abandon my vocation I would in some sense be abandoning the person I actually am. If 'can' is read in the permissiveness sense then this is the reading we get. There is also a metaphysical sense, that it is as a matter of identity-conditions that I act in the way I actually do. Here the seeming threat of determinism looms. But it actually turns out to be a trivial admission. Of course I can only act in the way that I actually do if I am to be the person I actually am, because what it is to be the person I actually am is to act in the way that I actually do. That doesn't show anything about any substantive view of my abilities *qua* free agent. Fichte here also hints at the inter-relational and inter-personal aspects of being a person – my relationships in some sense work to define me; so again if I did not stand in those relationships then I would be the person I actually am, but the dogmatist is going to have to work quite a bit harder to show that this involves any determinism.[468]

We can therefore see that Fichte is not concerned about issues to do with predestination as we ordinarily conceive of them. This is because it is not for him a proper question to ask

[468] As an analogy, the phrase 'P is true if it corresponds to the facts' is a good and correct way of stating what truth is, but carries with it (contrary to appearance) no substantive theory of what truth is, because every correspondence, coherence, identity, redundance etc theorist can and will accept such a formulation, as Walker (1989, 2, 21, 25) points out, with respect particularly to coherence theories, but the point generalises.

whether I am predestined to be good or bad. These things are just not *a priori* features of what it is to be a person. They are in a real sense up to me. Everyone has a vocation, but whether one responds to it appropriately is entirely in one's own power.

7) Conclusion

I have argued for a certain reading of Fichte's views on free will and related issues. In chapter 1, I argued that we can distil from Kant a coherent list of desiderata that Fichte wants to fulfil in his views on freedom. The most important one of these is having a coherent view that does not require noumenal entities or things-in-themselves. In chapter 2, I presented Fichte's form of the transcendental argument against determinism, in order to show his first philosophical move against the dogmatist, which can function as a call for the dogmatist to re-evaluate their own position and also give Fichte room to manoeuvre. I also gave an exegesis of Fichte's famous striving argument for the conclusion that practical reason is transcendentally more basic than theoretical reason. This can be true even if the faculties are necessarily identical. In chapter 3 I gave an exegesis and defence of the arguments that Fichte presents in the first and second deductions of the System of Ethics, supported by other texts from the later Jena period. This exegesis supports the view that Fichte is committed to real libertarian freedom, and that he rejects compatibilism. Because of Fichte's genetic method, he is also able to answer a metaphysical deduction and a transcendental deduction simultaneously, which means that he is able to move beyond the Kantian aporiea of being compelled to believe that we are free in a transcendental sense but not being able to know this to be true. In chapter 4 I turned to the question of Fichte on causation – an important area for anyone who believes that freedom is a kind of causal power. I argued that Fichte need not be an agent-causalist, and drew some parallels between Fichte's views and contemporary volitionism. In chapter 5 I gave an account of Fichte's views on moral evil, with particular attention to Schelling's brief criticisms as well as drawing out how great evil is possible, as well as the evils that occur by omission. I concluded that Fichte's account of evil as based on an assent to a natural drive or inertia is a coherent and plausible view. In chapter 6 I looked at a number of related issues concerning free will and time. I first argued that Fichte need not be worried about the possibility of a conflict between divine foreknowledge and free will, and that this potential worry is resolved by Fichte's denial that there are any truths about the future.

We can therefore see that Fichte has a coherent account of libertarian freedom as a causal power, which does not need to subscribe to the contentious doctrine of agent-causalism, as that is usually understood, which gives an account of how one can act both freely and wrongly, which comes with an affirmation of the truly open future, and which truly shows why Kant was right to call freedom the keystone concept in the system of reason.

Bibliography

Adorno, Theodor W. & Horkheimer, Max (2010[1944]) *Dialectic of Enlightenment* (London, Verso)

Allison, Henry, (1996[1989]) 'Kant's Refutation of Materialism', in *Idealism and Freedom: Essays on Kant's Theoretical and Practical Philosophy* (Cambridge, Cambridge University Press) 92-106

Ameriks, Karl, (2000) *Kant and the Fate of Autonomy*, (Cambridge, Cambridge University Press)

Armstrong, David M. (1993) *A Materialist Theory of the Mind* revised ed. (London, Routledge)

