University of Sussex

A University of Sussex PhD thesis

Available online via Sussex Research Online:

http://sro.sussex.ac.uk/

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details

Habermas's Project of Social Criticism: between Normativity, Institutions and Practices

David Martínez

DPhil Social and Political Thought

University of Sussex

January 2017

Contents

Title: Habermas's project of social criticism: between normativity, institutions and practices

Summary: This thesis maintains that Jürgen Habermas's moral and political theories rely on a modified version of Kant's notion of normativity. Taking this as a starting point, it examines this component in light of criticisms inspired by Hegel's critique of Kant. The thesis shows that Habermas can answer most of the criticisms that could arise from Hegel's critique. That said, Hegel's criticism of the will as a tester of maxims does apply to Habermas. This criticism states that Kant cannot connect the universal will of morality and the particular will of the empirical subject because he rules out particular contents as susceptible of being universalized. And it can apply to Habermas because he set strict limits to what can count as a content which may bleed into the justification of moral norms and, following Kenneth Baynes – in his interpretation of Habermas's theory –, of legal and political norms. To be justifiable, – according to Habermas – these norms need to embody generalizable interests and they cannot be based on particular interests. However, Habermas infers from this that norms can only be justified with impartial, that is agent-neutral reasons, and cannot be justified with agent-relative reasons. From this, emerges the question whether and to what extent a theory of this sort can successfully include particular contents (for example a particular agents' real interests, inclinations and needs). The strict version of the generalizability of norms seems to occlude this possibility. Nonetheless, it is possible to rebut this criticism by slackening the strong version of normative justification that Habermas has built into the theory. By means of an analysis of two elements that he incorporates into his reconstruction of the normative point of view, namely, the concept of ideal role taking and the notion of mutual recognition, it is possible to argue that the loosening of the strict notion of generalizability is a modification that does not contradict and actually coheres with Habermas's Kantian concept of moral reason, and this operation fortifies the theory in the face of the Hegelian criticism of the will as a tester of maxims. To develop these issues, this work is divided in two parts with two chapters each part. Part I is an analysis of Habermas's notion of moral reason and autonomy and it reconstructs its normative Kantianism. After that, it discusses Hegelian criticisms of Habermas's moral theory. Part II focuses on Habermas's political Kantianism in Between Facts and Norms and in the debate with Rawls and it examines Hegelian criticisms of that Kantianism.

Abbreviations

Works by Jürgen Habermas

BFN   Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy, trans. William Rehg (Cambridge: Polity Press, 1996).

BNR   Between Naturalism and Religion: Philosophical Essays, trans. Ciaran Cronin (Cambridge: Polity Press, 2008).

CES   Communication and the Evolution of Society. (Boston: Beacon Press, 1979).

JA   Justification and Application: Remarks on Discourse Ethics, trans. Ciaran Cronin (Cambridge, MA: MIT Press, 1993).

KHI   Knowledge and Human Interests. (Boston: Beacon Press, 1971).

MCCA   Moral Consciousness and Communicative Action, trans. Christian Lenhardt and Shierry Weber Nicholsen (Cambridge, MA: MIT Press, 1990).

MW   "'Reasonable' versus 'True', or the Morality of Worldviews," in J. G. Finlayson and F. Freyenhagen (eds,) Habermas and Rawls. Disputing the Political. (New York: Routledge).

PDM   The Philosophical Discourse of Modernity. (Cambridge, MA: MIT Press, 1985).

PC   The Post-National Constellation: Political Essays. (Cambridge, MA: MIT Press, 2001).

PT   Postmetaphysical Thinking. (Cambridge, MA: MIT Press, 1992).

RPR   "Reconciliation through the Public Use of Reason: Remarks on Rawls's Political Liberalism," in J. G. Finlayson and F. Freyenhagen (eds,) Habermas and Rawls. Disputing the Political. (New York: Routledge).

TCA1/TCA2   The Theory of Communicative Action, Volumes 1&2, trans. Thomas McCarthy (Cambridge: Polity Press, 1984 & 1987).

TL   Law and Morality. The Tanner Lectures on Human Values. Delivered at Harvard University. October 1 and 2, 1986.

IO   The Inclusion of the Other: Studies in Political Theory, ed. Ciaran Cronin and Pablo De Greiff, trans. Ciaran Cronin (Cambridge, MA: MIT Press).

STPS   The Structural Transformation of the Public Sphere. (Cambridge: MIT Press, 1989).

T&J   Truth and Justification, trans. Barbara Fultner (Cambridge: Polity Press, 2003).

Works by John Rawls

PL   Political Liberalism (New York: Columbia University Press, 1993; paperback edition, 1996).

RH   "Political Liberalism: Reply to Habermas," in J. G. Finlayson and F. Freyenhagen (eds,) Habermas and Rawls. Disputing the Political. (New York: Routledge).

TJ   A Theory of Justice (Cambridge, MA: Harvard University Press, 1971; revised edition, 1999).

Works by Kant

Die Metaphysik der Sitten, in AA 6: 203–493 [includes Tugendlehre (Doctrine of Virtue) and Rechtslehre (Doctrine of Right)]. Metaphysics of Morals, ed. and trans. Mary J. Gregor. Cambridge, 1997.

Grundlegung zur Metaphysik der Sitten, in AA 4: 385–463. Groundwork of the Metaphysics of Morals, ed. and trans. Mary J. Gregor. Cambridge, 1996.

Kritik der praktischen Vernunft, in AA 5: 1–163. Critique of Practical Reason, ed. and trans. Mary J. Gregor. Cambridge, 1996.

Kritik der reinen Vernunft, B edition in AA 3: 1–552, A edition in AA 4: 1–252. Critique of Pure Reason, ed. and trans. Paul Guyer and Allen W. Wood. Cambridge, 1998.

Vorlesungen über Moralphilosophie. Includes "Praktische Philosophie Herder," in AA 27/1: 1–89; "Moralphilosophie Collins," in AA 27/1: 237–471; "Moral Mrongovius," in AA 27/2.2: 1395–1581. [In part in:] Lectures on Ethics, trans. Peter Heath, ed. Peter Heath and J.B. Schneewind. Cambridge, 1997.

Zum ewigen Frieden, in AA 8: 341–386. Perpetual Peace, trans. Lewis White Beck. Indianapolis and New York, 1957.

"Über den Gemeinspruch: Das mag in der Theorie richtig sein, taugt aber nicht für die Praxis," in AA 8: 273–314. "On the Common Saying: That May Be Correct in Theory, but It Is of No Use in Practice," trans. Mary J. Gregor. In Practical Philosophy, ed. and trans. Mary J. Gregor. Cambridge, 1996, 273–309.

