



# On the Transcendental Deduction in Kant's Groundwork III

Marília Espírito Santo  
UFRGS

*Disputatio* Vol. 4, No. 30

May 2011

DOI: 10.2478/disp-2011-0004

ISSN: 0873-626X

# On the transcendental deduction in Kant's *Groundwork* III<sup>1</sup>

Marilia Espirito Santo  
UFRGS

## Abstract

The purpose of the third section of Kant's *Groundwork* is to prove the possibility of the categorical imperative. In the end of the second section, Kant establishes that a proof like this is necessary to show that morality is 'something' and 'not a chimerical idea without any truth' or a 'phantom' (1785: 445). Since the categorical imperative was established as a synthetic *a priori* practical proposition, in order to prove its possibility it is necessary 'to go beyond cognition of objects to a critique of the subject, that is, of pure practical reason' (1785: 440). Kant names this kind of proof a *deduction*. The present paper intends to (1) show the argument whose purpose is to justify the categorical imperative; (2) show that the argument is a transcendental deduction; (3) present the argument as it is reconstructed by Allison, and (4) show that, although it seems compelling, the position of the commentator could not be accepted by Kant himself.

## Keywords

Kant, *Groundwork*, transcendental deduction, moral law, categorical imperative.

## Introduction

The notion of a deduction plays a central role in Kant's critical project. Nowadays, it points to a meaning quite familiar to us: it refers to the logical procedure by means of which a conclusion is established

<sup>1</sup> The first version of this paper I wrote with a scholarship from CAPES/Brazil during a research period in New York under supervision of Beàtrice Longuenesse, to whom I'm thankful for her extremely helpful criticisms. The final version I wrote with a scholarship from CNPQ/Brazil. Special thanks are owed to Luciano Codato.

through the relationship between some premises. Kant was familiar with this logical usage of the notion 'deduction', but as Henrich (1989: 31) remarks, it was neither the only, nor the most common usage in the academic language in the 18<sup>th</sup> century.

During the 18<sup>th</sup> century, 'deduction' was a notion used by jurists to refer to the written claims exposed to the Court in legal proceedings. Considering the argumentative structure of a deduction, one of its peculiar characteristics is that it must refer to an origin. Since the aim of a juridical deduction was to justify the legitimacy of a possession or a usage, that is, the legitimacy concerning an acquired right, it was necessary to explain how this possession or usage came into being. With this, it could be possible to decide who between both parties in the controversial juridical claim was right. The origin of an acquired right should be found in a fact, which must exist before the right in question came into being. That is why the argumentation presented in a deduction should relate the origin to fundamental facts that constitute it. Kant used the term *deduction* having in mind the deduction writings and not the logical procedure (see Kant 1781: A XII; A84/B116; A751/B779; A752/B780).

Kant called *metaphysical deduction* the task of referring to a non empirical origin, or to identify this origin before justifying the legitimacy of a possession or a usage. Once the *a priori* origin is identified it is possible to go on to the task of justification, which Kant called *transcendental deduction*. What he was trying to justify by a deduction was the possibility of synthetic *a priori* judgments. And this is the main ground for calling his project the critical philosophy. According to him, when we come to *a priori* judgments we have to consider their origin in the nature of reason itself and so to justify them and explain their possibility. These tasks of justification and explanation belong to a critique of reason by itself.

In relation to the practical use of reason, Kant is trying to prove that we, human beings, can act morally and, consequently, judge our actions. His starting point is the common rational cognition. He agrees with everyman that moral judgments are bivalent, that is, that we do can say that some actions are right and others wrong. On the other hand, he disagrees with positivists that moral judgments can be verified or justified by appealing to experience. He internalizes the origin of the moral law and defends that human beings *ought* to act morally well because the moral law is a self-imposed one. So, he identifies the origin of the moral law in reason, and this means that

the moral law is an *a priori* one. But since from the mere analysis of the concept of a human being does not follow the concept of acting morally well, or obeying the moral law, the moral law (for human beings) is said to be a synthetic *a priori* principle. And, as we saw above, it has to be justified by a transcendental deduction.

In other words, willing the good action is not necessarily contained in the volition of a human being, endowed with reason and sensibility, that is, it cannot be analytically derived from the volition of such a being. In 1785: 420n, Kant justifies that the categorical imperative is a synthetically practical proposition *a priori*, because it 'does not derive the volition of an action analytically from another volition already presupposed (for we have no such perfect will), but it connects it immediately with the concept of the will of a rational being as something that is not contained in it'.