Anscombe, G. E. M. (1958) 'Modern Moral Philosophy' *Philosophy* 33(124): 1-19

Anselm of Canterbury (2000[1080-1085]) *Freedom of Choice [De Libertate Arbitrii]* in *Complete Philosophical and Theological Treatises of Anselm of Canterbury* trans. Jasper Hopkins & Herbert Richardson (Minneapolis MN, Arthur J. Banning Press)

Aquinas, Thomas, (1920) *Summa Theologica* 2nd Revised Edition, trans by the Fathers of the English Dominican Province, available at www.newadvent.org/summa (checked 30/5/19)

Aronszajn, Mark 'A Defense of Temporalism' *Philosophical Studies* 81(1): 71-95

Augustine of Hippo (2008) *Confessions* trans. Henry Chadwick (Oxford, Oxford University Press)

Ayer, A. J. (1982) 'Freedom and Necessity' in *Free Will* ed. Gary Watson (Oxford, Oxford University Press) 15-23

Baker, Lynne Rudder (2017) 'Naturalism and the Idea of Nature' *Philosophy* 92: 333-349

Beck, Lewis White (1965) 'Kant's Theory of Definition' in *Studies in the Philosophy of Kant* (Indianapolis, Bobbs-Merrill) 61-73

Beiser, Frederick (1987) *The Fate of Reason* (Cambridge MA, Harvard University Press)

Beiser, Frederick (2002) *German idealism 1781-1801: The Struggle Against Subjectivism* (Cambridge MA, Harvard University press)

Bonhoeffer, Dietrich (1959[1937]) *Creation and Fall: A Theological Interpretation of Genesis 1-3* trans. John C. Fletcher (London, SCM Press)

Brandom, Robert (2007) 'The Structure of Desire and Recognition: Self-consciousness and Self-constitution' *Philosophy and Social Criticism* 33(1): 127-150

Breazeale, Daniel (2003) 'Two Cheers for Post-Kantianism: A Response to Karl Ameriks' *Inquiry*, 46(2): 239-259

Breazeale, Daniel, (2013a) 'The Problematic Primacy of the Practical', in *Thinking Through the Wissenschaftslehre*: *Themes From Fichte's Jena Philosophy*, (Oxford, Oxford University Press) 404-440

Breazeale, Daniel, (2013b) 'The Interests of Reason', in *Thinking Through the Wissenschaftslehre*: *Themes From Fichte's Jena Philosophy*, (Oxford, Oxford University Press) 334-359

Breazeale, Daniel (2013c) 'The Divided Self and the Tasks of Philosophy' in *Thinking Through the Wissenschaftslehre*: *Themes From Fichte's Jena Philosophy*, (Oxford, Oxford University Press) 124-155

Breazeale, Daniel (2014) 'Against Nature? On The Status and Meaning of the Natural World in J. G. Fichte's Early *Wissenschaftslehre*' *Philosophia OSAKA* 9: 19-39

Breazeale, Daniel (2016) 'From Autonomy to Automata? Fichte on Formal and Material Freedom and Moral Cultivation' in Daniel Breazeale & Tom Rockmore (eds.) *Fichte's Addresses to the German Nation Reconsidered* (Albany, NY, SUNY Press) 21-54

Bruno, G. Anthony (2018) 'Genealogy and Jurisprudence in Fichte's Genetic Deduction of the Categories' *History of Philosophy Quarterly* 35(1): 77-96

Chisholm, Roderick (1976) *Person and Object: A Metaphysical Study* (London, George Allen & Unwin)

Chisholm, Roderick (1982) ‘Human Freedom and the Self’ in *Free Will* ed. Gary Watson (Oxford, Oxford University Press) 24-35

Chisholm, Roderick (1995) ‘Agents, Causes, and Events: The Problem of Free Will in *Agents, Causes, and Events: Essays on Indeterminism and Free Will* ed. Timothy O’Connor (Oxford, Oxford University Press)

Churchland, Patricia (1991) ‘Is Determinism Self-Refuting?’ *Mind* 90: 99-101

Clarke, Samuel (1717) *Remarks Upon a Book Entitled “A Philosophical Enquiry Concerning Human Liberty”* (London, James Knapton)

Clarke (2008) ‘Intrinsic Finks’ *The Philosophical Quarterly* 58: 512-518

Clarke, James (2013) ‘Fichte’s Critique of Rousseau’ *The Review of Metaphysics* 66: 495-517

Davidson (1963) ‘Actions, Reasons, and Causes’ *The Journal of Philosophy* 60(23): 685-700