Introduction

Many scholars working Habermas, the so called second generation Frankfurt School thinker, ask whether his critical theory is best seen as Kantian or Hegelian (or both). Among other things, some of them aim to rectify the shortcomings that they claim to have found in Habermas's theory by emphasizing either its Kantian or Hegelian components. Others criticize Habermas because he is a Kantian or because he is a Hegelian. Finally, some contribute to the understanding of which is the best possible interpretation of Habermas's project of social criticism as a Kantian or a Hegelian theory. Robert Brandom is not a philosopher who is normally associated or recognized for his contributions on these debates. At least, he is much better known for his work on other areas. However, in my view he is absolutely correct when he states that 'Habermas himself keeps a wary, careful distance from the Hegel of 1806 and after, and is far more comfortable associating himself with Kant when the "Kant oder Hegel?" question arises'. (2015: 34)

In light of this issue, in this work I claim that the normative foundations of Habermas's critical theory are Kantian, even while they have taken on board Hegel's critique of Kant. At first blush this claim might seem paradoxical. However, it makes more sense, when one knows the particular understanding of Hegel's critique of Kant that we endorse. For this purpose I rely on Robert Pippin (1991, 1997) and Terry Pinkard (1994, 1997) who propose the *Post-Kantian* interpretation of Hegel's critique of Kant, in which Hegel does not intend to eliminate Kant´s concept of moral autonomy (a view that to a certain extent is developed by Robert Brandom too). Rather, the issue is that pure reason cannot define the autonomy of the will independently of the institutions and practices of modern ethical life [*Sittlichkeit*]. I take this reading of Hegel's critique as the ground, which helps me to understand Habermas's complex, indeed somewhat fraught relationship with both Hegel and Kant.

That said, Habermas gives priority to the Kantian component in the reconstruction of the normative foundations of his critical theory. In this respect, Habermas's theory really is part of the Kantian family as Habermas himself, and various commentators have acknowledged. (MCCA, 195-197; Rehg, 1994: 2; Forst, 2011: 154; Baynes, 2016: 102; Laden, 2011: 135) And this means that a Kantian notion of moral reason takes the central stage in Habermas's moral and political theories. (Habermas, 2011: 284) This raises the question whether and to what extent Habermas can successfully rebut Hegelian criticisms inspired by Hegel's critique of Kant. At first appearance, as long as Habermas takes on board Hegel's critique, it seems that he can respond, if not to all, then to most of the Hegelian objections to his position. Habermas himself addresses this issue, insofar as it is levelled at his moral theory, in his article 'Morality and Ethical Life: Does Hegel's critique of Kant Apply to Discourse Ethics?' and of course the gist of his answer is that Hegel's critique does not apply to Discourse Ethics. (MCCA, 195-215; Finlayson, 1999: 29)[1]

Habermas has several convincing arguments to back up his claim. Among them, his notion of moral reason is a *modified* version of Kant's moral philosophy. Habermas's moral reason is based on a pragmatic notion of autonomy, whereas Kant's is grounded on a logical or metaphysical form of necessity. In Habermas, morality takes place in social space and historical time (T&J258), whereas in Kant is based on a metaphysical view which splits human agency between the *noumenal* self and the *phenomenal* self: the first, oriented only by the representation of practical principles and laws; the second, by the concrete person with her needs, inclinations and desires. (MCCA, 203; Apel 1983: 597) Kant's concept of autonomy discard the *phenomenal* self as a proper source of moral guidance, and argues that we ought to act from duty, and will maxims of action in virtue of their universal form. Certainly, in Kant we can act according to duty insofar as we are not only *phenomenal* selves, but also *noumenal* and as such fully rational and transcendentally free) selves. Habermas rules out this alternative because in his theory only concrete participants in discourse can legislate norms that can have moral worth. (MCCA, 130, 198) Roughly speaking, Kant's moral reason takes place in the *bifurcated* consciousness of the moral agent, so to speak, whereas in Habermas morality is based on intersubjective practices of mutual understanding between concrete individuals.

---

[1] According to Bohman and Rehg, 'Habermas's discourse theory of morality generally goes by the name "Discourse Ethics," a somewhat misleading label given that "ethics" has a distinct non-moral sense for him'. (2014) In this work I use the two labels because Habermas's moral theory has been largely labeled with the term 'Discourse Ethics' and this does not lead to misunderstandings

In his critique of Kant's moral philosophy, Hegel claims that the split between the *noumenal* self and the *phenomenal* self – between the *moral* will and the *empirical* will – discards the contents of the *empirical* will. (Hegel, 1991: §135) This is the criticism of the will as a tester of maxims[2] which pertains to a twofold dilemma: on the one hand, insofar as Kant's moral philosophy does not show how concrete content can be included in moral reasoning, then how can morality provide us with concrete guidance? On the other hand, due to this lack of content, how can morality explain the real motivation of empirical agents? Concerning the question of the lack of content, Kantians could answer this by reference to the formula of humanity. (4: 428-429, 435, 440) In this regard, humanity is the last source of all our moral actions, and humanity itself is a proper content: it is not a mean, or a principle, rather, it is the objective end of practical reason. And to answer the motivational question, Kantians have argued that good moral reasons seem to be enough to motivate concrete people to do what they ought to do. As Korsgaard asserts, 'the reasons why an action is right and why you do it are the same'. (1986: 10, see also Mackie 1977, 23–24)

In this work, I do not examine in detail the Kantian answers to Hegel's objection of the will as a tester of maxims. Rather, I ask whether a version of this criticism does apply after all to Habermas. The latter advances a lot when he develops the view that normative practices take place in social space and historical time. Therefore, when he *deflates* and socializes Kant's notion of moral reason he underpins his theory and this made it less vulnerable to Hegelian criticisms. Nevertheless, things get more complicated when Habermas asserts that moral norms can only rest on impartial and hence agent-neutral and not on agent-relative reasons. (IO, 7, 43) Bearing in mind this strict condition, I claim that Hegel's latter criticism applies to Habermas because he cannot give a full account of how the concrete contents of particular wills can be included in practical discourse. Hence, the aforementioned twofold dilemma returns. Still, in this work I aim to develop a defence of Habermas, which demands only that he slacken the strict distinction between agent-relative and agent-neutral reasons. He can do this because of the familiar point that some agent-relative reasons are universal.

According to Kenneth Baynes, Habermas's Kantian notion of moral reason not only pertains to his moral theory, but also has a central place in his political theory. (2016: 170,

---

[2] Of course, Hegel's criticisms of Kant as a moral philosopher are many. In this work, I do not only focus on the criticism of the will as a tester of maxims. Among other challenges, I develop the charge of the *empty formalism* of the categorical imperative as well.

179) Then, I bring a version Hegel's criticism of the will against both theories. Concerning the moral theory, I rebut the criticism by the modification already mentioned, namely, I argue that Habermas needs to loosen the strict distinction between agent-relative and agent-neutral considerations. With regard to Habermas's political theory, I propose a version of the same solution (namely, the slackening of the strict distinction between agent-relative and agent-neutral considerations in the justification of legal and political norms) to a slightly different problem. Now, it might be thought that the solution I propose is invalidated by the fact of reasonable pluralism. This is because, in the present historical context it is difficult to expect that all citizens are going to put into practice a particular notion of communicative reason, that is to say, that they will engage in discourse, and hence presuppose it pragmatically. To rebut this challenge, I end up claiming that it is reasonable to presume that citizens can act according to Habermas's notion of autonomy, and in doing so they are able to recognize themselves not only as *addressees* but also as *authors* of their legal and political community. Moreover, the *thin* and *weak* features of the concept of rational discourse in Habermas make the allegiance of the citizens to the normative core of legal and political legitimacy more tenable. This is because this practice is unavoidable in communication and it does not demand the commitment of the citizens to a particular religion, ideology or ethical doctrine.