Before we move on and analyze the deduction itself, it is important to bear in mind that the content and the origin of the moral law is something Kant has expounded in the first two sections of the *Groundwork*. In the first section, from an analysis of common rational cognition, he arrives at the condition for a moral action, which is obedience to a law. In the second section, he presents the content of the moral law by means of some formulations of it and identifies its origin in reason. The formulations are: (1) the formula of universal law, 'act only in accordance with that maxim through which you can at the same time will that it become a universal law' (1785: 421); (2) the formula of universal law of nature, 'act as if the maxim of your action were to become by your will a universal law of nature' (1785: 421); (3) the formula of humanity, 'so act that you use humanity, whether in your person or in the person of any other, always at the same time as an end, never merely as a means' (1785: 429); (4) the formula of autonomy, presented, initially, not as a command, but as 'the idea of the will of every rational being as a will giving universal law' (1785: 431), and (5) the formula of the kingdom of ends, 'that all maxims from one's own lawgiving are to harmonize with a possible kingdom of ends as with a kingdom of nature' (1785: 436). The formulas of universal law, humanity and autonomy are the primary formulations. The formulas of universal law of nature and of the kingdom of ends derived, respectively, from the formulas of universal law and of autonomy, can be called analogical formulas as Almeida suggests (2002).

The usual interpretation of the formula of universal law as *the* categorical imperative may lead one to believe that the formula of autonomy adds nothing to the other two primary formulations, that is, to those of universal law and of humanity. However, Kant does not just claim that the principle of autonomy follows from them, but also that the two primary formulations 'were only *assumed* to be categorical because we had to make such an assumption if we wanted to explain the concept of duty' (1785: 431). But those are not the only reason for us to take the formula of autonomy as the one that best expresses the unconditional duty. The principle of autonomy actually brings two ideas: one is the interest in acting on the categorical imperative; the other is the idea of conceiving not only oneself but all rational agents as universal legislators despite their own particular empirical ends. The idea of the moral agent not merely as acting in accordance with 'that maxim that at the same time can become a universal law', but as conceiving herself as a universal legislator and thus as the source of these maxims characterizes the interest in acting on the categorical imperative. Moreover, it states the rational nature not just as an objective end but also as an end that can motivate us. Besides this, by its analogical formula, the principle of autonomy brings the idea of not just oneself but all other rational agents as universal legislators, that is, as agents who despite their particular empirical ends have the same overriding interest in the universal law. It is only with the principle of autonomy that it is possible to understand how a manifold of agents with different empirical interests can accept a universal law. This idea, as Guyer reminds us (1998: 236), is necessary to prove that the categorical imperative is not just intentionally noncontradictory and coherent but also extensionally realizable, a requirement to demonstrate its real possibility.

The formula of universal law and its analogical universal law of nature express just the *form* of the principle of morality, which consists of universality. The isolated reading of this formula accuses Kant's moral theory of empty formalism. The formula of humanity, on the other hand, expresses just the *matter* of the principle of morality, which consists of the rational being, as an end by its nature and hence as an end by itself, as the limiting condition of all merely relative and arbitrary ends. But even with a matter to fulfill the form, it was still missing an interest in acting on the categorical imperative instead of an empirical interest in acting on hypothetical imperatives. Such an interest is introduced just with the formula of autonomy, together

with the understanding of a manifold of agents with diverse empirical particular ends but with the same overriding interest in a universal law, because it is only with the idea of autonomy that we have the complete determination of all moral maxims. This explains why, although Kant claims that the three primary formulas are just expressions of the very same law, they actually complement each other and that we really need the idea of autonomy to have access to the moral law.

Autonomy is the principle behind moral judgments and aims to be the condition for moral action. It expresses the essence of moral law, and it is the principle on which a rational agent would act if reason had full control over passion. That is, although Kant does not make it clear, the principle of autonomy does not necessarily need to take the form of a categorical imperative. Now, once we know which is the principle of morality, the next step is to ask about how it can be justified.

Kant first tries to justify it as a moral law and then as a categorical imperative. But since the autonomy expresses the essence of the moral law, which is the principle on which a rational agent would act if reason had full control over passion, and since it appears for a human being, who sometimes can act under the influence of passion, as a categorical imperative, that is, as a principle on which she *ought* to act, the question about the justification of the principle of autonomy can be expressed as 'how is a categorical imperative possible?'. And this is precisely the question Kant asks in the headline of subsection 4 of the third section of the *Groundwork*, and to which he will answer with a transcendental deduction.