Davidson, Donald (2005) *Truth and Predication* (Cambridge MA, Harvard University Press)

Dennett, Daniel, (1981) ‘The Abilities of Men and Machines’ in *Brainstorms* (Brighton, Harvester Press) 256-266

Di Giovanni, George & Harris, H. S. (1985) *Between Kant and Hegel: Texts In the Development of Post-Kantian Idealism* (Albany NY, SUNY Press, 1985)

Doris, John M. (2002) *Lack of Character* (Cambridge, Cambridge University Press)

Double, Richard (1996) *Metaphilosophy and Free Will* (Oxford, Oxford University Press)

Dryden, Jane (2013) ‘Embodiment and Vulnerability in Fichte and Hegel *Dialogue* 52: 109-128

Everett, Charles Carroll (1909) *Theism and the Christian Faith* ed. Edward Hale (London, Macmillan)

Farrer, Austin (1958) *The Freedom of the Will* (Adam & Charles Black, London)

Farrer, Austin (1972) ‘Grace and the Human Will’ in *Reflective Faith: Essays in Philosophical Theology* ed. Charles C. Conti (London, SPCK)

Fenn, William W. (1910) ‘The Theology of Charles Carroll Everett’ *The Harvard Theological Review* 3(1): 1-23

Fichte, J. G. (2001a[1792]), ‘Review of Leonhard Creuzer, *Skeptical Reflections on the Freedom of the Will, with Reference to the Latest Theories of the Same*, with a foreword by Professor Schmid’ trans. Daniel Breazeale, *The Philosophical Forum*, 32(4) (2001[1792]): 289-296

Fichte, Johann Gottlieb, (2001b[1792]) ‘Review of Friedrich Heinrich Gebhard, *On Ethical Goodness as Disinterested Benevolence’*, trans. Daniel Breazeale, *The Philosophical Forum*, 32(4): 297-310

Fichte, Johann Gottlieb, (2010[1792/1793]) *Attempt at a Critique of All Revelation*, trans. Garrett Green, ed. Allen Wood, (Cambridge, Cambridge University Press)

Fichte, J. G. (1982[1794]) *Foundations of the Entire Science of Knowledge* trans. Peter Heath & John Lachs (Cambridge, Cambridge University Press)

Fichte, J. G., (1996[1795]) ‘On the Linguistic Capacity and the Origin of Language’ trans. Jere Paul Surber, in *Language and German Idealism: Fichte’s Linguistic Philosophy* (New York, Humanity Books)

Fichte, J. G. (1988) *Early Philosophical Writings* ed. & trans. Daniel Breazeale (Ithaca & London, Cornell University Press)

Fichte, J. G. (2005[1798]) *System of Ethics* trans. Gunter Zoeller & Daniel Breazeale (Cambridge, Cambridge University Press)

Fichte, J. G. (1994) *Introductions to the Wissenschaftslehre and Other Writings 1797-1800* ed. & trans. Daniel Breazeale (Indianapolis, Hackett)

Fichte, J. G. (1992[1796/9] *Foundations of Transcendental Philosophy: Wissenschaftslehre Nova Methodo* ed. & trans. Daniel Breazeale (Ithaca & London, Cornell University Press)

Fichte, J. G. (2010[1799]) *Appeal to the Public* in Yolanda Estes (ed.) and *Curtis Bowman* (trans.) *J. G. Fichte and the Atheism Dispute (1798-1800)* (Farnham, Ashgate)

Fichte, J. G. (1987[1800]) *The Vocation of Man* trans. Peter Preuss (Indianapolis, Hackett)

Fichte, J. G. (1889a[1805]) *The Characteristics of the Present Age* in *The Popular Works of Johann Gottlieb Fichte* trans. William Smith (London, Trubner)

Fichte, J.G. (1889b[1806]) *The Way Toward the Blessed Life, or The Doctrine of Religion* in *The Popular Works of Johann Gottlieb Fichte* trans. William Smith (London, Trubner)

Fichte, J. G. (2008[1807-8]) *Addresses to the German Nation* trans. & ed. Gregory Moore (Cambridge, Cambridge University Press)

Förster, Eckhart (2012) *The Twenty-Five Years of Philosophy: A Systematic Reconstruction* trans. Brady Bowman (Cambridge, MA & London, Harvard University Press)

Franks, Paul (2005) *All or Nothing: Systematicity, Transcendental Arguments, and Skepticism in German Idealism* (Cambridge MA, Harvard University Press)