In Part I of this text, I examine Habermas's notion of practical reason and autonomy. I show that these concepts are based on a re-working of Kant's practical philosophy. This re-working relies on a *modified* and *attenuated* version of Kant's moral reason, because it has a much weaker form of necessity than the latter. Additionally, the Kantian normative component furnishes not only Habermas's moral theory, but also a central plank of his legal and political theory. After I have developed this basic framework in Chapter One, in Chapter Two, I discuss Hegelian criticisms of Habermas's moral philosophy. At first blush, the *modified* and *attenuated* Kantianism of his theory allows Habermas to claim that he can rebut, if not all, then most of the objections that arise from Hegel's critique. However, I show that Hegel's criticism of the will as a tester of maxims can still apply to Habermas's moral theory. This criticism states that Kant cannot connect the universal will of morality and the particular will of the empirical subject because he rules out particular contents as susceptible of being universalized. In Habermas´s theory something similar happens insofar as he sets strict limits to what can count as a content which may enter into practical

discourse. I show, however, that Habermas can rebut Hegel's criticism of the will if he introduces the modification already mentioned above.

In Part II, I focus on Habermas's political Kantianism in *Between Facts and Norms* and in the debate with Rawls. In Chapter Three, I show that Habermas´s theory of constitutional democracy gives a central place to a Kantian notion of normative justification contained in the principle of Discourse (D). (BFN, 107-108) Moreover, I examine Habermas´s co-originality thesis between private and public autonomy and I claim that, as Kant does, Habermas gives normative priority to the 'system of rights' before the concrete practical work of public autonomy is carried out. (RH, 63, 76; Habermas, 2001: 766-782; Habermas, 2011: 295) In Chapter Four, I show that, as with his moral theory, Habermas incorporates Hegelian components in his legal and political theory also. (BFN, 59, 63, 129, 132-133, 421) Nonetheless, according to Baynes, the Kantian notion of autonomy and morality enjoys a certain priority in Habermas's political theory. On the one hand, the incorporation of Hegelian components strengthened Habermas's political philosophy in the face of criticisms inspired by Hegel's critique of Kant. On the other hand, if Baynes's interpretation is correct, then the priority of the Kantian component seems to leave room for the return of Hegel's critique. I examine the criticism of the will as a tester of maxims this time in Habermas's political theory. I claim that Habermas can solve this criticism if he modifies some of the components of his legal and political theory.

Part I

Habermas's Kantianism and Hegelian criticisms of Discourse Ethics

1

The Kantian Foundations of Habermas's Discourse Ethics

This Chapter sets the basic framework of this thesis in which I examine Habermas's theory in light of Hegelian criticisms inspired by Hegel's critique of Kant. This stage is supported by the claim that I want to develop, namely, that Habermas's notion of normative justification relies on a modified version of Kant's notion of pure practical reason and autonomy.[1] Modified, in the sense that Habermas's account of communicative reason is an *attenuated* version of Kant's concept of pure practical reason because, among other things, the former has a much *weaker* form of necessity than the latter. Habermas's communicative reason is based on a pragmatic and social notion of justification, whereas Kant's is grounded on a logical or metaphysical form of necessity. According to Kant we must act in accordance with duty because we are rational beings and pure reason demands that we act in certain ways, say, we must not lie. For Habermas, as mature moral beings in a post-conventional society, there is no another functional alternative to regulating our moral lives and solving conflicts of interest between agents than by discourse. In this regard, the moral ought is a rational demand: we must not lie, because we are the beings for whom it is

---

[1] In this thesis I assume – following Kant and the Kantian tradition – that an autonomous will is shaped and coincides with a normative notion of moral reason (or in Kant's terms pure practical reason). In other words, the autonomy of the will pertains to a rational being who attaches her insights and actions to the procedures and rules of normativity. In this regard, Kant asserts in the *Groundwork* that 'a free will and a will under moral laws are one and the same'. (4: 447) Also there he states 'Reason must view herself as the authoress of her principles, indepently of alien influences, and must consequently, as practical reason, or as the will of a rational being, by herself be viewed as free'. (4: 448) Henceforth, the reader should not be confused because in some parts of this work I refer to pure practical reason (or moral reason) and in others to autonomy. These notions are conceptually related in the Kantian tradition. Habermas has a similar view. In *Truth and Justification* he asserts that 'Reason become "practical reason" insofar as it determines will and action according to principles'. (T&J, 94) Explicitly Habermas claims to follow Kant in the sense that by means of the concept of practical reason, moral and political autonomy are reconstructed in Discourse Ethics and in his theory of Constitutional Democracy in BFN. (See 2011: 284 where Habermas makes the connection explicit)

second nature to act on principle, and because otherwise we would precipitate ourselves into social conflicts with other agents. Hence, the moral ought is a social demand. Nevertheless, as I will show in this Chapter, Habermas's concept still has the Kantian hallmarks (what I also call the Kantian presuppositions of Discourse Ethics), namely, the Kantian notions of autonomy, equality, universalizability, cognitivism, deontologism and as in Kant, Habermas's conception of normativity is procedural (although, in this thesis I will centre on the notion of autonomy). In order to prove this, I focus on the two principles which frame Habermas's reconstruction of normativity: the principle of discourse (D) and the principle of universalization (U). In what follows, I show that bearing in mind these principles, it is possible to examine Habermas's Kantianism. These principles state:

> (D): Just those action norms are valid to which all possibly affected persons could agree as participants in rational discourses. (BFN, 107, See also MCCA, 66; BNR, 80)[2]

> (U): Valid moral practical norms must satisfy the condition that the foreseeable consequences and side-effects of their general observance for the interests of each individual must be acceptable by all those possibly affected in their role as participants in discourse. (BNR, 80, See also MCCA, 65)

In the early 90s Habermas settles on the view that these principles have a different status. (MCCA, 66; BFN, 107, 121; BNR, 84, 89) Albeit, one of his most important colleagues, Karl Otto Apel and other critics (Benhabib, 1990: 345; Larmore, 1995: 66-67) claim that the content of the moral principle is already contained in the principle of discourse (Apel, 1998). If this reading is correct, then (U) seems to be redundant and Discourse Ethics is well equipped with the principle of discourse. (Benhabib, 1986; 387; Benhabib and Dallmayr, 1990; See also Larmore, 1995) To a certain extent, Habermas has some of the responsibility for this interpretation. This is because in MCCA he claims that (U) and (D) should not be confused but that the principle of discourse already contains 'the distinctive idea of an ethics of discourse'. (MCCA, 66) As a matter of fact, in earlier versions of Discourse Ethics Habermas claims that (D) is the moral principle and (U) is just an elaboration of it. (Finlayson, 2016b: 3) Therefore, if (D) contains the essentials of

---

[2] The formulation of this principle in MCCA is identical to the one proposed in BNR in terms of content, 'Only those norms can claim to be valid that meet (or could meet) with the approval of all affected in their capacity as participants in a practical discourse'. (MCCA, 66)

Habermas's moral theory, then the question is where is the specific originality of (U) which supposedly reconstructs the moral point of view of this same theory. In this Chapter, following Habermas, I show that these principles are different. That said, I claim that the Kantian presuppositions of Discourse Ethics apply both to (D) and (U).

