Abstract: In *Groundwork* III, Kant attempts to give a deduction of the categorical imperative. There is widespread disagreement as to how Kant’s argument is supposed to proceed. Many commentators believe that Kant’s deduction fails because some of its argumentative moves are unjustified. In particular, Kant makes a mistaken inference from theoretical freedom to practical freedom, and his axiological ‘superiority claim’ regarding the noumenal world’s priority over the sensible world is unjustified. According to the standard incompatibilist story, Kant came to see that his deduction was flawed by the time he wrote the *Critique of Practical Reason*, at which point he claimed that the truth of the moral law does not require a deduction since it is a “fact of pure reason”. The moral law is no longer the conclusion of his argument; instead, it functions as the premise of an argument that establishes our freedom. Other commentators endorse a compatibilist reading, according to which the justifications of the moral law in *Groundwork* III and the second *Critique* are compatible because Kant never attempted to give the strong kind of deduction that he rightly rejects in the second *Critique*. On the view I develop here, the particular argumentative moves that the standard incompatibilist takes issue with are not flawed and incompatible with Kant’s second *Critique*. I argue for a compatibilist reading of these moves. I think the compatibilist is right to claim that the deduction Kant considered impossible in both the *Groundwork* and the second *Critique* is what I call a strong deduction. I also agree with compatibilists that the deduction he actually delivers in *Groundwork* III is only a weak deduction that makes use of a merely problematic conception of transcendental freedom. However, I do think that Kant’s argument in *Groundwork* III remains question begging in the final analysis. The facticity claim in the second *Critique*, by contrast, can provide a non-question-begging account of moral obligation. Here, I agree with the optimistic incompatibilist, who views the argument in the second *Critique* as an improvement on his argument in the *Groundwork*. However, in my novel account of Kant’s argument, I endorse what I call ‘radical incompatibilism’ because it concerns the roots of Kant’s approach to the justification of the moral law. What is novel about my account is the claim that the deduction in *Groundwork* III rests on the false assumption that practical cognition, like theoretical cognition, requires a critique of pure reason. In the second *Critique*, Kant

*Kontakt: Dr. Jochen Bojanowski*, University of Illinois at Urbana-Champaign, 105 Gregory Hall, 810 South Wright Street, Urbana, IL 61801, USA; job39@illinois.edu
revised his argument because he realized that, in contrast to synthetic a priori judgments of theoretical cognition, the possibility of synthetic a priori judgments of practical cognition can be derived from the actuality of a “deed”. With respect to pure practical reason, the second Critique proceeds metaphysically – i. e. dogmatically – rather than critically. Hence Kant came to view a deduction of the categorical imperative as unnecessary and abandoned the project of a critique of pure practical reason. We should, for this reason, resist the generality of Kant’s claim in the first Critique to the effect that, for all synthetic judgments a priori, “if not a proof then at least a deduction of the legitimacy of its assertion must unfailingly be supplied” (KrV, B 286).

**Keywords:** Kantian Ethics, Fact of Reason, Deduction of Freedom, Deduction of the Categorical Imperative, Groundwork III

It seems rather intuitive that the moral principle “Act in such a way that the maxim of your action can be at the same time a universal law” should require justification. Its justification should address at least two different sorts of scepticism: content scepticism and normative scepticism. A response to content scepticism first needs to provide an account of the formal content of the moral principle in question, and it needs to show how this principle underlies our particular moral judgments. Kant takes content scepticism to be a problem of mere analysis. He claims that the concept of ‘unconditional moral obligation’ necessarily leads to the formal principle in question (GMS, AA 04: Sec. I and II; KpV, AA 05: §§ 1–7). In oft-cited examples, he tries to make explicit that this principle is the form of all our particular moral judgments (GMS, AA 04: 422 f.; KpV, AA 05: 44).

However, even if Kant did have a satisfactory response to content scepticism, the categorical imperative could still be, as he puts it, a “chimerical idea without any truth” (GMS, AA 04: 445). The formal principle Kant provides might be an adequate articulation of unconditional moral obligation, but it is still an open question whether this principle has any bindingness on us as morally imperfect rational agents (i. e. whether it is true for us). Kant needs to justify why we are “necessarily bound” to the moral law as an imperative (GMS, AA 04: 440). I will call this kind of scepticism, which Kant tries to address in the third section of the *Groundwork*, normative scepticism. The goal of the “deduction” in *Groundwork* III is to explain (by means of a “critique of pure practical reason”) “on what grounds (woher) the moral law is binding” for us (GMS, AA 04: 450).

According to the standard story, Kant’s deduction in the *Groundwork* fails. The reasons for its failure are not unanimously agreed on, but two points of criticism remain perennial. The first is Kant’s illegitimate move from the spontaneity
of reason in its theoretical employment to practical freedom, and from there to
the moral law.\(^1\) The second is his question-begging axiological superiority claim
about the priority of the noumenal over the sensible world, which he mistak-
enly took to be an implication of his transcendental idealism.\(^2\) As is well known,
Kant’s deduction in the *Groundwork* was not his last word on the topic. As the
standard story continues, Kant realized in his second *Critique* that the deduction
in the *Groundwork* had failed. He admits failure, dismisses the “theoretical proof”
of freedom, and instead derives practical freedom from the moral law, thereby
reversing the order of his initial deduction. Kant now claims that the “reality of
the moral law cannot be shown by the means of a deduction” because it is “firmly
established of itself” (KpV, AA 05: 47). Consciousness of the moral law, Kant now
claims, cannot be “reason[ed] out” (*herausvernünfteln*) from “antecedent data of
reason”, but instead “forces itself on us as a synthetic a priori judgment” (KpV,
AA 05: 31). The truth of the moral law is no longer the conclusion of his argument
but rather functions as the premise of an argument that establishes our freedom.

In his widely received article, Dieter Henrich not only endorses the standard
story but also holds that the fact of pure reason is superior to his earlier argument
and so brings the story to a happy ending. He does, however, admit that Kant’s
deduction in the *Groundwork* is one of the “darkest” passages in the entire history
of philosophy.\(^3\) Henrich’s article provides only a prelude to what has become one
of the most researched texts in Kant’s entire oeuvre. In the past forty years since
the publication of Henrich’s article, Kant scholars have shed considerable light
on these “obscure” passages.\(^4\) Kant’s facticity claim in the second *Critique* has
also received growing attention.\(^5\) We now have a much more nuanced controversy
over the proof structure in *Groundwork* III, Kant’s facticity claim, and the relation-
ship between the two. The debate about the relationship between the *Ground-
work* and the second *Critique* can be roughly divided into two camps. The one
camp holds that the deduction in the *Groundwork* and Kant’s facticity claim in the
second *Critique* are compatible. For lack of a better term, this camp may be called
the compatibilist camp. Compatibilists hold that, both in the second *Critique* and
in the *Groundwork*, Kant only rejected what I will call a strong deduction. A strong
deduction would be grounded in a theoretical cognition of human freedom, the

---

2 Ameriks 2003, 177; Schönecker; Wood 2007, 205 f.; Timmermann 2007, 140 f.; Sussman 2008,
63.
3 Henrich 1975.
4 McCarthy 1982; Brandt 1988; Schönecker 1999; Ludwig 2008; Timmermann 2010; Allison 2012.
possibility of which Kant consistently rejected. The weak deduction, by contrast, which Kant developed in the *Groundwork*, is one that he never gave up. Kant’s deduction in the *Groundwork* and his facticity claim in the second *Critique* can be seen as complementary rather than mutually exclusive. The compatibilist camp is small, and its story is clearly a deviation from the standard story. The majority of scholars join Henrich in denying the compatibility of the deduction in *Groundwork* III and Kant’s facticity claim. Their account of the argument in *Groundwork* III is much more nuanced and refined than Henrich’s. In contrast to Henrich, many of them also give us good reason to believe that the fact of pure reason is not, as Hegel claimed, the “last undigested log” of Kant’s philosophy. But in spite of the progress we have made, the standard story concerning the development of Kant’s thought from the *Groundwork* to the second *Critique* has prevailed.