Franzel, Sean (2009) '"Welches Gesetz ist der Mensch in seiner Wirksamkeit?": Pedagogy and Media in Fichte's Encounter with Mesmerism' *The Germanic Review* 3-25

Gardner, Sebastian (2007) 'The Limits of Naturalism and the Metaphysics of German Idealism' in Hammer, Espen (ed) *German Idealism: Contemporary Perspectives* (Oxford, Routledge) 19-49

Gardner, Sebastian (2016) 'Kant's Third *Critique*: The Project of Unification' *Royal Institute of Philosophy Supplement* 78: 161-185

Gardner, Sebastian (2017) 'The Metaphysics of Human Freedom: From Kant's Transcendental Idealism to Schelling's *Freiheitsschrift*' *British Journal for the History of Philosophy* 25(1): 133-156

Gardner, Sebastian (2018) 'The Desire of the Whole in Classical German Philosophy' in Dina Emundts, Sally Sedgewick (eds.) *Begehren/Desire* (Berlin & Boston, De Gruyter) 233-256

Giladi, Paul (2014) 'Liberal Naturalism: The Curious Case of Hegel' *International Journal of Philosophical Studies* 22(2): 248-270

Girndt, Helmut (1992) 'Evolutionstheorie und transzendentaler Gedanke: Ein wissenschafts-theoretischer Problemaufriß' *Fichte-Studien* Vol.4, 68-85

Goh, Kien-How (2012) 'Between Determinism and Indeterminism: The Freedom of Choice in Fichte's Das System Der Sittenlehre' *European Journal of Philosophy* 23(3):439-455

Guyer, Paul (1987) *Kant and the Claims of Knowledge* (Cambridge, Cambridge University Press)

Guyer, Paul (2005) 'From Nature to Morality: Kant's New Argument in the "Critique of Teleological Judgement"' in *Kant's System of Nature and Freedom* (Oxford, Oxford University Press)

Guyer, Paul, (2015) 'Fichte's Transcendental Ethics' in Sebastian Gardner & Matthew Grist (eds.) *The Transcendental Turn* (Oxford, Oxford University Press), 135-158

Guyer, Paul (2017) 'The Struggle for Freedom: Freedom of Will in Kant and Reinhold' in Eric Watkins (ed.) *Kant on Persons and Agency* (Cambridge, Cambridge University Press, 2017)

Haldane, J. B. S. (1927. Repr. 1932) 'When I am Dead' in *Possible Worlds: And other essays* (London, Chatto & Windus) 204-210

Hanson, Norwood Russell, (1959) 'Copernicus' role in Kant's Revolution', *Journal of the History of Ideas*, Vol. 20, No. 2, 274-281

Harman, Gilbert, (1970) 'Knowledge, Reasons, and Causes' *The Journal of Philosophy*, Vol. 67, No. 21, 841-855

Hasker, William, (1973) 'The Transcendental Argument against Determinism' *Southern Journal of Philosophy* Vol. 11, No. 3, 175-183

Hasker (1999) *The Emergent Self* (Ithaca, Cornell University Press)

Hegel, G. W. F. (1995), *Lectures on the History of Philosophy: Vol. 3* trans. E. S. Haldane & Frances H. Simson (Lincoln, University of Nebraska Press)

Henrich, Dieter (2003) *Between Kant and Hegel: Lectures on German Idealism* ed. David S. Pacini (Cambridge MA & London, Harvard University Press)

Herman, Barbara (2009) 'A Habitat for Humanity' in *Kant's Idea For a Universal History with a Cosmopolitan Aim: A Critical Guide* ed. A. E. Rorty & James Schmidt (Cambridge, Cambridge University Press) 150-170

Hobart, R. E. (1934) ‘Free Will as Involving Determinism and Inconceivable Without It’ *Mind* 43(169): 1-27

Hofweber, Thomas (2017) ‘Conceptual Idealism Without Ontological Idealism: Why Idealism Is True After All’ in Tyron Goldschmidt & Kenneth L. Pearce (eds.) *Idealism: New Essays in Metaphysics* (Oxford, OUP) 124-141

Hoeltzel, Stephen (2014) ‘Nonepistemic Justification and Practical Postulation in Fichte’ in Rockmore, Tom & Breazeale, Daniel (eds) *Fichte and Transcendental Philosophy* (Hampshire, Palgrave Macmillan)