## I

In the third section of the *Groundwork*, Kant's argument is given in subsection 4, under the title 'How is a categorical imperative possible?' as follows:

A rational being counts himself, as intelligence, as belonging to the world of understanding, and only as an **efficient cause** belonging to this does he call his causality a *will*. On the other side he is also conscious of himself as a part of the world of sense, in which his actions are found as mere appearances of that causality; but their possibility from that causality of which we are not cognizant cannot be seen; instead, those ac-

tions as belonging to the world of sense must be regarded as determined by other appearances, namely desires and inclinations. All my actions as only a member of the world of understanding would therefore conform perfectly with the principle of autonomy of the pure will; as only a part of the world of sense they would have to be taken to conform wholly to the natural law of desires and inclinations, hence the heteronomy of nature. (The former would rest on the supreme principle of morality, the latter on that of happiness). But **because** the world of understanding contains the ground of the world of sense and so too of its laws, and therefore immediately lawgiving with respect to my will (which belongs wholly to the world of understanding) and must accordingly also be thought as such, **it follows** that I shall cognize myself as intelligence, though on the other side as a being belonging to the world of sense, as nevertheless subject to the law of the world of understanding, that is, of reason, which contains in the idea of freedom the law of the world of understanding, and thus cognize myself as subject to the autonomy of the will; **consequently** the laws of the world of understanding must be regarded as imperatives for me, and actions in conformity with these as duties. (1785: 453-4)

## II

The core of the proof begins with the adversative conjunction ‘but because...’ and goes to the end of the paragraph. To analyze the argument, we can break it down in two main parts: the first one is stated by the element ‘because’ and provides a reason; the second is stated by the expression ‘it follows’ and provides a conclusion. The premises of the argument can be rewritten as follows:

P1 – Because the world of understanding contains the ground of the world of sense;

Corollary of P1 – and so too of its laws [and because the world of understanding contains the ground of the laws of the world of sense];

P2 – and [because] is therefore immediately lawgiving with respect to my will (which belongs wholly to the world of understanding) and must accordingly also be thought as such;

Its conclusion, that can also be broken down in two parts, as follows:

C – It follows that I shall cognize myself as intelligence, though on the other side as a being belonging to the world of sense, as nevertheless subject to the law of the world of understanding, that is, of reason,

which contains in the Idea of freedom the law of the world of understanding, and thus cognize myself as subject to the autonomy of the will; C – consequently the laws of the world of understanding must be regarded as imperatives for me, and actions in conformity with these as duties.

The first premise is not difficult to understand. It states just that the world of understanding contains the ground of the sensible world. The meaning of the corollary that follows from P1 is also compelling. It is necessary, however, to pay attention to the meaning of *ground* in this passage, once Kant uses it with different meanings depending on the context. Sometimes *ground* is used as a synonym of *ratio*, reason, as well as cause, but it can still be found as a synonym of principle. In the context of the theoretical use of reason, specifically in the domain of his thesis about what it is a representation and how knowledge by representation is possible, Kant introduces a distinction between ground and cause (see Zingano 1989: 85). On the other hand, in the *Reflections on Metaphysics*, he presents *ground* as a *first cause* (see Kant 1769: R3972).

In the same manner, when he talks about practical grounds he claims that they are 'grounds of reason [that] provide the rule for actions universally, from principles, without influence from the circumstances of time and place' (1783: §53).

It is not at all our purpose here to list all the occurrences of the term *ground* and how it is used. The ambiguity of the term is solved in the context of its application, but it is necessary to be careful to avoid misinterpretations. In the text under analysis, Kant claims that 'the world of understanding contains (*enthält*) the ground of the world of sense'. *Ground*, here, points to the *efficient cause*, the first cause, the starting point. 'A rational being counts himself, as intelligence, as belonging to the world of understanding, and only as an *efficient cause* belonging to this does he call his causality a *will*'. The world of understanding *contains* the rational being as an efficient cause. That is, the rational being, through her reason, can bring about changes in the world of sense because it is the primary origin of movement. The rational being can be an efficient cause as long as she can be considered from a double standpoint: member of the world of understanding and part of the world of sense. Since she is also part of the world of sense, her property as an efficient cause could not be realized, that is why the necessity of a command, of an imperative.



The term *ground* can only be understood as *efficient cause* in the passage under analysis because Kant claims that the ‘world of understanding *contains* the ground of the world of sense’. If his claim were that the world of understanding *is* the ground of the world of sense, ground should be understood as *ratio*, reason and not cause. If the world of understanding were the *cause* of the world of sense, we would be admitting a transcendent use of the principle of causality. A use beyond the *phenomena* that would transgress the limits that Kant himself is trying to establish. So, if we took *ground* as *cause* we would be attributing to Kant dogmatism. But, since the world of understanding contains the rational being as an efficient cause, a rational being who is also part of the world of sense, we preserve Kant within the limits of criticism.