In this paper, I would like to show why there is an important lesson to be learned from both of these positions. Or, to put it negatively, I think we can resist the standard story without thereby falling into compatibilism. The alternative radical incompatibilist story I want to develop in this paper runs as follows. Kant’s deduction in the *Groundwork* was based on the false belief that both practical cognition and theoretical cognition require a critique of pure reason. In the *Groundwork*, Kant still thought it necessary to explain how the categorical imperative, as a synthetic judgment a priori, is possible. He gave up his deduction not because, as the standard incompatibilist story has it, he mistakenly inferred practical freedom from theoretical freedom in *Groundwork*. Nor did he abandon his so-called “superiority claim” regarding the priority of the noumenal over the sensible world. What Kant came to realize in the years between the *Groundwork* and the second *Critique* was that, in contrast to theoretical reason, practical reason can prove the objective reality of its synthetic a priori judgments through a “deed”. And since this deed is already an actuality, the question of its possibility – the very question the *Groundwork* was supposed to answer – does not need to be addressed. Compatibilists are right to claim that the deduction Kant

---

6 Beck 1960; McCarthy 1982; Wolff 2009.
7 Hegel 1836, 369.
9 Sussmann 2008; Schönecker 1999; Timmermann 2010; Allison 2012; Ameriks 2003. Heiner Klemme (2010) also belongs to this camp, but on his novel account the main difference between the *Groundwork* and the second *Critique* is that Kant came to recognize that an antinomy of pure practical reason arises even within practical cognition, and that it requires its own critical solution. Klemme raises an important issue, but in this paper I want to confine myself to the problem of the deduction of the categorical imperative.
considered impossible both in the *Groundwork* and in the second *Critique* is what I call a *strong* deduction. A strong deduction takes as its starting point theoretical cognition of transcendental freedom. By contrast, the deduction he delivers in the *Groundwork* is a weak deduction, which is based on a merely problematic conception of transcendental freedom. This does not mean that Kant could still endorse the argument in *Groundwork* III if a critique of pure practical reason were required. Since Kant moves in this argument from our conception of pure practical reason to transcendental freedom, and from there to our consciousness of our autonomy, the argument in *Groundwork* III ultimately begs the question. His argument in the second *Critique*, by contrast, can give a non-question-begging account, and in this respect we should view it as an improvement on his argument in the *Groundwork*. However, the radical difference between his approaches in the *Groundwork* and in the second *Critique* concerns the very idea of a critique of pure practical reason, and with it the very nature of pure practical cognition. This is why his attempt at a deduction in *Groundwork* III and his facticity claim are fundamentally incompatible. For the same reason, we should resist the generality of Kant’s claim from the first *Critique* to the effect that, for all synthetic judgments a priori, “if not a proof then at least a deduction of the legitimacy of its assertion must unfailingly be supplied” (KrV, B 286).

## 1 The Task of a Deduction

Kant maintains that the categorical imperative is a synthetic judgment a priori (GMS, AA 04: 420; KpV, AA 05: 31, 46). The demand for a justification, or “deduction”, arises from the fact that, for Kant, only in synthetic judgments does the meaning of the predicate term go beyond the meaning of the subject term.\(^\text{10}\) So,

---

\(^{10}\) Bernd Ludwig questions whether Kant ever attempted to give a deduction of the categorical imperative. He calls the expression “deduction of the categorical imperative” an “artefact of the history of Kant interpretation”. The real objects of the deduction in the *Groundwork* are “freedom”, the “moral law”, and the “idea of a pure will”, which are then supposed to explain how a categorical imperative is possible (Ludwig 2008, 457). Admittedly, Kant’s use of ‘deduction’ is not very precise. However, when Kant speaks of the deduction of the principle of pure practical reason in the second *Critique*, he equates it with the question of how this principle is possible, which is the very question he asks in the third section of the *Groundwork* (KpV, AA 05: 46). The question is asking for a justification of the binding force of the moral law. An answer to this question is supposed to show why the moral law is an imperative. So there does seem to be at least some textual support for describing Kant’s project as a deduction of the moral law as a categorical imperative.
only in synthetic judgments do we learn something we did not already know when we know the meaning of the subject term. In order to justify the connection of the predicate term with the subject term, we cannot simply refer to the meaning of those terms. What we need instead is a “third thing”, as Kant puts it, which functions as the justificatory ground of the synthetic judgment. In the case of synthetic judgments a posteriori, this “third thing” is actual experience (KrV, B 193 ff.; cf. B 778). The categorical imperative, however, is not a synthetic judgment a posteriori but rather a synthetic judgment a priori; hence an empirical justification of this judgment via actual experience is in principle impossible. It is well known that for Kant the possibility of philosophy as a science in its own right turns essentially on the question whether synthetic judgments a priori are possible. His answer in the first Critique is that these judgments cannot be justified with reference to actual experience but must instead be justified with reference to possible experience (ibid.). To give a well-known example: only under the condition that all events (changes in time) have a cause is experience (in the sense of determinate empirical knowledge) possible.

However, the categorical imperative is not a theoretical synthetic judgment a priori but rather a synthetic practical judgment a priori (GMS, AA 04: 420). In practical judgments, we determine the subject term not as it is but as it ought to be. Hence a justification with reference to possible experience is ruled out. Moreover, the subject term of a practical judgment is the agent herself rather than an object given to us from elsewhere; practical judgment in its core case is self-determination. The subject term of the categorical imperative is the human agent. Human agents by nature do not necessarily will what is practically universal. But in being guided by the categorical imperative, we demand precisely this; we oblige ourselves only to will in such a way that the principles of our action are practically universal. In judging morally, we demand of ourselves that we always act from universalizable maxims, even though this universalizability is not already contained in our actual volition (cf. GMS, AA 04: 420; 454). In other words, the categorical imperative connects the concept of a finite rational being by means of the copula ‘ought’ with an idea which is beyond any possible experience: the idea of moral perfection, which is agency from universalizable maxims.

Kant’s universalization procedure has been the subject of much criticism. The goal of a deduction in Groundwork III is not to justify this procedure. Kant does not address content scepticism with regard to the moral principle.11 Instead, he takes it for granted that the first two sections of the Groundwork have established the universalization procedure as a proper articulation of pure practical cognition

11 Korsgaard 1986, 5.
(moral reasoning). The deduction takes it for granted that the categorical imperative in its different formulations is the correct analysis of the unconditional practical necessity which, according to Kant, is expressed in our practical judgments. What the deduction seeks to justify, instead, is the synthesis in those practical judgments. By tracing moral obligation back to its origin, it seeks to show why finite rational beings like us are obliged to act under the kind of moral law Kant has already provided an account of. The goal of the deduction is, in other words, to explain why the categorical imperative has binding force for us – or, as Kant puts it, “on what grounds (woher) the moral law is binding” (GMS, AA 04: 450).

II The Deduction of Moral Obligation

The logical form of the categorical imperative is not the point of contention between the *Groundwork* and the second *Critique*. In both writings, Kant holds that the categorical imperative is a synthetic practical judgment a priori (GMS, AA 04: 420; KpV, AA 05: 31). The disagreement lies in the justification of this judgment. The first two sections of the *Groundwork* only show that those who hold “morality to be something and not a chimerical idea without any truth must also admit the principle of morality [i. e. the moral law] brought forward”. Conversely, those who hold morality to be a mere “chimerical idea” do not need to accept that the moral law has any binding force for us as finite rational beings (GMS, AA 04: 445). “Well-disposed souls” might gladly concede its bindingness (GMS, AA 04: 453), but what about those who put the “absolute necessity” (duty) of morality into doubt (GMS, AA 04: 445)? In the third chapter of the *Groundwork*, Kant attempts to establish the “third thing”, which functions as a justificatory ground for the synthetcity of our moral judgments. By means of this “third thing”, Kant wants to gain the entitlement to connect a finite will with the idea of a will that necessarily wills what is universalizable.

12 Carol Voeller’s claim that the goal of *Groundwork* III is to show how freedom is possible in a world governed by natural law misconstrues the problem *Groundwork* III is supposed to solve from the outset (Voeller 2001, 104). Voeller’s question is dealt with by Kant in a footnote in subsection III. Here, he exempts himself from dealing with this “theoretical” issue by leaving it “unsettled”. This is precisely what the solution of the third antinomy in the first *Critique* achieved; freedom is a problematic concept. In *Groundwork* III, Kant presupposes this theoretically problematic concept and only makes use of the concept of freedom in a practical sense. “We can thus liberate ourselves from the burden that weighs upon theory” (GMS, AA 04: 448).
So far, we have been mainly concerned with the task of the deduction in *Groundwork* III. We now need to turn to the actual argument Kant delivers in *Groundwork* III. There is considerable disagreement about each move in his argument. But in order to be in a position to critically adjudicate between Kant’s deduction in *Groundwork* III and his later account in the second *Critique*, we need to get a clear grasp of what his argument in *Groundwork* III consists in. In this section, I would like to lay out, in four argumentative steps, Kant’s deduction as I understand it.

1 The Reciprocity Thesis

In a first step, Kant argues for his so-called reciprocity thesis: “(A) free will and a will under moral laws are one and the same” (GMS, AA 04: 447). This thesis has long been taken to imply that an act of willing is free if and only if it is morally good. I will briefly outline a reading that can avoid this absurd consequence.