Holton, Richard (2009) ‘Determinism, Self-Efficacy, and the Phenomenology of Free Will’ *Inquiry* 52(4): 412-428

Kant, Immanuel, (1990[1781/1787]) *Critique of Pure Reason*, translated by Norman Kemp Smith, (London, Macmillan)

Kant, Immanuel, (1998[1795]) *Groundwork for the Metaphysics of Morals*, in *Practical Philosophy* trans. Mary Gregor (Cambridge, Cambridge University Press)

Kant, Immanuel (1996) *Religion and Rational Theology* trans. & ed. Allen W. Wood & George di Giovanni (Cambridge, Cambridge University Press)

Kant, Immanuel (2000[1790]) *Critique of the Power of Judgement* ed. Paul Guyer, trans. Paul Guyer & Eric Matthews (Cambridge, Cambridge University Press)

Kant, Immanuel, (1998[1787/8]) *The Metaphysics of Morals* in *Practical Philosophy*, trans. Mary Gregor (Cambridge, Cambridge University Press)

Kemp, R. S. ‘How to Become an Idealist: Fichte on the Transition from Dogmatism to Idealism’ *British Journal for the History of Philosophy* 25:6: 1161-1179

Kinlaw, C. Jeffrey (1992) ‘Fichte’s Kenotic Christology’ *Idealistic Studies* 22(1): 39-51

Kinlaw, C. Jeffrey (2002) ‘Reflection and Feeling and the Primacy of Practical Reason in the Jena *Wissenschaftslehre*’ in Breazeale, Daniel & Rockmore, Tom (eds) *New Essays on Fichte’s Later Jena Wissenschaftslehre* (Evanston, IL, Northwestern University Press)

Korsgaard, Christine (1996) *The Sources of Normativity* (Cambridge, Cambridge University Press)

Korsgaard, Christine (2009) *Self-Constitution: Agency, Identity, and Integrity* (Oxford, Oxford University Press)

Kosch, Michelle (2015a) 'Agency and Self-Sufficiency in Fichte's Ethics' *Philosophy and Phenomenological Research* 91(2) 348-380

Kosch, Michelle, (2015b) 'Fichtean Kantianism in Nineteenth-Century Ethics', *Journal of the History of Philosophy*, Vol. 53, No. 1 (2015), 111-132

Kosch, Michelle (2018) *Fichte's Ethics* (Oxford, Oxford University Press)

Kripke, Saul (1980) *Naming and Necessity* (Malden, Blackwell)

Laird, John, (1917) *Problems of the Self* (London, Macmillan)

Langlois, Luc (2017) 'The Meaning of Life According to Fichte (1796-1800)' *Fichte Studien* 44(1): 17-30

Langton, Rae (1999) *Kantian Humility* (Oxford, Oxford University Press)

Leibniz, G. W. (1989) 'Meditation on Knowledge, Truth, and Ideas' in *Philosophical Writings* ed. Roger Ariew & Daniel Garber (Indianapolis, Hackett) 23-27

Leslie, John (1979) *Value and Existence* (Oxford, Blackwell)

Lockie, Robert (2018) *Free Will and Epistemology: A Defence of the Transcendental Argument for Freedom* (London, Bloomsbury Academic)

Lowe, E. J. (2008) *Personal Agency* (Oxford, Oxford University Press)

Lucas, J. R. (1961) 'Minds, Machines, and Gödel' *Philosophy* 36(137): 112-127

Lucas, J. R. (1970) *The Freedom of the Will,* (Oxford, Oxford University Press)

Luther, Martin (1957) 'Disputation Against Scholastic Theology' trans. Harold J. Grimm in *Luther's Works Vol 31: Career of the Reformer I* ed. Harold J. Grimm (Philadelphia, Concordia Press) 9-16

Luther, Martin (1979) *On the Bondage of the Will [De Servo Arbitrio]* trans. & ed. Philip S. Watson & B. Drewery in *Luther and Erasmus: Free Will and Salvation* (London, SCM Press)

Lütterfelds, Wilhelm (1992) 'Hat die Idealistische Erfahurng eine naturalistische Basis? Fichtes frühe transzendentalphilosophie und die evolutionäre Erkenntnistheorie' *Fichte-Studien* 4: 86-118

Moore, A. W. (2012) *The Evolution of Modern Metaphysics* (Oxford, Oxford University Press)