In *Between Facts and Norms* and in *Between Naturalism and Religion* the distinction concerning the principle of discourse and the principle of universalization is notably sharpened. (BFN, 107; BNR, 80) In both books, Habermas states that (U) is a moral procedure of testing of norms and the principle of discourse (D) is defined as a *weak concept of normative justification* (IO, 45; BNR, 87) that only incorporates an idea of 'norm-justification in which individuals are viewed as mutually accountable agents'. (Baynes, 2016: 115) (D) does not include *thicker* components, for example, 'the equal consideration of interests of all possible affected' (BNR, 86) which is an element that has been built into the principle of universalization (U). Rather, the principle of discourse expresses the post-conventional 'need of justification only in very general terms with respect to action norms as such'. (BNR, 80) This broader scope leaves room for further specifications of the domain in which this principle is applied – i.e., politics, law and morality. (BNR, 80-81) In this way, concerning the principle of discourse, 'despite its normative content, it lies at the level of abstraction that is still neutral with respect to morality and law'. (BFN, 107)

The difference between these principles has far reaching consequences for the architecture of Habermas's project of social criticisms, inasmuch as by means of this distinction political legitimacy is differentiated from morality. On the one hand, Habermas's notion of legitimacy derives from the interpenetration [*Verschränkung*] between the principle of discourse and the legal form. (BFN, 82-131)[3] On the other hand, Habermas's notion of moral validity depends on principle (U) which he claims can be derived from the combination of the principle of discourse and the rules of discourse. (MCCA, 96-97)

In this thesis, I do not challenge the distinction between (D) and (U) that Habermas eventually wants to draw at the centre of the reconstruction of the normative point of view. Hence, I do not aim to blur the differentiation of spheres of validity that he asserts he

---

[3] The legal form or the legal medium concerns the institutions and practices which confer legitimacy to upon legal and political norms. Among its features we find that modern law: is positive, in the sense that its norms 'stem from the changeable decisions of a political lawgiver' (IO, 254); 'it has been passed by a legally body correctly, according to its rules […] it is enforceable by various legitimate means'. (Finlayson, 2011: 10) Finally, it protects subjective rights. (Habermas, 2011: 285)

builds. Rather, my claim is weaker and I maintain that either we look at (D) or (U) and distil their main contents what we have at bottom is a *modified* and *attenuated* Kantian understanding of practical reason.[4] If that statement is correct, then a Kantian notion of autonomy frames discourses of justification of moral norms and of legal and political norms. In this way, Habermas distinguishes between moral autonomy and civic or political autonomy. In the latter realm, he distinguishes between private autonomy and public autonomy. Private autonomy refers to doing what one wants provided other people can too. Public autonomy concerns to being part of a self-legislating community so one is both *addressee* and *author* of the law. Finally, moral autonomy is acting deliberatively and self-consciously on a principle that is universally valid according to (U).

The aim of this Chapter is to examine the appropriation of the Kantian notion of autonomy – that is connected to all these forms of autonomy – and the 'invasive intervention' (T&J, 87) that it suffers in Habermas's theory.[5] Therefore, I deny from the outset that the Habermasian concept of normative justification is identical to the principles developed by Kant, namely, the categorical imperative and the formula of humanity. (4:421, 429)

In this work, once we have proved that Habermas's *oeuvre* endorses and gives priority to *modified* and *attenuated* Kantian presuppositions in its concept of normativity and morality (Chapter One) and in its notion of legal and political legitimacy (Chapter Three) the second movement is to examine criticisms inspired by Hegel's critique of Kant of this Kantian component in Habermas's moral theory (Chapter Two) and in his political theory (Chapter Four). Particularly in Chapter One, I examine Habermas's re-working of Kant's concept of pure practical reason and autonomy. To develop this issue, I divide this Chapter into five sections.

---

[4] It is possible to give an interpretation of this principle from a Hegelian point of view, in the sense that it arises in the medium of social institutions and practices. Honneth speaks of practices of mutual recognition where individuals ascribe the status of reason-giver to one another. (Honneth, 2014: 42) This coincides with Robert Brandom's pragmatism in which rational agency is fundamentally a normative status dependent on social practices and the attitudes displayed by individuals in the context of those practices. (Brandom, 1994; in Baynes, 2016: 85) These similarities pertain to the analytical reading of Hegel's critique of Kant in which Hegel is continuing and supplementing Kant but not rejecting him. (Pippin, 1989, 1991; Pinkard, 1994, 1999) Hence, the Kantian idea of autonomy finds its expression in the medium of social institutions and practices. In front of this interpretation of Hegel's critique of Kant it reminds the question whether and to what extent Hegel here is more or less an intersubjectivistic version of Kant. Certainly, I cannot examine in detail this problem here. This way to put the issue was formulated by Habermas to Terry Pinkard. (Pinkard, 1999)

In the first section, I show that the Kantian component sets at the front in Habermas's *oeuvre*. This Kantianism arises from his habilitation *Structural Transformation of the Public Sphere*, to one of his less Kantian books *Between Facts and Norms*.[6] In STPS Habermas studies the origins of the public sphere in the 18th century, and in this social space Kantian concepts such as autonomy, the public use of reason and equality were embedded in the medium of institutions and practices of an enlightened society. In BFN, the idea of political self-legislation, the tension between *Facticity and Validity*, the concepts of autonomy and equality also show this Kantianism. In this section, I do not only examine these two books. I also focus on Discourse Ethics and the program of universal pragmatics, and the strategies developed in *Truth and Justification* to detranscendentalize Kant (1);

In the second section, I claim that this Kantian component has a central place in Habermas's critical theory. Habermas defines the *Theory of Communicative Action* as 'the beginning of a social theory that is concerned to validate its own critical standards' (TCA1, xxxix). According to Honneth, the moral theory developed in the program of Discourse Ethics has been built as the justificatory ground of these critical or normative standards. (Honneth, 1991: 282) In Discourse Ethics, Habermas has been explicit that his moral theory is Kantian. Consequently, if the normative ground of the critical theory is framed in Discourse Ethics, then the normative ground of Habermas's critical theory is Kantian. In other words, the Habermasian Kantian notion of moral reason and autonomy lies at the heart of Habermas's broader project of social criticisms. Additionally, I initially develop some of the essential features of the Habermasian reconstruction of the moral point of view, namely, its formalism, cognitivism and universalism; and its relationship with Kohlberg's model of stages of moral development (2);