We may read the reciprocity thesis the following way: a free capacity of volition is identical with a capacity to act under the representation of moral laws. I take this to imply that our free capacity of volition is not merely a two-way capacity that can be exercised either in morally good or in evil action. Kant defines freedom positively as “the capacity of pure reason to be by itself practical” (MS, AA 06: 214). This definition makes clear that our capacity of volition is directed at the

---

13 Kant claims literally that “free will” and “a will under moral laws” are identical. Yet their identity is not complete, since it concerns only a few of their properties. More precisely, a biconditional relation holds between them; a will is free if and only if it is “under” moral laws. So both of the following claims are correct: “If a will is free, then it is under moral laws”, and “if the will is under moral law, then it is free”. Now, if the relation between “free will” and “will under the moral law” were merely analytical, it would be possible to read the relation merely in one direction (e.g. from the analytic judgment “All bachelors are men”, it does not follow that “All men are bachelors”). It is not false to speak of “free will” and “will under moral laws” as analytical (Schönecker 1999, 2013). However, one ought to add that the analytic containment holds in both directions. This is what the term “reciprocity” is supposed to capture (Allison 2012, 294 f.).


15 Henry Allison claims that Kant equates negative freedom with “transcendental freedom” or “absolute spontaneity”. Positive freedom, by contrast, is autonomy (Allison 2011, 288). This way of drawing the distinction is misleading. To put the point briefly: Kant distinguishes between the speculative idea of freedom and practical freedom. The speculative idea is the idea of a first causality. It can be defined negatively and positively. Negatively defined, it is the independence of a necessitating cause. Positively, it is the beginning of “a series of occurrences entirely from itself” (KrV, B 562). Both definitions are merely nominal definitions. Practical freedom is the freedom of a capacity of volition. It can also be defined positively and negatively. Negatively, it is the “inde-
good. We only exercise this capacity properly if we act morally well. Morally evil action is a misuse of this capacity. But again, this does not imply that when we misuse this capacity we do not act freely.\(^\text{16}\)

If the reciprocity thesis holds, it seems a promising strategy to first argue for the claim that we as finite rational agents are indeed free and then deduce from there the possibility of acting under the representation of the moral law. This seems to be precisely the strategy employed by Kant. What is less clear, however, as we will see in the second and third steps, is how he can argue for freedom in a non-question-begging way, for Kant attempts to give a “deduction of the concept of freedom from pure practical reason” (GMS, AA 04: 447, my emphasis).

\section{Deduction of the Freedom of Pure Rational Beings in General}

Since the binding force of the moral law is supposed to get a grip on the agent independently of any given desires she happens to have, Kant first points out that the kind of freedom \textit{moral} agency requires is not a relative kind that all living beings with a faculty of choice (\textit{Willkür}) possess. Instead, the freedom presupposed by the unconditional necessity of the moral law is the absolute kind of freedom that rational beings possess only in virtue of their (pure) practical rationality.\(^\text{17}\) It is

\begin{itemize}
\item \textit{pendence of the power of choice from necessitation by impulses of sensibility”}; positively, it is the capacity to act from pure practical reason (MS, AA 06: 214). Only this last definition is a real definition. It expresses a cognition to which we have access through the moral law. The central question for Kant’s ethics is whether our practical freedom is indeed transcendental freedom or whether it is merely the “comparative” or relative kind (KpV, AA 05: 96). If we equate transcendental freedom with negative freedom, we fail to see that autonomy itself is, from the theoretical perspective, both the independence of a necessitating cause and the beginning of “a series of occurrences entirely from itself”.


\textit{17} Kant does not speak explicitly of pure practical reason in subsection two. However, in subsection one he says that he will attempt to give a “deduction of the concept of freedom from pure practical reason” (GMS, AA 04: 447, my emphasis). Moreover, in subsection two Kant claims that one “cannot possibly think of a reason that would self-consciously receive guidance from any other quarter with regard to its judgments”. This claim is more plausible if one assumes that Kant speaks here of a reason that has the capacity to act from pure practical cognition. Accordingly, I take Kant’s claim to be that consciousness of a capacity of pure practical volition necessarily implies that we act under the idea of freedom. For the “idea” of freedom is not merely a representation but a concept of pure reason. Freedom as a concept of pure reason is transcendental (incompatibilist) freedom. Only the capacity to act from pure reason implies that we can act from principles that don’t presuppose any given desire. I take it that this explains “why Kant never gives compatibilism a chance [here]” (Ameriks 2003, 174). In other words, instrumental-
important to note that in *Groundwork* III Kant merely articulates the conception a rational agent has of herself. He thinks he can fend off worries of determinism by implicitly referring back to the solution of the third antinomy, which left it theoretically “unsettled” whether we are free and only demonstrated its logical possibility (GMS, AA 04: 448). If the absolute notion of freedom can indeed be thought without contradiction and the articulation of our rational agency does provide us with a reason to assert the concept of freedom, then, he thinks, we are entitled to claim the reality of that concept.\(^{18}\)

His argument for the assumption of the idea of freedom, as I read it, runs as follows. It is incompatible with the concept of pure reason to “attribute the determination of [a] judgment not to [... ] reason but to an impulse” (GMS, AA 04: 448). According to Kant’s anti-naturalist conviction, ‘pure reason’ implies that it is a spontaneous and not a receptive capacity that cognizes the form of our cognition. His anti-instrumentalist view commits him to the further claim that reason as a capacity of cognition has two fields of exercise: theory and practice. As pure practical reason, it cognizes the end of rational volition. Since reason is the source of those principles of action, a rational agent has to think of herself also as a being with a free capacity of volition. On the other hand, if we presuppose the freedom of a capacity of volition, then, as Kant showed in the first step of the argument, “morality together with its principle follows from it by mere analysis of its concept” (GMS, AA 04: 447). So, in short: Kant moves from the concept of a capacity of pure practical reason to the idea of freedom under which this capacity

\(^{18}\) Cf. Bojanowski 2006, Part II.
acts, and from there to its autonomy. The main conclusion of this argument is that all rational beings necessarily act under the idea of freedom. This marks the end of the first deduction. As we will see in the next step, the conclusion of this deduction will function as the major premise of a second deduction: the deduction of the categorical imperative.

3 Deduction of the Freedom of Human Agents in Particular

So far, the argument has only been concerned with rational beings in general. Kant has only shown that we may “think” of a being with a capacity of pure rational volition as a free being (GMS, AA 04: 448). It seems that Kant’s argument thus far has been circular. The circle consists in the claim that, on the one hand, we (a) “consider ourselves as free [...] in order to think ourselves under moral laws, [...] but afterwards [...]” (b) we think ourselves as subject to these laws because we have ascribed to ourselves freedom of will” (GMS, AA 04: 450).

This circle is not a circlus in probando but rather a petitio principii. “Well-disposed souls might gladly concede” the bindingness of the moral law, and they...
therefore simply presuppose the idea of freedom without giving an argument for it (GMS, AA 04: 453), but someone who is not well disposed could easily question this presupposition. So in order to provide a non-question-begging argument, Kant does not simply want to assume the bindingness of the categorical imperative. Instead, he wants to explain its origin or source, so that we can “see on what grounds (woher) the moral law is binding” (GMS, AA 04: 450). To proceed with his syllogism, he needs to find a reason to assume that human beings are in fact free, for so far it has only been shown that we “think” of rational beings in general as free, but it could not be “proven […] as something real in us” (GMS, AA 04: 448). Perhaps I’m trivializing the argument, but I think Kant’s next move is simply to argue for the claim that human beings are in fact rational beings:

Now the human being actually finds in himself a faculty through which he distinguishes himself from all other things, and even from himself insofar as he is affected by objects, and this is reason. [...] Reason, under the name of the ideas, shows such a pure spontaneity that it thereby goes far beyond everything that sensibility can provide it, and proves its most excellent occupation by distinguishing the world of the senses and the world of the understanding from one another [...].

With this minor premise in hand, we can finish off the syllogism by concluding that human beings necessarily act under the idea of freedom. And if the reciprocity thesis is correct, we can also infer that they have the capacity to act from the representation of (moral) laws. So the solution to the circle is to give an argument for why we do have a reason to consider ourselves free which does not presuppose the validity of the moral law, and this reason is our self-consciousness of our pure rational activity, which governs not only our theoretical knowledge but also our rational volition. However, since this argument presupposes our capacity of pure practical reason, and since we only get to know this capacity through its moral use, Kant has not really given us a non-question-begging account. Thus I take it that Kant’s solution to the circle fails, and I think that this is precisely the point where the second Critique can be seen as an improvement on his argument. There, he provides a non-question-begging account of the moral law as a fact.

25 “Nun findet der Mensch in sich wirklich ein Vermögen, dadurch er sich von allen andern Dingen, ja von sich selbst, so fern er durch Gegenstände afficirt wird, unterscheidet, und das ist die Vernunft. [...] Vernunft unter dem Namen der Ideen eine so reine Spontaneität zeigt, daß sie dadurch weit über alles, was ihr Sinnlichkeit nur liefern kann, hinausgeht und ihr vornehmstes Geschäfte darin beweiset, Sinnenwelt und Verstandeswelt von einander zu unterscheiden [...].” GMS, AA 04: 452. I want to point out again that Kant’s argument here does not need to be read as a move from theoretical to practical freedom. He rather starts with the concept of reason in general, which has two employments: theoretical cognition and practical cognition.
of pure practical cognition, from which he can then derive our consciousness of freedom. At this point in the *Groundwork*, the question still stands whether Kant has only assumed freedom because of the “importance” of the moral law (GMS, AA 04: 450).