Mann, William (2015) 'Divine Simplicity' in *God, Morality, and Modality* (Oxford, Oxford University Press) 20-43

Martin, C. B. (1994) 'Dispositions and Conditionals' *The Philosophical Quarterly* 44:1-8

Martin, Wayne M., (1996) "Without a Striving, No Object is Possible" Fichte's Striving Doctrine and the Primacy of Practice', in Daniel Breazeale & Tom Rockmore (eds) *New Perspectives on Fichte* (New Jersey, Humanities Press) 19-33

Martin, Wayne, (1997) *Idealism and Objectivity: Understanding Fichte's Jena Project* (Stanford, Stanford University Press)

Martin, Wayne (2009) 'Ought but Cannot' *Proceedings of the Aristotelian Society* CIX, 103-128

McTaggart, John M. E., (1934) 'Introduction to the Study of Philosophy' in *Philosophical Studies* ed. S. V. Keeling (London, Edward Arnold), 183-210

Meier, Georg, Freidrich (2016) *Excerpt From the Doctrine of Reason* trans. Aaron Bunch, Axel Gelfert, Riccardo Pozzo (London, Bloomsbury Academic)

Midgley, Mary (1984) *Wickedness: A Philosophical Essay* (London, Routledge & Kegan Paul)

Millikan, Ruth, (1984) *Language, Thought, and Other Biological Categories: New Foundations for Realism* (Cambridge, MIT Press)

De Molina, SJ, Luis (1988) *On Divine Foreknowledge (Part IV of the Concordia)* trans. Alfred J. Freddoso (Ithaca, Cornell UP)

Moore, G. E. (1898) 'Review' *International Journal of Ethics* 9(1): 92-97

Neuhouser, Frederick, (1990) *Fichte's Theory of Subjectivity* (Cambridge, Cambridge University Press)

O'Connor, Timothy 'Agent Causation' in *Agents, Causes, and Events: Essays on Indeterminism and Free Will* ed. Timothy O'Connor (Oxford, Oxford University Press, 1995) 173-200

Paul, L. A. (2014) *Transformative Experience* (Oxford, Oxford University Press)

Paul, L. A. (2015) 'What you can't expect when you're expecting' *Res Philosophica* 91(2): 1-23

Pearce, Stuart & Pickard, Hanna (2010) 'Finding the Will to Recover: Philosophical Perspectives on Agency and the Sick Role' *Journal of Medical Ethics* 36(12): 831-833

Perry, John (1979) 'The Problem of the Essential Indexical' *Nous* 13(1) 3-21

Perovich Jr., Anthony N. 'Fichte and the Typology of Mysticism' in *Fichte: Historical Contexts/Contemporary Controversies,* Daniel Breazeale & Tom Rockmore (eds.) (New Jersey, Humanities Press, 1994) 128-141

Pickard, Hanna (2015) 'Psychopathology and the Ability to Do Otherwise' *Philosophy and Phenomenological Research* 90(1), 135-163

Pinkard (2002) *German Philosophy 1760-1860: The Legacy of Idealism* (Cambridge, Cambridge University Press)

Plantinga, Alvin (2000) *Warranted Christian Belief* (Oxford, Oxford University Press)

Plantinga, Alvin, (2004) 'Evolution, Epiphenomenalism, Reductionism' *Philosophy and Phenomenological Research* Vol. 68, No. 3, 602-619

Prior, A. N. (1959) ‘Thank Goodness That’s Over’ *Philosophy* 34(128) 12-17

Quine, W. V. O. (1980) ‘Two Dogmas of Empiricism’ in *From a Logical Point of View* (Harvard, Harvard University Press) 20-46

Reinhold, Karl Leonhard (2011[1789]) *Essay on a New Theory of the Human Capacity for Representation* trans. Tim Heigan & Barry Empson (Berlin & New York, de Gruyter)

Richard, Mark (1981) ‘Temporalism and Eternalism’ *Philosophical Studies* 39(1) 1-13

Rockmore, Tom (1980) *Fichte, Marx and the German Philosophical Tradition* (Carbondale IL, Southern Illinois University Press)

Rockmore, Tom (2010) ‘Fichte, German Idealism and the Thing in Itself’ in Breazeale, Daniel & Rockmore, Tom, *Fichte, German Idealism, and Early Romanticism* (New York, Rodopi) 9-20

Rödl, Sebastian (2018) *Self-Consciousness and Objectivity: An Introduction to Absolute Idealism* (Cambridge MA & London, Harvard University Press)