In the third section, I expound and report Habermas's own view of Discourse Ethics. I examine the chief principle of Habermas's moral theory, namely, the principle of universalization (U), which is the specification of the principle of discourse (D) in the moral sphere of validity. According to Habermas, (U) reformulates 'the basic intuition contained in Kant's categorical imperative […] to ensure that only those norms are accepted as valid which express a general will'. (MCCA, 63) Nevertheless, the principle of

---

[6] I say that BFN is one of his less Kantian books because there Habermas wants to find the support for the justification of the 'system of rights' that compose the *Recht-Staat* in an intra-legal notion of legitimacy, 'overly immanent to law'. (Forst, 2011: 173) Forst and Flynn (2003; 2011), are critical of this strategy because for them the 'system of rights' necessarily needs moral support. This discussion is going to be at the centre especially on the second part of this work.

universalization is an *attenuated* and *modified* version of Kant's moral principle which is not based on the transcendental notion of *freedom* and agency as it was in Kant. Moreover, Kant's reconstruction of the moral point of view refers to a monological account of pure practical reason, whereas in Habermas relates to a dialogical notion of mutual understanding [*Verständigung*] (3);

In the fourth section, I assert that the principle of discourse and the principle of universalization share the Kantian features that I have discussed so far. Interestingly enough, according to Habermas the chief difference between these principles is that the principle of discourse reconstruct the practice of giving reasons in general – it does not have a particular sphere of validity but it is indifferent *vis à vis* morality and law (BFN, 107, 121; BNR, 84, 89) – whereas the principle of universalization regulates interpersonal relationships that concern the moral domain of normative validity. Nonetheless, the essential features of a modified version of Kant's concept of practical reason and autonomy shapes both principles and assures their normative nature (4);

In the fifth section, I discuss the strategies of deduction of the principle of universalization of Discourse Ethics (U). I conclude that at the end Habermas needs to build an historical argument to explain the derivation of this principle and this resembles Kant's strategy of deduction of the categorical imperative as a 'fact of reason'. Here I follow Kenneth Baynes, who asserts that the doctrine of the 'fact of reason' refers to what, 'from a practical point of view, is already familiar to ordinary humans'. (Baynes, 2016: 91) Thus, both Habermas and Kant use this strategy to justify the moral point of view (5).

1. Kantian and Hegelian components in Habermas's *oeuvre*

Kantian components set at the front in Habermas's *oeuvre,* in the sense that he has been always committed to some deeply Kantian presuppositions. Certainly, it is not difficult to trace this relationship and does not require complex strategies to do so. One obvious reason that explains this is that Habermas has openly recognized the connection of the version of critical theory that he develops with Kant. (See MCCA, 68) Thus, the latter is not an obscure figure or the blind spot of the philosophy of the former. Nevertheless, this raises the question of the way in which and the extent to which Habermas endorses Kant's philosophy. In this thesis, I argue that Habermas builds a modified version of Kant's

concept of autonomy which among other things has taken on board Hegelian insights, namely, the idea that the Kantian component needs the support of social institutions and practices of a modern ethical life [*Sittlichkeit*]. (See PR § 135, 153)

Notwithstanding the fact that Habermas tries to take on board Hegel's criticisms of Kant, in this work I maintain that the Kantian component has priority. In further chapters, I discuss this element in light of criticisms inspired by Hegel's critique of Kant's moral philosophy. In this section, I reconstruct Habermas's Kantianism and the incorporation of Hegelian components in a preliminary and introductory fashion, considering some relevant milestones of Habermas's *oeuvre*.

<center>***</center>

The 1962 publication of Habermas's habilitation *The Structural Transformation of the Public Sphere* (STPS) shows an early interest in the domain of civil society, where Kantian values like publicity, inclusion, equality and autonomy take a central place. Here, Habermas describes a detailed social history of the development of the bourgeois public sphere from its origin in the 18th century. In this context, Habermas is neither arguing that at that time the values of the enlightenment were fully developed nor that they were merely ideology, as a Marxist critique would argue. Rather, in Habermas's view modern societies have an immanent rationality that affords the possibility of building the critical point of view. In this regard, this work develops the view that the Kantian ideals need to be embedded in the medium of social institutions and practices. (Hegel, 1991) Henceforth, Habermas's habilitation shows a central motive of his philosophy, namely, the dialectical relationship between Kantian and Hegelian components.

These motives are present three decades later in Habermas's second major book *Between Facts and Norms*.[7] Here, he develops some of the initial insights that shaped STPS. In this way, Habermas argues that constitutional democracy has never been fully developed but its rationality is part of the self-understanding of modern societies. This rationality – which is expressed in the ideas of self-legislation and equality – is embodied in the legal medium of modern democracies. (BFN, 82-131) Thus, considering STPS and BFN, Habermas is

---

[7] It is well known that the first major work of Habermas is the *Theory of Communicative Action*. (TCA1, TCA2) Just to clarify, at this point I am referring to the second major work *Between Facts and Norms* because there Habermas develops forward some of the thesis that were present in an embryonic state in STPS.

developing Kantian contents at the centre of his theory of legal and political legitimacy which in both works are embedded in the medium of social institutions and practices.

Few years after STPS, Habermas's 1968, *Knowledge and Human Interest* (KHI) shows the attempt to appropriate Kant's critical philosophy without endorsing some of its more controversial claims – i.e., its world-constituting subject and the two world-metaphysics. In this regard, KHI 'in some ways Habermas's least Kantian work, opens with an appreciation of Kant's enterprise'. (Baynes, 2016: 83):

> The critique of knowledge was still conceived in reference to a system of cognitive faculties that included practical reason and reflective judgement as naturally critique itself, that is, a theoretical reason that can dialectically ascertain not only its limits but also its own idea. (KHI, 3)

In the 1976 *Communication and the Evolution of Society* (CES), Habermas defines universal pragmatics by the task of identifying universal conditions of possible understanding. This program is similar to the Kantian project which aims to examine the conditions of the experience of the objects of thought. However, Habermas distinguishes his own version of universal pragmatics from the more fully Kantian one developed by Apel. According to Habermas, his theory develops a 'quasi-transcendental' account of universal pragmatics, whereas Apel's constructs a transcendental one. The latter reconstruction of universal pragmatics means, 'What we must necessarily always presuppose in regard to ourselves and others as normative conditions of the possibility of reaching understanding (Verständigung); and in this sense, what we must necessarily always already have accepted'. (Apel, in CES, 2) The aprioristic [*immer schon*: always already] adds a mode of necessity that expresses a transcendental constraint to which we are subject when we perform or respond to a speech act.