### 4 The Normativity Claim

What the argument has shown thus far is that we consider ourselves part of the “world of the understanding” because we have the capacity to act from principles that have their source in pure reason. But since we are not only rational beings, we may consider our actions from two different standpoints:

> When we think of ourselves as free we transfer ourselves into the world of understanding as members of it and cognize autonomy of the will along with its consequence, morality; but if we think of ourselves as put under obligation we regard ourselves as belonging to the world of sense and yet at the same time to the world of the understanding.  

Kant’s argument has given us reason to think that we are entitled to consider ourselves “members of the world of the understanding”. And, on the other hand, we consider ourselves part of the “sensible world” because we are also always affected by “desires and inclinations” (GMS, AA 04: 453). This by itself does not seem to be sufficient to prove the bindingness or the “obligation” of the categorical imperative. What Kant still hasn’t yet explained is how the obligation arises – or, to put it differently, why we don’t think that it is up to our discretion to follow the moral law, and why we are instead obliged to act from universalizable maxims. It has been claimed that Kant’s solution to the bindingness problem relies on the assumption that the “world of the understanding” is axiologically “superior” to the sensible world. Only on the basis of this assumption is Kant entitled to the claim that it is not merely permissible to follow either our inclinations or the rules of reason. However, the superiority claim, so the objection goes, is unfounded and highly speculative. Indeed, Kant does not give us a worked-out argument for the superiority claim. Here is Kant’s argument:

---

26 “Denn jetzt sehen wir, daß, wenn wir uns als frei denken, so versetzen wir uns als Glieder in die Verstandeswelt und erkennen die Autonomie des Willens sammt ihrer Folge, der Moralität; denken wir uns aber als verpflichtet, so betrachten wir uns als zur Sinnenwelt und doch zugleich zur Verstandeswelt gehörig.” GMS, AA 04: 453.

27 Schönecker; Wood 2007, 205 ff.
Because the world of understanding contains the ground of the world of sense, and hence also of its laws, and thus is immediately legislating with regard to my will (which belongs wholly to the world of understanding), and thus must also be thought as such, I shall cognize myself as an intelligence, though on the other side as a being belonging to the world of sense, as still subject to the law of the former – i.e. the world of reason, which contains its law in the idea of freedom – and thus to the autonomy of the will; and must consequently view the laws of the world of understanding as imperatives for me, and actions that conform with this principle as duties.28

The claim that needs to be elucidated is the grounding relation between the “world of the understanding” on the one hand and the “sensible world” on the other. I think what Kant is saying here is that what it means to consider the world from a rational perspective is to consider reason to be normative. It just does not make sense to say that we have a decisive reason to do \( x \) and at the same time to question the normative force of rationality. Through rational deliberation (and on Kant’s view a priori cognition of the moral law), we determine the reality of the sensible world and are in this sense its “ground”. We do not simply act from given desires in accordance with the laws of nature; rather, we have the capacity to act from the representation of principles of pure reason. This capacity puts us in a position to shape the sensible world in a way that coheres with the principles of reason. So, qua rational beings, we consider ourselves free and not necessitated by natural causes. As such, we are not merely part of the sensible world; we can act from the representation of laws. This capacity makes us part of the “world of the understanding” in which agents operate according to principles of pure practical reason. So, as soon as we consider ourselves absolutely (not merely relatively) free agents, we “cognize our autonomy”. For we then come to realize that our purely formal idea of a universal law is the decisive reason for our action. In other words, we can adopt maxims in which the realization of our desires is dependent upon their practical universality. The “world of the understanding” is in this sense normative for our desires. Thus only because we are members of the world of the understanding are we entitled to connect the subject and the predicate term in our synthetic practical judgment a priori. Only because we are

28 “Weil aber die Verstandeswelt den Grund der Sinnenwelt, mithin auch der Gesetze derselben enthält, also in Ansehung meines Willens (der ganz zur Verstandeswelt gehört) unmittelbar gesetzgebend ist und also auch als solche gedacht werden muß, so werde ich mich als Intelligenz, obgleich andererseits wie ein zur Sinnenwelt gehöriges Wesen, dennoch dem Gesetz der ersteren, d.i. der Vernunft, die in der Idee der Freiheit das Gesetz derselben enthält, und also der Autonomie des Willens unterworfen erkennen, folglich die Gesetze der Verstandeswelt für mich als Imperativen und die diesem Princip gemäße Handlungen als Pflichten ansehen müssen.” GMS, AA 04: 453f.
members of the world of the understanding are we entitled to connect our imperfect will by means of the copula ‘ought’ with the idea of a perfectly rational will.

It is important to note that the rational perspective is our own perspective. We do not need to be reasoned into it. Kant’s argument does not give an externalist justification of morality. He does not have an argument against someone who takes up a standpoint deliberately beyond good and evil (morality). It would be hopeless to look for an argument here that could convince a moral sceptic of this sort that she ought to adopt a moral standpoint.29 Instead, Kant argues that, from the standpoint of a rational agent, the obligation – the ‘ought’ – is in fact a willing, and we consider reason to be the lawgiver (GMS, AA 04: 455). He does not argue against the sceptical worry about why we ought to take up the rational perspective. But since it is hard to see how this worry can be coherently spelled out – for how can the sceptic make sense of ‘ought’ without already taking up the very perspective he tries to call into question? – Kant may rightly leave this worry to the side. He instead argues that as finite rational beings we already judge our actions from this rational perspective. The rational perspective is not a perspective we might choose as we choose between different perspectives on a painting, because it is constitutive of our volition as such. Finite rational volition as such is guided by the standards of rationality. We do not need to take this claim to imply that we always do what we judge to be practically good or necessary. To be sure, we can indeed violate the moral law, and hence the moral law takes on its imperatival form. However, even the violation of the moral law is always accompanied by our recognition of its general bindingness.

Let us briefly take stock and review the core of Kant’s deduction in the *Groundwork*:

1. A free capacity of volition is identical with a capacity to act under the representation of unconditional moral laws.
2. An agent with a capacity of pure practical reason necessarily acts under the idea of freedom.
3. Humans are agents with a capacity of pure practical reason.
4. Therefore, human agents act under the idea of freedom (first deduction: deduction of freedom).
5. Therefore, human agents act under the representation of unconditional moral laws.
6. Reason is normative for our inclinations (“the world of the understanding is the ground of the sensible world”).

Therefore, the moral law is an imperative (second deduction: deduction of the categorical imperative, i.e. unconditional moral obligation).

All of the premises can be called into question. The goal of this paper is not to call these premises into question, however, but to adjudicate between two different justificatory strategies on the basis of Kant’s general metaethical framework. Before we turn to Kant’s argument in the second *Critique*, we should keep in mind how he determines the scope of his deduction in his “Concluding Remarks”. He points out that he does not deliver a positive proof of the highest principle of his deduction: freedom (GMS, AA 04: 463). Instead, he only gives a “deduction of the concept of freedom from the concept of pure practical reason” (GMS, AA 04: 447). Kant merely articulates the standpoint of a finite rational agent. Theoretical cognition of the freedom of our capacity of volition is in principle impossible. “It is therefore”, as Kant writes, “no censure of our deduction of the supreme principle of morality, but a reproach that must be brought against human reason in general, that it cannot make comprehensible [begreifen] as regards its absolute necessity and unconditional practical law (such as the categorical imperative must be)” (GMS, AA 04: 463). To comprehend the bindingness of the moral law, we would have to be in a position to cognize the condition of the causality of freedom. However, transcendental or libertarian freedom as original causality is what is itself unconditioned and therefore impossible to cognize. Hence, it remains impossible to cognize the condition of the moral law and, accordingly, to comprehend its practical necessity. This insight, which I call the incomprehensibility claim, is at the bottom of Kant’s final remark in the *Groundwork*: “And thus we do not indeed comprehend the practical unconditional necessity of the moral imperative, but we nevertheless comprehend its incomprehensibility” (GMS, AA 04: 463).

If we follow Kant on this point, we can distinguish between a strong deduction of the categorical imperative and a weak one. The strong deduction would be a deduction that justifies the bindingness by “comprehending” the practical necessity of the moral law. Kant explicitly rules out the epistemic possibility of such a strong deduction. What he provides instead is what we may call a weak deduction. The weak deduction can only “discover the concept that is compatible with [the absolute necessity of the moral law]” (GMS, AA 04: 463). This concept is the problematic concept of freedom. We have a reason to affirm this concept when our action is guided by pure practical reason. Since pure practical reason is only constitutive of moral principles, in the *Groundwork* Kant already implicitly takes moral cognition to be the “ratio cognoscendi” of freedom. By acting under the idea of freedom, we consider ourselves to be affected but not necessitated by natural causes; we are, as Kant puts it, “members of the world of the understand-
ing”. Our membership in the world of the understanding is the “third thing”,
which entitles us to connect our finite and sensibly influenced will with the idea
of a perfect rational will. Since reason is normative for our sensible inclinations,
we take this will to be normative for us.