Schellenberg, J. L. (2013) *Evolutionary Religion* (Oxford, Oxford University Press)

Schelling, F. W. J. (1978[1800]) *System of Transcendental Idealism* trans. Peter Heath (Charlottesville, VA, University of Virginia Press)

Schelling, F. W. J. (2006[1809]) *Philosophical Investigations into the Essence of Human Freedom* trans. Jeff Love & Johannes Schmidt (Albany NY, SUNY Press)

Schopenhauer, Arthur, (1988[1811-2]) ‘On the Facts of Consciousness and the Doctrine of Science’ in *Manuscript Remains, Vol. II: Critical Debates (1809-1818)* ed. Arthur Hübscher, trans. E. J. F. Payne (Oxford, Berg)

Scribner, F. Scott (2010) *Matters of Spirit: J. G. Fichte and the Technological Imagination* (University Park, Pennsylvania State University Press)

Sedgewick, Sally, (2000) ‘Introduction: Idealism From Kant to Hegel’ in *The Reception of Kant’s Critical Philosophy: Fichte, Schelling, and Hegel* (ed) (Cambridge, Cambridge University Press) 1-18

Seidel, George J. (1993) *Fichte's Wissenschaftslehre of 1794: A Commentary on Part 1* (Indiana, Purdue University Press)

Slagle, Jim (2016) *The Epistemological Skyhook: Determinism, Naturalism, and Self-Defeat* (New York & Oxford, Routledge)

Stang, Nicholas (2016) *Kant's Modal Metaphysics* (Oxford, Oxford University Press)

Stroud, Barry, (1968) 'Transcendental Arguments' *The Journal of Philosophy* Vol. 65, No. 9, 241-256

Talbot, Ellen Bliss, (1907) 'The Philosophy of Fichte in its Relation to Pragmatism' *The Philosophical Review*, Vol. 16, No. 5, 488-505

Temple, William (1917) *Mens Creatrix* (London, Macmillan)

Van Cleve, James (1990) *Problems from Kant* (Oxford, Oxford University Press)

Van Inwagen, Peter (1983) *An Essay on Free Will* (Oxford, Oxford University Press)

Vater, Michael G. & Wood, David W. (2012) *The Philosophical Rupture Between Fichte and Schelling* (Albany NY, SUNY Press)

Walker, Ralph C. S. (1989) *The Coherence Theory of Truth: Realism, Anti-Realism, Idealism* (London, Routledge)

Ware, Owen (2015) 'Agency and Evil in Fichte's Ethics' 15(11): 1-21

Ware, Owen (2018) 'Fichte's Normative Ethics: Deontological or Teleological?' *Mind* 127, 565-584

Wick, Warner (1964), 'Truth's Debt to Freedom' *Mind* Vol. 73, No. 292, 527-537

Wood, Allen (1984) 'Kant's Compatibilism' in Allen Wood (ed.) *Self and Nature in Kant's Philosophy* (Ithaca, Cornell University Press) 73-101

Wood, Allen, (2000) ‘The “I” as the Principle of Practical Philosophy’ in *The Reception of Kant’s Critical Philosophy: Fichte, Schelling, and Hegel* (ed) (Cambridge, Cambridge University Press), 93-108

Wood, Allen W. (2014) ‘The Evil in Human Nature’ in Gordon E. Michalson (ed.) *Kant’s Religion Within the Boundaries of Mere Reason: A Critical Guide* (Cambridge, Cambridge University Press)

Wood, Allen, (2016) *Fichte’s Ethical Thought* (Oxford, Oxford University Press)

Zanelotti, Kevin (2008) ‘How Not to Read Fichte’s *Anweisung zum Seligen Leben* (1806): Against the Mystical Reading’ in Daniel Breazeale & Tom Rockmore (eds.) *After Jena: New Essays on Fichte’s Later Philosophy* (Evanston, IL, Northwestern University Press) 131-150

Zöller, Gunter, (1997) ‘An Eye for an I: Fichte’s transcendental experiment’ in *Figuring the Self: Subject, Absolute, and Others in Classical German Philosophy*, David E. Klemm & Gunter Zöller (eds.), (New York, SUNY Press), 73-95

Zöller, Gunter, (1998) Fichte’s Transcendental Philosophy: The Original Duplicity of Intelligence and Will (Cambridge, Cambridge University Press)