The second premise is a little bit more problematic. One might think that because of P1 and its corollary, that is, because of the relation between the worlds of understanding and of sense established in P1, the world of understanding is also immediately legislative for my will, which belongs entirely to the world of understanding. However, this cannot be true.

To better understand P2 we can rewrite it in the following way:

P2 – In relation to my will, which belongs wholly to the world of understanding, the world of understanding is directly legislative, and it must also be conceived as containing the ground of actions and laws of the world of sense.

Actually, we can break P2 down in two premises. The sentence which appears between parentheses should be read as an independent premise. So, we have P2 and P3 as follows:

P2 – In relation to my will, the world of understanding is directly legislative, and it must also be conceived as containing the ground of actions and laws of the world of sense.

P3 – my will belongs wholly to the world of understanding.

The meaning of the P2 itself is not problematic, what is problematic is the relation between it and P1; better, the problem is relative to the element ‘because’ in the beginning of P1 and the function of the expression ‘and therefore’, which introduces P2. That is, one might think that P2 is a conclusion that follows from P1. However, this would be a complete misinterpretation of what Kant is arguing for. It

would make no sense to support that *because* the world of understanding contains the ground of the world of sense and of its laws, it is also directly legislative for the will. It would make no sense because the will belongs entirely to the world of understanding (P3); it does not belong to the world of sense. So, there is no such type of relation between P1 and P2. The simplest way to solve this misinterpretation is to put a 'because', or any other element that indicates a reason, in the beginning of P2. So, it would be properly read as a premise in addition to P1.

Moreover, it is important to stress, that the 'it', in the second part of P2, refers to the world of understanding. Hence, what P2 expresses is that (1) the world of understanding is directly legislative to the will and (2) in relation to my will, the world of understanding 'must also be conceived as containing the ground of actions and laws of the world of sense'. Thus, from P2 and P3, it is possible to say that the will, as part of the world of understanding, contains the ground of the world of sense and its actions and laws. And this is the gist of the deduction: that is, that the pure will as part of the world of understanding contains the moral law as a categorical imperative for this will affected by desires and inclinations, as part of the world of sense.

The conclusion of the argument is a little bit easier to understand, although it is not completely evident. To analyze it, it is possible to break it down in two parts, which are separated by the occurrence of the element 'consequently'.

[...] **it follows** that I shall cognize myself as intelligence, though on the other side as a being belonging to the world of sense, as nevertheless subject to the law of the world of understanding, that is, of reason, which contains in the idea of freedom the law of the world of understanding, and thus cognize myself as subject to the autonomy of the will; [...]

In this first part, Kant maintains that the human being (*I*), inasmuch as she considers herself as intelligence and *at the same time* as a being that belongs to the world of sense, is subject to the law of the world of understanding and to the autonomy of the will. It is important to bear in mind the conjunction *and at the same time* since the beginning of the

sentence, although it does not appear in this passage<sup>2</sup>. It is not the being only as intelligence that is subject to the law of the world of understanding and to the autonomy of the will; it is the being that is both: intelligence and sensible, that is subject to reason (which contains the law of the world of understanding) and to the autonomy of the will. It would be a misunderstanding to take Kant to be supporting that a rational being *only* as intelligence is subject to the law of the world of understanding and to the autonomy of the will. If it were this, it would be impossible to explain how a human being is subject to the moral law and can take it as a motive to her actions, that is, it would be impossible to prove morality under human conditions. Moreover, the element ‘nevertheless’ would not be necessary. For a being who is only intelligence, it is not necessary to consider herself in an adversative way subject to the law of the world of understanding and to the autonomy of the will.

In relation to the second part of the conclusion, ‘**consequently** the laws of the world of understanding must be regarded as imperatives for me, and actions in conformity with these as duties’, it is possible to say that what is a law of the world of understanding must appear, or be considered by a human being (*me*), as an imperative to such a being. This is because she is not only intelligence, she is also part of the world of sense, so her will can also be affected and hence motivated to act by a law of this last world. And since an imperative is just ‘the formula of a command of reason and is expressed by an ‘ought’ (Kant 1785: 413), the actions according to it are called duties.