With his claim that the moral law is a “fact of pure reason”, Kant calls into
question the need for the deduction just given. In order to be in a position finally
to show why, in the second Critique, Kant realized that his practical epistemology
was incompatible with the project of a deduction of the moral law, in the follow-
ing section I will lay out the alternative approach offered in the second Critique
and point to main argumentative differences.

III The Primacy of Moral Cognition

In the second Critique, Kant no longer makes any attempt to explain the possibil-
ity of the categorical imperative or, what amounts to the same thing, to give any
sort of a deduction (strong or weak). Kant now calls the moral law a “fact of pure
reason” (’KpV, AA 05: 31, 47), and he claims that “something different and quite
paradoxical takes the place of this vainly sought deduction of the moral prin-
ciple, namely that the moral principle itself serves as the principle of the deduction
of an inscrutable faculty [...], namely the faculty of freedom” (KpV, AA 05: 47).
Many have claimed that the main difference between the Groundwork and the
second Critique is Kant’s reversal in the order of the justification of freedom.30 He
no longer attempts, as he did in the Groundwork, to deduce the moral law from
some prior cognition of freedom. Instead, the moral law is now that which is prior
in cognition and is itself the ratio cognoscendi, the ground of cognition, of the
freedom of our capacity of volition (KpV, AA 05: 04). I think that this suggestion
goes in the right direction; in claiming that the moral law is a fact of pure reason,
Kant is concerned with the order of justification. However, since Kant already
attempted in the Groundwork to give a “deduction of the concept of freedom from
pure practical reason” (GMS, AA 04: 447), and since practical reason is nothing
but the capacity for moral cognition, we need to further specify the reversal claim.
I would like to focus on a passage where Kant explicitly explains why he calls the
moral law a ‘fact’ of reason:

[The moral law is 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 judgment that is not based on any intuition, either pure or empirical, although it would be analytic if the freedom of the will were presupposed; but for this a positive concept, an intellectual intuition would be required, which certainly cannot be assumed here.  

This passage is very dense. In the following, I want to lay out five different claims that I take Kant to be making here. Bringing them into focus will help us to see more clearly what distinguishes Kant’s argument in the second Critique from his argument in the Groundwork.

1 The Facticity Claim

Kant puts forward a facticity claim with regard to moral consciousness. The reason why consciousness of this law is called a fact of reason here is that it cannot be “derived from some given data which is prior to it”. “Data” in this context does not mean something given to us in intuition; the fact of reason is “not based on any intuition, either pure or empirical”. Instead, Kant explicitly says that the only possible candidates from which our consciousness of the moral law could be derived are “data of reason” (my emphasis). Otherwise, the categorical imperative’s claim to universality would be unfounded (cf. KpV, AA 05, §§ 2–4). It has been noted that the phrase ‘fact of pure reason’ ought to be understood not as a factum brutum but rather in the literal sense of ‘factum’ as an activity or act of reason, which is the source of this law. This reading is certainly at least partly correct, for the moral law is a fact not of intuition but of “reason”. Since reason is a spontaneous capacity, the activity reading must be correct. And if the cognition of the moral law is not the cognition of something given to us from elsewhere,
reason must cognize itself – or, to put it slightly differently, the fact of pure reason must be a “self-consciousness of pure practical reason”.

However, this is not the full story. The phrase ‘fact of pure reason’ remains ambiguous. The fact of pure reason is not only an activity of reason; as Kant also explicitly holds, the fact of the moral law means that it is “given” to us in our consciousness (KpV, AA 05: 31, 47, 91). I take this to mean that “as soon as we adopt maxims” we are “immediately conscious” of the normative requirements of the moral law (KpV, AA 05: 29). If we were only to aim at the satisfaction of our desires and did not “lay out maxims” for our actions, the moral law would not be given to us in our consciousness. Only because we make the attempt to act according to principles do we recognize the authority of the moral law, which requires us to act on principles that can be willed by all rational beings. In acting on maxims, we come to recognize that some of those maxims are formed in such a way that they cannot be willed by all rational beings and therefore lack practical universality. Those maxims are deficient because they cannot function as proper principles in a practical syllogism. In other words, in rational willing we become self-consciously aware of the proper form of those principles that guide the activity of rational willing or agency. The main point we need to keep in mind is that the claim the categorical imperative makes on us is embedded within the structure of rational agency as such. Or, to put it slightly differently, the claim the categorical imperative makes on us is constitutive of such agency. The moral law is a “fact of pure reason” in the sense that our self-consciousness of the normative requirements of rational volition is given in our consciousness through the activity of reason. And since we are thereby also aware that our inclinations are not already in line with what is practically universal, we are also conscious of the bindingness of the moral law. Kant makes this clear when he says that the moral law “forces itself upon us” and when he speaks of the syntheticity of this consciousness. I will come back to this syntheticity claim below.

35 Let me once again point out that this is not the place to address the difficulties faced by the application of the so called “universalizability test” (Wood 1999, 97–110). Nor do I want to discuss the question whether Kant’s theory of action, which, among other things, involves the claim that actions are conclusions of practical syllogisms, is plausible (Bittner 2001, chap. 3.). The leading question remains whether Kant’s claim that the moral law is a fact of pure reason is correct within his normative and metaethical framework.
2 The Epistemological Primacy Claim

The facticity claim implies the epistemological primacy of the consciousness of moral obligation over our consciousness of freedom or any other kind of intuition (rational or sensible). Kant claims that consciousness of the moral law is also the “ratio cognoscendi” of absolute freedom (KpV, AA 05: 04). This primacy claim is not global but is instead limited to a primacy in consciousness or cognition. The epistemological primacy claim does not commit Kant to an ontological primacy claim. On the contrary, Kant explicitly points out that freedom is still, as it is in the Groundwork, the “ratio essendi” – the essence – of the moral law. Moreover, Kant identifies as prior not the moral law as such, but rather “consciousness” of the moral law. Thus, Kant’s claim need not be taken to imply that there are moral facts out there in the world, independent of human cognizers.36

3 The Reversal Claim

The primacy claim has been taken to imply a reversal claim with regard to Kant’s position in the Groundwork. Kant’s argument in §§ 5 and 6 of the second Critique is similar to his argument in subsection one of Groundwork III. In both places, he argues that the freedom of our capacity for rational volition analytically entails that this capacity is directed at the moral law. It is conceivable that consciousness of the moral law could be derived from the transcendental freedom of this capacity, for, as Kant argues in §§ 2–4 and in Groundwork II, a capacity of rational volition that is transcendentally free must be guided by a purely formal principle of volition, and the moral law turns out to be just this principle. But in the fact of reason passage quoted above, Kant again, for epistemic reasons, excludes the option of employing freedom as a ground for the deduction of the moral law. Instead, what is prior in consciousness is cognition of the moral law itself.

It is a common view that the order of this inference turns the argument of the Groundwork upside down. However, in the Groundwork, Kant doesn’t only deny the possibility of theoretical cognition of freedom. He also denies the possibility of consciousness of freedom that is prior to the acts of “pure practical reason”.36

---

36 This is one reason why I don’t think that Kant’s claim that the moral law is a fact of pure reason necessarily implies that Kant is a moral realist (cf. Ameriks 2003, ch. 11; Schönecker 2013). If the moral law is “self-consciousness of pure practical reason”, as Kant puts it elsewhere (KpV, AA 05: 29), it seems more appropriate to think of Kant’s moral epistemology as an idealist position. Unlike theoretical laws, the reality of a practical law depends on the awareness of those creatures that are subject to it (cf. Bojanowski 2012).
The deduction of freedom he delivers is explicitly a “deduction of the concept of freedom from pure practical reason” (GMS, AA 04: 447). So even in the *Groundwork*, the starting point of Kant’s deduction is pure practical reason. Our entitlement to ascribe pure practical reason to ourselves is grounded in our consciousness that we can determine our volition from concepts of pure reason (“ideas”) (GMS, AA 04: 452). In other words, our consciousness of these ideas of pure practical reason is the reason for taking ourselves to be transcendentally free. I don’t think that the reversal in the second *Critique* concerns the relation of our mere consciousness of pure practical reason to the idea of freedom. What we should say instead is that Kant argued in the *Groundwork* from our mere consciousness of pure practical reason to the idea of freedom, and from there (with some additional premises) to the bindingness of the moral law. In the second *Critique*, by contrast, Kant no longer thinks that the question how a categorical imperative is possible needs to be addressed, for we already know through a “deed” that moral obligation is actual (KpV, AA 05: 03). He starts with the actuality of our cognition of moral obligation and derives our consciousness of freedom from it. In doing so, transcendental freedom, which in the *Groundwork* remains a regulative idea, has “objective reality” through our cognition of the categorical imperative, and it can now be counted among those things of which we have “knowledge” (KpV, AA 05: 05; cf. KU, AA 05: § 91).