According to Habermas, Apel's justification of universal pragmatics is similar to Kant's transcendental argument for freedom. (See MCCA, 77 & 95 and also BNR, 77) In the *Groundwork of the Metaphysics of Morals* transcendental freedom and hence noumenal agency must be located outside time and space. In contrast, Habermas's 'quasi-transcendental' argument appeals to what given certain social practices, is extremely difficult to imagine doing without. (Baynes, 2016: 91) In this respect, as I am going to develop in the last

section, Habermas's strategy of deduction of communicative reason bears a strong similarity with Kant's doctrine of the 'Fact of Reason' to give a foundation to the moral law. (Baynes, 2016: 91) By means of this strategy of argumentation, it seems that the Hegelian insight that reason is embedded in social practices and historical time also inflects Habermas's understanding of universal pragmatics.[8]

In the program of Discourse Ethics the continuities between Kant and Habermas are even clearer. Habermas's moral theory is located in the Kantian family. (MCCA, 67, 68 & 195; Rehg, 1994, 2, 114 & 123; McCarthy, 1978: 326; Forst, 2011: 154; Baynes, 2016: 102) In these moral theories, the main task is to construct the moral point of view, as the procedure through which norms can be tested for their universal validity. Kant constructs the moral point of view as the procedure from which a rational subject can test the universal validity of a maxim. Insofar as the maxim does, so accord and the subject acts upon it, (and thus acts for the sake of duty) the subject is fully autonomous, because as Kant says, the moral will and the autonomous will are one and the same. (4: 447) Likewise, the discursive theory of morality builds the Kantian concept of autonomy but based this time on an intersubjective account of practical reason.[9]

This version of autonomy opens up a challenge for Habermas because he needs to show how a moral subject can be fully autonomous, and yet form their moral will in the process of dialogue. The answer is what Habermas calls, following Mead, 'the larger self' or the 'decentred self' in which the subject brings her own will into line with what everyone can agree in discourse. (MCCA, 65, 121 & 198) This is not a heteronomous process of conforming to what everyone thinks, but a process of determining what the particular agent has most reason to do. This notion of intersubjective autonomy is going to be developed in this Chapter. (See Section 2 and 3)

In the following sections, I examine in detail the continuities between Habermas and Kant concerning their moral theories. Nevertheless, at this point I would like to mention that Habermas's theory is Kantian not only because it aims to build the moral point of view to reconstruct universalizable norms and because puts at the centre the concept of autonomy. Rather, it is also Kantian because Discourse Ethics is formalist, cognitivist and

---

[8] The strategies of deduction of the moral point of view in Discourse Ethics are going to be discussed in detail in the last section of the Chapter.
[9] This connection is examined in detail in further sections of this Chapter.

deontological.[10] Concerning the Hegelian components of Discourse Ethics, they are examined in detail in Chapter Two. For the sake of my argument at this point I limit myself to refer to Habermas's MCCA where he argues that post-conventional moralities need the support of a life-world context that meet them halfway. (MCCA, 207)

Certainly, it requires further analysis to find the coherence between these continuities in Habermas and Kant, and the fact that the former has a critical reading of the philosophy of the subject or consciousness that is present in the latter. Baynes rightly states the dilemma because the relationship amid these philosophers 'at least raises the questions of the extent to which one can follow Kant without likewise embracing the philosophy of the subject'. (Baynes, 2004: 195) However, every relevant philosopher is engaged critically with the tradition that he/she elaborates in an original way. Otherwise, there would not be progress. To my mind, this is what Habermas does with Kant.

Habermas endorses a Kantian pragmatism which aims to 'detranscendentalizing' Kant. (T&J, 84, 175-176; Fultner, 2003: xii) Among other things, this means that Kantian notions like autonomy are this time embedded in the medium of social space and historical time. In this way, Habermas (again) incorporates a Hegelian insight that fortifies his theory in front of criticisms inspired by Hegel's critique of Kant. Albeit, as I will argue in this thesis, I claim that the Kantian component of Discourse Ethics has priority over the Hegelian one.

Moreover, in *Truth and Justification* Habermas addresses the post-Kantian – and Hegelian – task of 'detranscendentalizing' Kant. Moreover, he wants to avoid the pitfalls of other forms of pragmatism in which, he claims, lifeworlds or linguistic frameworks are given too much constitutive authority – i.e., Heidegger's 'Being' or Hegel's 'spirit'. Habermas attempts to accomplish this by means of his notion of communication. Via Thomas McCarthy (1991), Habermas argues that there are 'genealogical' connections between idealizing presuppositions of communicative action and the Kantian 'ideas of reason'. I propose two interpretations to read the idea of 'genealogical' connections. A weak reading implies that there is a *family resemblance* between the presuppositions of communication and Kant's concepts. (T&J, 87) A strong interpretation would mean that these presuppositions are identical to the Kantian concepts. In Habermas's theory the latter interpretation cannot apply, because he works with a *modified* and *attenuated* view of reason compared with Kant.

---

[10] These features are described in more detail in the next section of this Chapter.

Where pure practical reason in Kant is based on the philosophy of the subject, communicative reason in Habermas is grounded on an intersubjective paradigm which pertains to practices of mutual understanding [*Verständigung*] that are explained within the framework of the linguistic turn. (PT, 6, 21; T&J, 1, 220)

In what follows, I develop these 'genealogical' connections. According to Habermas, Kant's idea of the 'cosmological unity' of the world, the idea of 'freedom' as a postulate of practical reason, and the idea of the 'unconditioned' (or God) correspond to three formal-pragmatic presuppositions of communicative action, namely, the common supposition of an objective world, the rationality that acting subjects mutually attribute to one another, and the unconditional validity they claim for their statements with speech acts. (T&J, 87)

Concerning these presuppositions of communicative action, Habermas asserts that they 'refer to one another and form aspects of a desublimated reason embodied in everyday communicative practice'. (T&J, 84) In other words, they are not fully Kantian in the sense that transcendental ideas are necessary constructs of pure reason, that structure and regulate the empirical world. Rather, they are practical presuppositions embedded in the medium of social space and historical time. Habermas describes these 'genealogical' connections in the following terms. (T&J, 87):

(1) between the "cosmological idea" of the unity of the world (or the totality of in the sensory world) and the pragmatic presupposition of a common objective world;

(2) between the "idea of freedom" as a postulate of practical reason and the pragmatic presupposition of the rationality of accountable agents;

(3) between the totalizing movement of reason that, as a "faculty of ideas," transcends all that is conditioned toward an unconditioned and the unconditionality of the validity claims raised in communicative action; and

(4) finally, between reason as the "faculty of principles," which takes on the role of the "highest court of appeal for all rights and claims," and rational discourse as the unavoidable forum of possible justification.

In this section, I focus mainly in the first and the second relationship. The third is referred as long as it helps to explain these two. We can ignore the fourth, since it is not relevant to our purposes.

The first genealogical connection, the common objective world, alludes to the totalizing anticipation of the entirety of objects of possible experience. (T&J, 88) This idea of reason does not make cognition possible, but guides it. This cosmological idea is a methodological principle of completeness and is a regulative idea. Metaphysical thinking, argues Habermas, falls victim to the dialectical illusion of hypostatized world order because it uses this regulative idea constitutively. The reifying use of theoretical reason confuses this idea of reason with an object that is accessible to experience. According to Habermas, the difference between the 'world' and the 'innerworldly' must be preserved 'even if the transcendental subject loses its position outside time and space and is transformed into a multitude of subjects capable of speech and action'. (T&J, 88) The process of detranscendentalization implies that knowing subjects are socialized in the context of a shared life-world, and the pragmatic thesis that knowledge of the world is entwined with speech and action.