Thus, almost without realizing, the reader is faced with a complete deduction<sup>3</sup>. This deduction is the answer Kant provides to the

<sup>2</sup> In a number of passages throughout the *Groundwork*, Kant emphasizes the simultaneity of both perspectives (intelligible and sensible) in relation to human beings and the imperative character of the moral law. This simultaneity plays a fundamental role here.

<sup>3</sup> For Liddell, the paragraph we analyzed and took to be the whole deduction is just the second part of Kant’s deduction in the *Groundwork*. The author affirms that the deduction begins in the subsection ‘Freedom must be presupposed as a property of the will of all rational beings’, and its second part is presented in the subsection ‘How is a categorical imperative possible?’ (1972: 401-2). We cannot agree with Liddell. For us, the argument presented in the subsection ‘Freedom must be presupposed (...)’ is important for the deduction, but it is just a preparatory argument together with the subsection ‘Of the interest attaching to the ideas of

question 'How is a categorical imperative possible?'. The gist of the deduction is that the pure will as part of the world of understanding contains the moral law as a categorical imperative for this will as part of the world of sense. Behind the deduction there is the idea that somehow, the rational perspective of a human being is superior (in the sense of being an efficient cause) to her sensible perspective, and that the pure practical will contains the supreme condition of the will affected by sensible desires.

In the paragraph following the one we analyzed, Kant again answers the question 'how is a categorical imperative possible?' (see 1785: 454), and this answer can be taken as a summary of the deduction just presented, one of the formal characteristics of a good deduction as Henrich (1989: 34) points out. Next, Kant claims, with a concluding remark that, 'the practical use of common human reason confirms the correctness (*Richtigkeit*) of this deduction' (1785: 454)<sup>4</sup>.

Now we can conclude that (1) if the purpose of a transcendental deduction for Kant is to justify the legitimacy of a possession or a usage of a synthetic *a priori* judgment or principle; (2) if the autonomy, under human conditions, was proved to be a categorical imperative and hence a synthetic *a priori* principle; (3) if membership in the world of understanding is what justifies that the human being, a finite rational being, is autonomous and hence can act morally well, and (4) if the paragraph we analyzed is the argument that proves that the human being is a member of the world of understanding, therefore this paragraph presents a complete transcendental deduction.

morality'. This seems to be also Kant's idea, since he affirms, at the end of the first subsection 'The concept of freedom is the key to the explanation of the autonomy of the will', that he cannot yet answer the question of how a categorical imperative is possible, because 'some further preparation is required'. So subsections 2 and 3 are the preparatory argument to the answer that will be given in subsection 4, where Kant presents the deduction itself.

<sup>4</sup> In the second *Critique*, Kant again confirms the success of the deduction of the *Groundwork*: 'It [the *Critique of Practical Reason*] presupposes, indeed, the *Groundwork*, but only insofar as this constitutes preliminary acquaintance with the principle of duty and provides and **justifies** a determinate formula of it; (...)' (1788: 8).

## III

It is important to note that our interpretation is distinguished from interpretations of well-known Kant scholars. Our present purpose is to analyze Allison's thought in relation to the deduction of the third section of the *Groundwork* and show why we think Kant himself would not concur<sup>5</sup>.

According to Allison, the deduction, whose pivotal point is the move from possession of reason to membership in the intelligible world, can be presented in seven steps:

- (1) 'Now I assert that every being who cannot act except under the idea of freedom is by this alone – from a practical point of view – really free' (Kant 1785: 448).
- (2) 'And I maintain that to every rational being possessed of a will we must also lend the idea of freedom as the only one under which we can act' (Kant 1785: 448).
- (3) All laws 'inseparably bound up with freedom' are valid for every being with reason and will.
- (4) But the Reciprocity Thesis establishes that the moral law is 'inseparably bound up with freedom'.
- (5) Therefore, the moral law is valid for every being with reason and will.
- (6) Since beings such as ourselves have reason and will, the moral law is valid for us.
- (7) Since we do not necessarily follow the dictates of the law (these dictates being 'objectively necessary' but 'subjectively contingent'), the law for us takes the form of a categorical imperative, that is, we are rationally constrained, although not causally necessitated, to obey it.

Steps 1 and 2 consist of a preparatory argument, as the commentator calls it. Step 7 consists of 'a distinct deduction of the categorical

<sup>5</sup> Allison's work on Kant's moral theory has been criticized for some time now. That is why it is important to point out that, as we understand them and as far as it goes with our reading, the criticism raised here differs completely from those raised by Stephen Engstrom, Andrews Reath, Karl Ameriks and Paul Guyer to whom Allison replies in his 'Kant on freedom: a reply to my critics', 1996.

imperative' (Allison 1995:224). Steps 3 to 6<sup>6</sup> consist, therefore, of a deduction of the moral law. And here it is the first distinction between our and Allison's analysis.