4 The Syntheticity Claim

The passage reaffirms the *syntheticity claim* made in the *Groundwork*. Kant makes clear that even though he no longer thinks that the categorical imperative needs a deduction, its logical form remains the same. In other words, the reason why the categorical imperative does not need a deduction is not that it turns out to be analytic. Kant still considers it a “synthetic judgment a priori”. But he no longer thinks that the syntheticity claim entails that the categorical imperative is in need of deduction. In the first *Critique*, Kant holds that for a synthetic judgment a priori, “if not a proof then at least a deduction of the legitimacy of its assertion must unfailingly be supplied”, otherwise “our understanding would […] be open to every delusion” (KrV, B 286). I take it that this claim must be limited in scope to *theoretical* judgments. In other words, synthetic-practical judgments a

---

37 I fully agree with Ludwig that this is the decisive change in Kant’s argument (Ludwig 2008, 460).
priori don’t require a proof or a deduction.\textsuperscript{38} As we will see in the next section, the reason for this limitation in scope is that Kant ultimately holds that a critique of \textit{pure} practical reason is unnecessary.

5 The Incomprehensibility Claim

The syntheticsity claim leads Kant, finally, to the \textit{incomprehensibility claim}, which is also in line with his position in the \textit{Groundwork}. In deliberating about the counterfactual possibility that the categorical imperative is analytic, he reveals that he still holds that if we had an intellectual intuition, and if we could cognize transcendental freedom, we could also comprehend the unconditional necessity (bindingness) of the categorical imperative. Now if this is still true, Kant’s approach in the \textit{Groundwork} of deriving the bindingness of the moral law from the concept of freedom is not completely misguided. Comprehending the transcendental concept of freedom would put us in a position to provide what I call a \textit{strong} deduction. As we saw in the previous section, Kant only attempts to provide a weak deduction in the \textit{Groundwork}. This weak deduction, I take it, does not presuppose any intellectual intuition, because it presupposes a merely problematic concept of transcendental freedom. Some have taken this to imply that Kant did not give up on a deduction of the moral law after all. The \textit{Groundwork} and the second \textit{Critique} are compatible, they argue, because in both works Kant holds on to the incomprehensibility claim, and he rejects only the strong (and not the weak) deduction. So, when Kant in the second \textit{Critique} speaks of the “in vain sought deduction”, he is referring not to his own attempt in the \textit{Groundwork}, but to the previous paragraph, where he speaks of the strong deduction.\textsuperscript{39} Clearly, Kant’s argument in the second \textit{Critique} at AA 05, 46 f. is directed against the possibility of a strong deduction rather than the weak deduction he delivers in the \textit{Groundwork}. However, this does not mean that he still holds on to the project of a deduction as such. For, as I have argued before, in the second \textit{Critique} he rejects

\textsuperscript{38} Wolff, in his truly insightful analysis of the fact of reason, rightly points out that the moral law is “the only fact of pure reason” (KpV, AA 05: 31, my emphasis) and that there are not any valid theoretical facts of reason (Wolff 2009, 535). However, he still holds on to the claim that even the categorical imperative as a synthetic judgment a priori is in need of a deduction. Wolff thinks that it would be “unacceptable” for Kant to “silently abandon his methodological maxim” (Wolff 2009, 541; similarly Prauss 1983, 66 ff.). I think the special status of the moral law as the only fact of reason (which Wolff acknowledges) already indicates that a deduction is not required.

\textsuperscript{39} Cf. Wolff 2009, 541.
as misguided the whole question of *Groundwork* III concerning the possibility of the categorical imperative.

It should be obvious that the first three claims – the facticity claim, the epistemological primacy claim and the reversal claim – do set the second *Critique* apart from the *Groundwork*. However, they do not render Kant’s approach in *Groundwork* III a regression to dogmatic rationalism. It is not that Kant accuses himself of overstepping the bounds of sense by deducing the bindingness of the moral law through the cognition of an absolute notion of freedom in the intelligible world. The incomprehensibility claim and the syntheticity claim, which both indicate the limitations of human cognition, were explicitly already at work in *Groundwork* III (as we saw in the previous section). The concept of freedom from which Kant attempts to derive the bindingness of the categorical imperative is merely problematic. The difference between the *Groundwork* and the second *Critique* is that in the latter Kant rejects the idea of giving a deduction of the moral law altogether. For, as we shall see in the final section, Kant came to understand more clearly why the special nature of pure practical cognition does not demand a critique of its pure employment, and therefore why a deduction of its fundamental law is unnecessary.

### IV Exposition Instead of Deduction

Thus far, we have restricted our attention to the meaning of Kant’s phrase ‘fact of pure reason’. But for many, it remains an open question whether Kant is entitled to the claim that the categorical imperative has any binding force. Kant tries to address this sceptical worry in *Groundwork* III. To put the worry slightly differently, even if it is the case that human beings are guided in action by the moral law, this only answers the *quid facti* question; the *quid juris* question is still open. Hence, critics demand a deduction, and they presume that Kant’s initial attempt in the *Groundwork* was a move in the right direction.⁴⁰ Some claim that Kant never gave up on his weak deduction in the *Groundwork*.⁴¹ Others maintain that when Kant came to realize that his deduction in the *Groundwork* had failed and he found himself unable to improve his argument, he falsely took the mere exposition of the moral law to be a sufficient justification of it.⁴² According to a widespread

---

⁴¹ Beck 1960; McCarthy 1982; Wolff 2009 540 f.  
⁴² Unfortunately, we cannot solve the question by simply pointing to the heading of the section where Kant deals with the deduction issue, which reads as follows: “Of the Deduction of the
view, consciousness of the moral law could simply be a historically contingent fact. What right does Kant have to claim that the “reality of the moral law cannot be shown by means of a deduction” because it is “firmly established of itself” (KpV, AA 05: 47, cf. 94)? Or, as he also says, that consciousness of the moral law cannot be “reason[ed] out” (herausvernünfteln) from “antecedent data of reason”, but rather “forces itself on us” (KpV, AA 05: 31)? What is consciousness of the moral law like, such that it allows for non-perceptual knowledge (as a fact of pure reason rather than sensibility)? Or, to put the question in an interpretive context, is Kant right to replace the “deduction” of the categorical imperative provided in the *Groundwork* with a mere “exposition” of it (KpV, AA 05: 46)?

The demand for a deduction of the moral principle Kant puts forward raises at least three different kinds of worries, which should be made explicit and kept distinct. First, the formal content of the principle requires a derivation (content scepticism). Second, Kant needs to give an argument for why acting from this principle is possible (predeterminist scepticism). Finally, he needs to show how the principle has any binding force – or, to put it in more Kantian terms, how it determines moral agents (normative scepticism). Let me explain what role each of these worries plays in the deduction.

The first worry concerns the so-called “universalization procedure” of the categorical imperative. More precisely, this is the worry about whether the categorical imperative is in fact the proper articulation of the form of our moral judgments. Even though it is far from obvious that the categorical imperative in all its different formulations is indeed a proper articulation of our moral judgments, I don’t think that this is the question Kant was worried about when he raised the deduction issue in *Groundwork* III. This issue had already been addressed in the first two sections of the *Groundwork*. If Kant’s account of moral judgment is correct, he can take it for granted, in the beginning of the third section, that the categorical imperative is the form of practical judgment. Similarly, content scepticism is not at stake when Kant considers the deduction issue in the second *Critique*. He again takes it for granted that our moral judgments make a claim to strict universal validity. His argument in sections 1–8 is supposed to establish mainly two claims: first, only the formal law of practical reason can hold up to the

---

Principles of Pure Practical Reason*. It has been claimed that “in speaking ‘of’ a deduction here Kant may mean simply to call attention to a significant possibility that he now thinks we cannot realize” (Ameriks 2003, 186). However, the heading in the first *Critique* where Kant undoubtedly does deliver a deduction is phrased identically (KrV, B 116). So at least we cannot draw any inference as to whether Kant does deliver a deduction.

claim of strict universality; second, the principle of this law is autonomy, which is freedom in its absolute sense.