The second genealogical connection, the idea of *the accountability of the subjects*, refers to the interpersonal relationships of language users who take one another 'at their word' and hold one another to 'be answerable' for their words and deeds. In their cooperative relationships, they must mutually expect one another to be rational, at least provisionally'. (T&J, 94) This presupposition could be unwarranted and contrary to this expectation, it may happen on any particular occasion, that an agent does not (and cannot) give reasons for her action or claims. However, in contexts of communicative action 'this kind of frustration can occur only against the background supposition of rationality that anyone engaged in communicative action must assume'. (T&J, 94) This supposition of rationality is connected with pure practical reason that determines the agency of the subject according to principles. In Kant's view, by means of the categorical imperative, the idea of freedom acquires its own causality.

In Kant's philosophy freedom is transcendental, a force that originates beyond social space and historical time, but that is supposed to have observable causal effects in the empirical social world. The condition of possibility of this notion of freedom is based on the

doctrine that the causality of the autonomy of the subject depends on the respect for the moral law. This account of freedom acquires legislative force for every rational being. Moreover, in Kant's account of agency there is a distinction between freedom of choice [*Willkür*] and free will [*Wille*]. The former is related to actions in general, namely, moral, pragmatic, prudential and even evil acts. The latter, obeys universally binding norms that have been constructed from the moral point of view.

In Habermas, autonomy is understood as the accountability of the subjects to render validity claims in discourse. Furthermore, communicative action is connected with a broader spectrum of reasons than in Kant. In Habermas, not only pragmatic and moral reasons but also epistemic, ethical and legal considerations are included in practical discourses. Another difference between Habermas's Kantian pragmatism and Kant's philosophy is that in the former the autonomy of the agents is a defeasible claim which means that experience can contradict this presupposition. In contrast, in Kant it is an *apriori* claim that human beings have a natural capacity to cognize the moral law and therefore to recognize themselves as autonomous beings.[11]

In all these various respects, we can see that the Kantian component is essential in Habermas *oeuvre,* in the sense that he has been always committed with Kantian presuppositions. This component arises from in his early works on the public sphere (STPS), his project of universal pragmatics (CES), in Discourse Ethics (MCCA, JA), in his political theory (BFN) and in the presuppositions of communicative action. (T&J) In this regard, in this section I have examined two 'genealogical' connections between the idealizations of communicative action and Kant's 'ideas of pure reason'. Albeit, Habermas has taken on board the Hegelian insight that the Kantian element needs the support and be embedded in the middle of social institutions and practices. However, as I have shown, Habermas has been committed at a deep level to a broadly Kantian notion of autonomy. In this way, in a relatively recent formulation Habermas openly admits to follow Kant in assuming that, with the concept of autonomy, 'the practical reason shared by all persons

---

[11] It is important to notice that in Kant's moral philosophy there is an internal relationship between freedom and the moral law. In the *Groundwork* the strategy was to deduce the moral law from a non-moral notion of freedom. Nevertheless, in the *Second Critique* Kant appeals to the doctrine of the 'fact of reason' which pertains to the claim that 'the moral law neither allows nor needs a deduction in his technical sense. [Rather], on reflection we do accept the moral law as an authoritative standard'. (Reath, 1997: x) Thus, as far as I know the existence of the moral law - as a fact of reason - I know the possibility of freedom.

offers a reliable guide both for morally justifying individual actions and for the rational construction of a legitimate political constitution for society'. (2011: 284)

After the discussion of these connections, in the following sections of this Chapter I focus on Habermas's moral philosophy. Somehow, this theory was referred in this section when I examined the second 'genealogical' connection. The accountability of the subjects refers to the capability of them to render validity claims in discourses of justification. Certainly, this accountability can take place in discourse of moral justification.

Henceforth, in the next parts of this Chapter I examine the specification of the competence to render validity claims in moral discourse. Nevertheless, before I move into that discussion, a preparatory section is needed to illuminate the place that Discourse Ethics has in Habermas's broader project of critical theory. Otherwise, the decision of developing Habermas's moral theory would have been a merely contingent choice to discuss his Kantianism. Rather, my claim is that Discourse Ethics is central in Habermas's *oeuvre* because there he develops the basic components that furnish the normative element of his critical theory. If it is true that Discourse Ethics is central to Habermas's *ouvre*, and it is presupposed for example by his political theory, then arguments that apply to Discourse Ethics will also apply to BFN and to his debate with Rawls. In the next section, I develop the place that Discourse Ethics has in the broader project of social criticism of Habermas and I examine in further detail arguments that show the underlying Kantianism in his moral theory.

2. The place of Discourse Ethics in Habermas's critical theory

The Kantian component in Habermas's theory is a modified version of Kant's notion of pure practical reason which becomes visible in Habermas's moral theory. In this respect, Habermas asserts:

> In recent years Karl-Otto Apel and I have begun to reformulate Kant's ethic by grounding moral norms in communication, a venture to which I refer as "discourse ethics". (MCCA, 195. See also MCCA, 67 & 68)

The relationship between Kant's moral philosophy and Discourse Ethics can be examined bearing in mind the question for the normative grounds of Habermas's critical theory. As I will show, this issue shows the place of Discourse Ethics in Habermas's broader project of social criticism.

The problem concerning its own normative foundations is one of the central tasks of the theory of communicative action. (TCA 1, xli) In this respect, Habermas's relationship with the tradition of critical theory is one of continuity through rupture. On the one hand, Adorno and Horkheimer – the first generation of critical theory – laboured over the issue concerning the justification of the foundations of their critical theory. (TCA 1, 374) On the other hand, in Habermas's view they did not succeed in doing this. He explains this failure because Adorno and Horkheimer were caught in a model of rationality based on the philosophy of the subject. The first generation of critical theory was trapped in a paradox, namely, their concept of rationality could not explain the point of view from which their theory was able to elaborate the social criticisms of modern societies. In other words, they could not reconstruct the concept of reason from which their critical theory was grounded. Rather, it was not rationality but its radical 'other' – i.e., the body or aesthetics – the source of resistance and critique. (PDM, 285, 291)

According to Honneth and Benhabib, Discourse Ethics is the answer to the problem of the normative foundations of Habermas's critical theory. (Honneth, 1991; 282; Benhabib, 1986: 224-335; Rehg, 1994: 21; Finlayson, 2013: 519-520)[12] Therefore, if the moral theory is essentially framed in Kantian terms, then the normative standards of Habermas's broader project of critical theory belong to this family. As I am going to show, the normative kernel of Habermas's philosophy embodies Kantian ideals like autonomy, equality and universalizability.[13]

---

[12] In the opening Chapter of *Truth and Justification* Habermas asserts that 'the pragmatic approach to language [*Sprachpragmatik*] helped me to develop a theory of communicative action and of rationality. It was the foundation for a critical theory of society and paved the way for a discourse-theoretic conception of morality, law, and democracy'. (T&J, 1) The reading of this passage seems to cast some doubts over Benhabib's and Honneth's thesis because here Habermas claims that his theory of language is the ground of the critical theory and of everything else he has written. Although, I think that this way of putting things does not necessarily overshadow Benhabib's and Honneth's claim because Habermas's moral theory has been framed considering the essentials of his pragmatic approach to language.