Allison notes that the third section of the *Groundwork* is one of the most enigmatic of Kantian texts. Although it is clear that its main purpose is to justify the supreme principle of morality, articulated in the first two sections, and for that Kant appeals to a deduction; it is not clear whether the deduction is of the moral law, the categorical imperative, freedom, all three, or even whether it can be properly characterized as a deduction at all. Allison's argument is for a deduction of the moral law, and his underlying presupposition is the reciprocity thesis.

As already indicated, our first disagreement with this interpretation is about what Kant is trying to justify by a deduction. We argue for a deduction of the categorical imperative, Allison argues for a deduction of the moral law.

In section two of the *Groundwork*, Kant claims that the principle of autonomy is a categorical imperative 'cannot be proved by mere analysis (...), because it is a synthetic proposition' (1785: 440). For such a proof 'one would have to go beyond cognition of objects to a critique of the subject, that is, of pure practical reason' (1785: 440), a business that 'does not belong in the present section' (1785: 440) says Kant. Moreover, in the end of the same section, Kant emphasizes that the proof that morality is not a chimera to human beings 'follows if the categorical imperative (...) is true and absolutely necessary as an *a priori* principle' and this 'requires a possible *synthetic use of pure practical reason*, which use, however, we cannot venture upon without prefacing it by a *critique* of this rational faculty itself, the main features of which we have to present, sufficiently for our purpose, in the last section' (1785: 445).

In the first part of section three, Kant affirms that 'if (...) freedom of the will is presupposed, **morality** together with its principle **follows** from it **by mere analysis** of its concept. But the **principle of morality (...) is nevertheless always a synthetic proposition (...)**' (bold added) (1785: 447). Then he goes on to state that free will must be attributed to every rational being (1785: 448). It

<sup>6</sup> According to Allison, steps 1 and 2 constitute the explicit argument Kant provides, steps 3 to 7 constitute the natural extension of the argument he does not make.

would be contradictory to suppose a rational being who could be regarded as not free; that is, if we deny freedom, we necessarily deny reason. So, if we have to attribute free will to every rational being, if from freedom of the will follows morality by mere analysis, we can say that from the concept of a rational being follows morality by mere analysis. We must note that this is true when we think about a pure rational being. To a finite rational being, the principle of morality is 'always a synthetic proposition', because by analysis of its concept does not follow to act morally well. Given its finitude, for such a being it is possible to act different from what morality dictates. That is why to complete the deduction of the categorical imperative Kant has to appeal to a distinction between two worlds (world of understanding and world of sense) as two standpoints from which imperfectly rational beings may regard themselves. Moreover, in the same section, right after the argument that justifies the possibility of the categorical imperative, Kant claims that 'the practical use of common human reason confirms the correctness of this deduction' (1785: 454).

Now, according to the passages we quoted it seems reasonable to regard Kant's attempt to prove the possibility of the categorical imperative, and not of the moral law, by a deduction. Moreover, since a deduction is to prove the transcendental conditions of a possession or a usage of a synthetic *a priori* judgment or principle, and since the moral law follows analytically from the concept of reason, a deduction would not be necessary to justify it. A deduction of the moral law, that is, a deduction of an analytic principle does not have the character of justification, but of demonstration, and once more this is not the task Kant intends to develop here.

But, even if Allison's suggestion was right that the deduction is of the moral law<sup>7</sup> and not of the categorical imperative, he argues for a

<sup>7</sup> What Allison may have in mind (he does not make it clear or explicit) when he affirms that the deduction is of the moral law is that what Kant is trying to justify by a deduction is why **this** one (autonomy, 'the idea of the will of every rational being as a will giving universal law' (Kant 1785: 431)) and not another is **the** moral law. The answer we can provide to him is that **this is the moral law** because of the *complete determination*. It is only because of autonomy, whose principle according to Kant (1785:431) follows from the conjunction of the principles of the universal law and of the humanity, that a rational agent can conceive herself as capable of renunciation of all interest in volition from duty, which is the specific mark of a moral

failure of the deduction due to a fatal ambiguity in two central notions. The first is in that of the intelligible world and the second in that of the will.

In relation to the ambiguity in the notion of an intelligible world, Allison asserts that Kant refers to both a *Verstandeswelt* (world of understanding) and an *intelligibele Welt* (intelligible world) and shifts from the former to the latter without sufficient justification. In doing this, Kant cannot avoid providing a 'non-question-begging deduction of the moral law in *Groundwork III*' (Allison 1995:228).