At this point of the argument in the *Groundwork*, Kant still thought it necessary to address a different kind of worry. Even if the categorical imperative is an adequate articulation of unconditional practical necessity, it is still an open question whether human beings are subject to its demands. We might call this worry “normative scepticism”. As we saw in section II of this paper, Kant’s strategy is to trace consciousness of the moral law’s bindingness back to consciousness of freedom. He first argues for the analytic connection between freedom and the moral law. He then tries to establish the claim that, as rational beings with a capacity of pure volition and conscious of freedom, we are entitled to view our actions from two standpoints. Since rationality is normative for our sensibility, Kant finally concludes that we consider ourselves obliged to the categorical imperative. We also saw that this argument turns out to be circular, for it presupposes the capacity of pure practical reason, which is known to us only via cognition of the moral law.

In the second *Critique*, Kant is still convinced that the final part of his response to normative scepticism is correct. He even implicitly refers to what is in all likelihood his argument in *Groundwork* III: “[I]t has been sufficiently proved elsewhere that freedom, if it is attributed to us, transfers us into an intelligible order of things” (KpV, AA 05: 42). So Kant does not abandon his idea of the two standpoints, which is supposed to explain, as we saw in section II, the imperative form of the moral law.

However, since he endorses the reversal claim in the second *Critique* (as we saw in the previous section), he has to give up the first part of his response to normative scepticism. He no longer tries to argue from consciousness of freedom to consciousness of the binding force of the moral law. Instead, he now claims that “[t]he most ordinary attention to oneself confirms that this idea [i.e. the moral law] is really, as it were, the pattern for the determination of our will” (KpV, AA 05: 44). He now simply refers to our self-consciousness of practical universality, and he illustrates this consciousness with the very same examples he used when he introduced the universal law formula in *Groundwork* II. In other words, the starting point of Kant’s argument in the second *Critique* is the actuality of consciousness of the moral law. He infers from this actuality that it must be possible and concludes that the problem of the possibility of the moral law (the problem which the deduction is supposed to answer) does not need to be addressed.

Now at this point, one might wonder whether there is room for a fourth kind of scepticism. This time, the scepticism would be psychological, and it would proceed as follows: it might be the case that many people are in fact conscious of the moral law when they determine their will; however, this consciousness could just be a mere historical accident of the particular time we or Kant happen
Kant would respond to psychological scepticism by pointing out the formal character of the moral law and its a priori origin. The requirements of objective and subjective universality are not historically contingent, but are rather constitutive of cognition as such. Without those requirements, practical as well as theoretical cognition would be impossible. Assuming that Kant can demonstrate how the categorical imperative determines the practical universality of any of the maxims put forward by the sceptic, and therefore how it functions as the condition of the possibility of practical cognition, psychological scepticism is unfounded. If the demand for a deduction arises because Kant’s claim about the fact of pure reason has been taken to give rise to psychological scepticism, then Kant’s analysis of practical cognition in sections 1–8 of the second Critique can be seen as a deduction. The “exposition” of the categorical imperative as an a priori principle of pure practical cognition would be at the same time its “deduction”, i.e. a justification against psychological scepticism. But if we want to stick as closely as possible to Kant’s vocabulary, it would be most appropriate to call Kant’s argument in §§ 1–8 a “metaphysical exposition”. For an exposition is “the clear, though not detailed, representation of that which belongs to a conception; and [it] is metaphysical when it contains that which represents the conception as given a priori”.

The deduction in Groundwork III is not supposed to be a response to psychological scepticism. When Kant worries that the categorical imperative could be a chimerical idea without truth at the end of section II, he is worried that the categorical imperative might not apply to finite rational beings. As we will see in the final section, a fundamental difference between theoretical and practical cognition renders this deduction of the categorical imperative unnecessary.

V Dogmatism Instead of a Critique of Pure Practical Reason

I claim that the demand for a deduction of the categorical imperative only arises because Kant erroneously modelled his argument for the principles of practical cognition in the Groundwork on the deduction of principles of pure theoretical cognition. Hence he tried to provide a “critique of pure practical reason” (GMS, AA 04: 440, 446, my emphasis) – a critique that he deemed unnecessary in the second Critique (KpV, AA 05: 03). So if we want to see why Kant gave up on the

44 Marx 1845, 31 ff.; Nietzsche 1886, 188 ff.; 1887, 258–274.
deduction, we need to get a better grip on the difference between theoretical and practical cognition.

Judgments of pure theoretical reason need to be restricted to possible experience, for otherwise they are transcendent and do not have any meaning. Accordingly, it is not sufficient to show that the fundamental concepts of pure theoretical reason are thinkable; Kant also needs to show that they map on to the world as constitutive principles of possible experience. In other words, criticizing reason in its pure employment means restricting its employment to possible experience. The principles of practical cognition, by contrast, are not restricted to possible experience, for in practical cognition we do not determine “objects that may be given to reason from elsewhere”. Instead, cognition of the moral law is the self-consciousness of pure practical reason (KpV, AA 05: 29), and, in the paradigm case, this cognition makes the object of its cognition “actual” (wirklich) (KrV, B IX f.). As Kant puts it elsewhere, practical cognition is the “ground of the existence of objects” (KpV, AA 05: 46). So, when Kant claims that practical cognition is concerned with the determination of our faculty of volition (KpV, AA 05: 19), he means not merely that we judge our maxims to be good or evil, but that cognition thereby brings about the action and is thus also causally efficacious. Practical cognition, in its paradigmatic form, requires that the principle of legislation also be the principle of execution.

This is rather abstract, and it is not immediately obvious how this account of practical cognition can explain why the categorical imperative does not require a deduction. It might help to elucidate the claim with one of the examples to which Kant alludes when he claims that we become conscious of the moral law as the determining ground of our volition by simply attending to our own consciousness (see above). Let us consider the following maxim: “When I am in need of money, I will borrow money and promise to pay it back even though I know this will never happen.” Again, this is not the place to respond to content scepticism and to critically assess Kant’s universalization procedure. But Kant’s main idea is that the end of this maxim (borrowing money) could not be achieved if the maxim became a universal law. There wouldn’t be such a thing as borrowing money if it became a universal law that everyone in need of money falsely promised to pay it back. There is nothing problematic about a world without promises as such. But if an agent did have this maxim and wanted to borrow money, she could not bring about her end if it became a universal law. So, from a universal or rational standpoint (which is our own standpoint), this maxim can be neither willed nor even thought without contradiction. Now, a proper maxim is supposed to function as a principle in a practical syllogism. For this principle to be justified, it should hold for every action subsumed under it (objective universality), and all other rational beings should be willing to agree to it (subjective universality). But
if the maxim cannot even be thought without contradiction, it cannot function as a sound principle of a practical syllogism and therefore does not hold up to a maxim’s claim to universality. A maxim that cannot be universalized lacks practical necessity and is morally impermissible; it is, in Kant’s own modal terms, “contrary to duty”.

What this is supposed to show is that in practical cognition the merely “problematic [...] thought a priori of a possible general legislation” is the determining ground of our will (KpV, AA 05: 31), and therefore of the reality of some object: the good (or evil) action. Through this merely problematic thought, we determine whether a principle counts as a principle of practical cognition. This is not to say that in practical cognition we gain positive insight into the intelligible world, cognizing the true end of moral agency or receiving insight into this end by divine revelation. On the contrary, even in moral cognition it is merely a formal end of which we have cognition a priori: the merely problematic thought of general legislation. Moral cognition would remain empty if our maxims did not contain given desires or, as Kant likes to say, inclinations. For as I pointed out earlier, only when we lay out maxims (which is our way of determining under which conditions we choose to satisfy our desires) do we become conscious of the moral law. And only in having particular desires can our reasoning guide our action. Without inclination, the moral law would be ineffective. It is not as though we must repress all our desires when we act morally; we must only make their realization depend upon their possible universalizability or law-likeness. But in determining those desires, we determine them not as they are but as they ought to be, for practical cognition is volition under the condition of its practical universality. Volition is practically universal if it is valid for all actions that fall under the rule (objective universality) and if it holds for every rational cognizer (subjective universality). Since our inclinations are not already in line with what is practically universal, practical reason issues its demands in the form of imperatives. Now on Kant’s view, in order for the cognition to become practical, it needs to give rise to a feeling which is able to strike down all competing desires. Kant calls this feeling “respect for the moral law” (KpV, AA 05: 73 ff.). Only if reason is in fact able to determine our volition in such a way that the cognition is itself the ground of the object’s existence are we entitled to say that reason is practical and not merely theoretical. For even in technical-practical volition, the rule is always determined by a presupposed desire. Only in moral-practical cognition is the determining ground of our volition not some presupposed desire but a merely problematic thought: the thought of universal legislation.