The *Verstandeswelt* is to be understood negatively as encompassing whatever is nonsensible or 'merely intelligible', that is, whatever is thought to be exempt from the conditions of sensibility (the *noumenon* in the negative sense). The *intelligibele Welt*, on the other hand, is to be understood positively as referring to a supersensible realm governed by moral laws, a 'kingdom of ends' or 'the totality of rational beings as things in themselves' (Kant 1785: 458) (the *noumenon* in the positive sense).

Allison notes that Kant's goal is to show that human beings are members of an *intelligibele Welt* because this would entail that they stand under the moral law. The problem is that the possession of reason only gets us to a *Verstandeswelt*, and since this world is an indeterminate concept, it cannot provide any conclusion about the nature of the rational being as a whole nor about her will.

The second difficulty, related to an ambiguity of the notion of the will, is a corollary of the former. The main point is that given the identification of will and practical reason, the claim that rational beings possess a will can mean (1) merely that reason is practical or (2) that pure reason is practical. The former (practical freedom) is sufficient for us to affirm that we are genuine rational agents rather than automata; but it is the second (transcendental freedom) that is necessary to establish our autonomy.

The problem, again, is that the membership in the *Verstandeswelt* provides support just for practical freedom, but it is transcendental freedom that is necessary and sufficient to establish morality on the basis of a nonmoral premise about our rationality.

will, and understand how herself and all others rational agents in spite of their particular empirical ends have an overriding interest in a universal law.



Thus, Allison concludes for a failure of the deduction, which, according to him, Kant himself may have recognized. Assuming that, the commentator claims that

we can see why he <Kant> would abandon the attempt to establish the practicability of pure reason on the basis of any nonmoral premise. Thus, instead of beginning with the concept of a rational agent and moving from this first to the presupposition of freedom and then, via the Reciprocity Thesis, to the moral law, Kant there <*Critique of Practical Reason*> moves directly from the consciousness of the moral law as the ‘fact of reason’ to the practicability of pure reason and the reality of transcendental freedom (1995: 228). (Emphasis mine)

#### IV

Our critique of Allison is based on three points. First, knowing the way in which Kant uses the terms, it is more reasonable to consider him to be using the terms of a *Verstandeswelt* and of an *intelligibele Welt* not in a univocal sense, but, sometimes, interchangeably. Second, that in the *Groundwork*, it is not Kant’s purpose to prove the reality (objective validity), but the real possibility<sup>8</sup> of the categorical imperative, and for this the *noumenon*, in the negative sense, is necessary and sufficient. Finally, that which gets us to the intelligible world it is not just the possession of reason but also the *consciousness of the spontaneity of reason*.

The *noumenon* in the negative sense is a being of understanding ‘in-sofar as it is not an object of our sensible intuition’ (Kan 1781: B307), but it can be under determination of space and time, that is, it is the

<sup>8</sup> See the passage of the first *Critique* where Kant talks about the work of the jurists (1781: A84/B116), that they distinguish between what is lawful (*quid juris*) and what concerns the fact (*quid facti*) and that they call the first a deduction. In the third section of the *Groundwork* Kant is working with a lawful question, the *quid fact* he will deal with only in the second *Critique*. Where, by the way, he claims that ‘the moral law is given, as it were, as a **fact of pure reason** of which we are *a priori* conscious and which is apodictically certain (...). Hence the **objective reality** of the moral law **cannot be proved by any deduction** (...)’. (1788: 47). And further he claims again that ‘the objective reality of a pure will or, what is the same thing, of a pure practical reason is given *a priori* in the moral law, as it were by a fact – for so we may call a determination of the will that is unavoidable even though it does not rest upon empirical principles’ (1788: 55).

object that *can appear* (as a *phaenomenon*). Admittedly, it is the *noumenon* in the negative sense that Kant needs to justify the real possibility of the categorical imperative because the categorical imperative is the way the moral law, a law of a being of understanding, appears to a being that is not only a being of understanding, but it is also a being of sense, a being that is under determination of space and time. To prove the possibility of the categorical imperative is to prove that the sensible affected will can give meaning, through its actions, which appear in space and time, to the rules of the pure rational will, and for that, the *noumenon* in the negative sense is necessary and sufficient. If we have just the *noumenon* in the positive sense, that is, 'the object of a non-sensible intuition (...), namely intellectual intuition' (1781: B307), we would have two different worlds and no connection between them. Hence, it would be impossible to justify how a being that is also part of the sensible world could be motivated to act by a law of the intelligible world. It is the *noumenon* in the negative sense that allows us to understand the intelligible world and the sensible world as a double standpoint of the same world, a double standpoint that the human being considers herself and allows to understand why she has to act morally well.

Finally, our last objection to Allison is that what gets us to the intelligible world it is not just the possession of reason but also the *consciousness of the spontaneity of reason*. And this seems to give positive content to our thought of ourselves as members of the intelligible world and, hence, a positive content to the concept of an intelligible world itself. That is, it seems reasonable to support that the consciousness of the spontaneity of reason presupposes a law different from that of nature, and this allows a positive characterization of the intelligible world and, therefore, a characterization of the human being as a *noumenon* in the positive sense.

Notwithstanding, Almeida (2009: 45) notes that the characterization of the human being as a *noumenon* in the positive sense, allowed by the consciousness of the spontaneity of reason, can lead us to another problem. The problem is that although such a rational being can 'transport' herself to the intelligible world by the consciousness of the spontaneity of theoretic reason this is not sufficient to ascribe the same spontaneity to practical reason. For this, an independent moral premise would be necessary. And this, according to him, is the unsolvable problem that made Kant abandon his attempt to prove the

supreme principle of morality by a deduction and appeal to a *fact of reason* in the second *Critique*.

However we will need to address this problem in future work. In the beginning of this paper, we stated as our objectives to show the argument whose purpose is to prove the real possibility of the categorical imperative; to show that the argument is a transcendental deduction; to present the argument as it is reconstructed by Allison, and, finally, to show that, although it seems compelling, the position of the commentator would not be accepted by Kant himself.

Marilia Espirito Santo  
Universidade Federal do Rio Grande do Sul (UFRGS)  
Programa de Pós-graduação em Filosofia  
Av. Bento Gonçalves, 9500 – prédio 43311, bloco Al, sala 110  
CEP 91501-970 – Porto Alegre, R.S., Brazil  
mariliae@yahoo.com.br

## References

- Allison, Henry. 1995. *Kant's theory of freedom*. Cambridge: Cambridge University Press.
- Allison, Henry. 1996. Kant on Freedom: a reply to my critics. In *Idealism and Freedom: essays on Kant's Theoretical and Practical Philosophy*. Cambridge: Cambridge University Press.
- Almeida, Guido. 2002. Sobre as "Fórmulas" do Imperativo Categórico. In DOMINGUES, Ivan., PINTO, Roberto Paulo M., DUARTE, Rodrigo. *Ética, Política e Cultura*. Belo Horizonte: ed. UFMG.
- Almeida, Guido. 2009. *Fundamentação da Metafísica dos Costumes / Immanuel Kant*. Tradução com introdução e notas por Guido Antônio de Almeida. São Paulo: Discurso Editorial: Barcarolla.
- Guyer, Paul. 1998. The Possibility of the Categorical Imperative. In *Kant's Groundwork of the Metaphysics of Morals: critical essays*. Lanham: Rowman & Littlefield Publishers.
- Henrich, Dieter. 1989. Kant's notion of deduction and the methodological background of the first Critique. In Foster, E. (ed.) *Kant's Transcendental Deductions. The Three "Critiques" and the "Opus Postumum"*. Stanford, California: Stanford University Press.
- Kant, Immanuel. 1769. R3972. In *Kant's gesammelte Schriften*. Berlin: Walter de Gruyter & Co, 1926. V 17.

- Kant, Immanuel. 1781. *Critique of Pure Reason*. Translated by Paul Guyer and Allen Wood. Cambridge: Cambridge University Press, 1998.
- Kant, Immanuel. 1783. *Prolegomena to Any Future Metaphysics*. Translated by Gary Hatfield. Cambridge: Cambridge University Press, 2007.
- Kant, Immanuel. 1785. *Groundwork for the Metaphysics of Morals*. Translated by Mary Gregor. Cambridge: Cambridge University Press, 2009.
- Kant, Immanuel. 1788. *Critique of Practical Reason*. Translated by Mary Gregor. Cambridge: Cambridge University Press, 2009.
- Liddell, Brendan. 1972. Kant's 'deduction' in the '*Grundlegung*'. In *Proceedings of Third International Kant Congress*: 401-406, ed. by Beck, Lewis White. Dordrecht-Holland: D. Reidel Publishing Company.
- Zingano, Marco Antonio. 1989. *Razão e História em Kant*. São Paulo: editora Brasiliense.