One might wonder how this can explain why the moral law does not require any deduction, for in Kant, and especially in Kant, there is a fundamental similarity between theoretical and practical cognition. Even theoretical cognition is not
something that we are passively saddled with. It is only through the spontaneity of our understanding and reason that theoretical cognition is possible. And one might think that even in practical cognition the matter of our maxims does not (or at least does not entirely) depend on practical reason but is instead given to us from elsewhere. As we have just seen, however, in sharp contrast to theoretical cognition, practical cognition is not concerned with the determination of an independently existing manifold. Through the forms of our practical cognition, the categories of freedom, we do not simply determine what we in fact desire. Instead, in practical cognition we cognize the formal requirements of rational volition, and the cognition of these requirements is, in the paradigm case, also “the ground of the existence of the objects” (KpV, AA 05: 46). We can call this view Kant’s causal efficaciousness view of practical reason and distinguish it from the mere representationalist view. Reason as a causally efficacious capacity shapes the world not only formally but also materially.

These two views lead to two different procedures for justifying the fundamental principles of cognition. For although we can become conscious of the apriority and necessity of the fundamental principles in both practical and theoretical cognition by simply “attending” to our judgments (KpV, AA 05: 30), the fundamental principles of theoretical cognition are still in need of a deduction. In theoretical cognition, Kant holds a representational view of reason. He therefore needs to show that the fundamental principles are not mere “chimera” (Erdichtungen) but do in fact map onto the world. It transpires that their “objective reality” can be “quite easily and evidently proved through examples from the sciences” (KpV, AA 05: 91, 47). The a priori principles of theoretical reason are the condition of the possibility of experience, but they do not bring the object of their cognition into existence. Our theoretical cognition is therefore the cognition not of a divine intellect – an “intuitus originarius” which brings the object of its cognition into existence by cognizing it – but of an “intuitus derivativus” (KrV, B 72). In practical cognition, we are closer to the “intuitus originarius” insofar as the fundamental principle of practical reason does not simply represent given desires but is rather itself the origin of the object of our volition. This does not mean that we are merely finite as theoretical cognizers but divine as practical agents; rather, our practical finitude is constituted by the fact that we are subject to the influence of inclination.

Now since practical cognition is the origin of the object of our willing, and since willing is considered a causal force in nature, reason “proves” the reality of the categorical imperative “through a deed and all ratiocination against its possibility is in vain”. In other words, a deduction of the principles of pure reason is the attempt to answer the question how those principles are possible. If it turns out that the principles of this cognition are already actual, the question of their possibility does not even arise (KpV, AA 05: 03).
I think that what Kant came to realize in the second *Critique* was that the critical project in his practical philosophy was not, as he had envisioned it in the *Groundwork*, a critique of *pure* practical reason but was instead (simply) a critique of practical reason; pure reason in its practical use is not to be criticized, because it is by its very nature transcendent. Kant merely wants to show “that there is pure practical reason [...]. If it succeeds in this it has no need to criticize the pure faculty itself in order to see whether reason is merely making a claim in which it presumptuously oversteps itself (as does happen with speculative reason)” (KpV, AA 05: 03). I take this to mean that Kant only needs to show that there is a fundamental practical principle constitutive of pure practical cognition. The question whether and how it relates to an object need not be addressed, for this cognition first brings the object of its cognition into existence. As an a priori cognition of what ought to be, it is by its very nature transcendent. Kant calls “dogmatic” any method that does not “at first examine the capacity or incapacity of reason” (KrV, B 7). I think we are therefore entitled to the rather bold claim that, with respect to *pure* practical reason, Kant’s method in the second *Critique* is dogmatic. With this qualification in mind, Kant can again endorse a claim that Erich Adickes dates to between 1769 and 1771: “Pure reason is dogmatic in morals” (Refl, AA 17: 520, R 4362). The third section of the *Groundwork* was supposed to make a transition from metaphysics, which proceeds in dogmatic fashion (KrV, B 7), to a critique of pure practical reason. Kant abandoned this idea. With respect to *pure* reason, the *Critique* of Practical Reason remains metaphysical. The *critical* project in the second *Critique* is limited to two tasks: first, Kant needs to fend off all attempts to derive moral principles from given desires; second, he needs to defend his absolute notion of freedom against predeterminist scepticism. It turns out that the self-consciousness of pure practical reason is consciousness of an unconditional requirement of practical volition. This unconditional requirement cannot in principle depend on any given desire, but must instead have its source in reason. As soon as we adopt maxims, we become aware of this requirement. Since ‘ought’ implies ‘can’, we are entitled to infer our absolute freedom from the consciousness of this requirement. This is the only deduction of the fundamental principle of pure practical reason offered by Kant in the second *Critique* – a deduction of freedom. The only further argument needed at this point is an argument that

---

45 Some think that because the moral law functions as the ground of this deduction it is a “credential” even for its own validity (Beck 1960, 174; Engstrom 2002, xiv; Ameriks 2003, 184; Sussman 2008, 66); the moral law validates itself by validating freedom. “The credential is furnished by the fact that the moral law is itself the source of a deduction” (Engstrom 2002, xiv). I fail to see how this could do any justificatory work in favour of the moral law. I agree with the view that the
will fend off determinist sceptics. This scepticism is theoretical in origin, since it questions the possibility of absolute freedom in the world of appearances. Against this kind of scepticism, Kant attempts to prove not the theoretical actuality of the concept of transcendental freedom but rather its logical possibility. Whenever this scepticism arises, he reminds his readers of his solution to the third antinomy (GMS, AA 04: 455 f.; KpV, AA 05: 30, 48 ff., 94–106; KU, AA 05: 175; RGV, AA 06: 39 f., 49 f.). However, if the self-consciousness of pure practical reason as the moral law is a cognition, then not only can we ascribe the merely problematic idea of freedom to a finite rational being, but we can even positively determine this idea through our non-perceptual moral cognition, thereby proving – from a practical point of view – the objective reality of freedom. How is consciousness of freedom possible? Through consciousness of the moral law. The famous example of the gallows is supposed to confirm the order of this deduction.

Conclusion

In this paper, I have tried to develop a radical incompatibilist story about the relationship between the *Groundwork* and the second *Critique*. This incompatibilism is radical because it concerns the roots of Kant's approach to the justification of the bindingness of the principles of practical cognition. I don't think, as standard incompatibilists maintain, that Kant rejected his argument in *Groundwork* III because he detected a mistaken inference from theoretical freedom to practical freedom, or because he became aware of the fact that his axiological “superiority claim” concerning the priority of the noumenal over the sensible world was unjustified. My analysis of *Groundwork* III has shown that those particular arguments can be spelled out such that they are perfectly compatible with his view in the second *Critique*. The radical incompatibilism I want to endorse instead concerns the false belief that practical cognition, like theoretical cognition, requires a critique of *pure* reason. In the second *Critique*, Kant revised his argument because he came to realize that, in contrast to synthetic a priori judgments of theoretical cognition, the possibility of synthetic a priori judgments of practical cognition can be derived from the actuality of a “deed”. With respect to *pure* practical reason, the second *Critique* proceeds not critically but metaphysically, i. e. dogmatically. Hence Kant abandoned the idea of a deduction of the categori-

---

Text is corrupt here, and it is obvious from the context that “credential” here refers to freedom and not to the moral law; our practical cognition of the bindingness of the moral is a credential for our being transcendently free (cf. Bojanowski 2006, 78–81; Timmermann 2008).
cal imperative and with it the idea of a critique of pure practical reason: the very project he had started to undertake in the third section of Groundwork III. I agree with the compatibilist story that the deduction Kant considered impossible both in the Groundwork and in the second Critique is what I have here called a strong deduction – a deduction that takes as its starting point theoretical cognition of transcendental freedom. And I also agree with compatibilists that the deduction he actually delivers in Groundwork III is only a weak deduction, which makes use of a merely problematic conception of transcendental freedom. However, I don’t think that Kant’s argument in Groundwork III succeeds entirely, for Kant moves from our conception of pure practical reason to transcendental freedom, and from there to our consciousness of autonomy; hence, his argument remains question begging. As argued above, the facticity claim in the second Critique, by contrast, can provide a non-question-begging account of moral obligation. Here I agree with the optimistic incompatibilist, who views the argument in the second Critique as an improvement on his argument in the Groundwork. However, more importantly, the radical difference between his approaches in the Groundwork and the second Critique, which the standard incompatibilist account overlooks, concerns the very idea of a critique of pure practical reason, and with it the very nature of pure practical cognition. This is why his attempt at a deduction in Groundwork III and his facticity claim are fundamentally incompatible. And, as I point out in the beginning of this paper, we should for this very reason resist the generality of Kant’s claim in the first Critique to the effect that, for all synthetic judgments a priori, “if not a proof then at least a deduction of the legitimacy of its assertion must unfailingly be supplied” (KrV, B 286).

Acknowledgement: My work on this article was funded by the Netherlands Organization of Scientific Research (NWO).

Works Cited

McCarthy, Michael H.: Kant’s Rejection of the Argument of Groundwork III. In: Kant-Studien 73, 1982, 169‒188.
Prauss, Gerold: Kant über Freiheit als Autonomie. Frankfurt am Main 1983.
Schönecker, Dieter: Kant: Grundlegung III. Die Deduktion des kategorischen Imperativs. Freiburg/München 2010.