[13] See this and the following sections. There are other Kantian ideals that are incorporated in Habermas's theory. For example, the concept of Human dignity. (2010) In this thesis, I focus mainly on the concepts of practical reason and autonomy.

Habermas's moral theory is part of the Kantian tradition. (MCCA, 195-197; Rehg, 1994: 2; Forst, 2011: 154; Baynes, 2016: 102) Among other Kantian features, Discourse Ethics is formalistic, cognitivist and deontological. (MCCA, 196-197) It is formalistic, in the sense that it does not produce or defend any substantive content, but only reconstructs a procedure through which norms of action can be morally justified. It is cognitivist in various senses. It shows morality to be a kind of knowledge. According to Habermas the justification of moral norms has an analogous structure to the validation of scientific statements: there is a thoroughgoing analogy between truth and rightness. (MCCA, 196). That said, it is not cognitivist in the strict sense that moral statements can be literally true or false. (See on this Finlayson, 2005) In both types of discourses it is possible to speak of progress in learning. Finally, it is deontological because it assumes the priority of the right over the good. And the procedure of moral justification does not presuppose a specific conception of the good life. Rather, moral theories within the Kantian tradition – i.e., Scanlon, Rawls and Habermas – assume the fact of the plurality of worldviews and the impossibility of grounding a universalistic morality in particular conceptions of the good life.[14]

These features of Habermas's moral theory of discourse are exhibited by the principle of universalization (U) which is formalist, cognitivist and deontological. In further sections, I examine in more detail this principle. Now, what interest me is to show that the reconstruction of the moral point of view in Habermas presupposes the post-conventional level of normative justification which is a re-working of Kohlberg's theory of moral development. (1981, 1984)

According to Habermas, Kohlberg explicitly endorses Rawls's *Theory of Justice*, Kant's moral theory and the modern tradition of natural law. (MCCA, 119) However, these relationships need more careful treatment because Kohlberg used to think that Kant's categorical imperative was the best model of the post-conventional level 6, given the formal criteria that defined it – universalizability, reversibility and formalism. Afterwards he switches to Rawls's model of the choice of principles in the original position. (Kohlberg, 1981: 197)[15] Certainly, this shift is not too problematic because Rawls himself understands his own

---

[14] According to Finlayson, Habermas only begins to address the question of pluralism and different conceptions of the good in the late 80s and the early 90s. (Finlayson, 2016b: 11)

[15] Nevertheless, Kohlberg Mistakenly reads Rawls's theory of justice as a moral theory, namely, as a general theory of right conduct. According to Finlayson, *Theory* should be read as a political theory. (Finlayson, 2010: 6)

theory as Kantian. (TJ, xviii, 11 &221) Particularly, in Kohlberg's theory the Kantian moral point of view becomes the highest stage of moral development that the agents can reach and refers to the post-conventional level. Kohlberg argues that truly moral reasoning involves Kantian features – which are present in Rawls's theory too – equality, autonomy, universalizability, reversibility and prescriptivity. (1981; 1984) At this point, when Habermas was working on Discourse Ethics in the 80s he gets embroiled in a debate with Kohlberg over the structure of Stage 6 of moral consciousness. Habermas proposes a modification of Kohlberg's model through the filter of the theory of communicative action. For Habermas, this reworking is necessary because he thinks it is not clear whether 'Kohlberg's social perspectives can be linked with stages of interaction in such a way as to permit a plausible grounding of moral stages in a logic of development'. (MCCA; 156) It seems that the reformulation was necessary because Habermas persuaded Kohlberg to change his view and built into the theory the standpoint of practical discourse as the post-conventional level of justification. (See Kohlberg, 1986) For that reason, I examine the post-conventional level as Habermas understands it, and its relationship with Kant's notion of pure practical reason.

Habermas has built into his moral theory the essential features of Kohlberg's model of moral development. Thus, the moral point of view proposed by Discourse Ethics:

> Reflects the very operations Kohlberg postulates for moral judgement at the postconventional level: complete *reversibility* of the perspectives from which participants produce their arguments; *universality*, understood as the inclusion of all concerned; and the *reciprocity* of equal recognition of the claims of each participant by all others. (MCCA, 122)

Henceforth, the post-conventional level of justification, both in Kohlberg's and in Habermas's version, endorses Mead's pragmatism and Kantian concepts. In this respect, it is important to notice that Kohlberg was always influenced by Mead, but Habermas claims that his shift to the Rawlsian model of the moral point of view loses sight of the social dimension of Mead's conception of the self that is socialized into individuality. (MCCA, 119) Concerning the post-conventional level of normative justification, the idea of *reversibility* refers to Mead's 'ideal role taking'– which is going to be discussed in detail in the following section –; the notion of *universality* – which is going to be discussed in detail in

the following section as well – relates to the notion that moral norms should be valid for all affected by it; and finally, *reciprocity* means equal treatment.

The notion of agency that Habermas advances concerning the post-conventional level assumes that agents adopt a hypothetical and reflexive attitude towards norms of action. Additionally, the norms that can be justified at the post-conventional level are shaped in the actual practice of discourse. In other words, they are not pre-given moral norms. Despite Lafont's interpretation of Discourse Ethics (Lafont, 2004), Habermas is not a moral realist. (Finlayson, 2005: 342) Rather, the practical discourse shaped by the moral point of view offered by Discourse Ethics 'are not just heuristic aids but, on the contrary, "constitute" or establish the status of a norm as a moral norm'. (T&J, 258; Baynes, 2016: 106) In this regard, Habermas asserts:

> Our moral convictions must ultimately rely on the critical potential of self-transcendence and decentering that […] is built into the practice of argumentation and the self-understanding of its participants. (T&J, 109)

At this post-conventional level of normative justification, which is Habermas's re-working of the Kantian moral point of view, the agents move towards a '*progressively decentered understanding of the world*'. (MCCA, 168) And agents at Stage 6 (the post-conventional level) do not so much pursue their own interests and good, but rather understand their deepest interests as entwined with the interests of all others, and so they resolve conflicts on the basis principles that are in the equal interests of all. At this stage, the validity of moral norms is reconstructed in the light of procedures of justification. Discourses Ethics, proposes the principle of discourse and the principle of universalization to reconstruct or 'constitute' (Baynes, 2016: 107) the moral point of view that establishes procedures of justification. Having described some of the main Kantian features of Habermas's reconstruction of the moral point of view, in what follows I examine in detail one of the central components of its Kantianism, namely, its procedural character.

3. The categorical imperative and the principle of universalization

We have established that Discourse Ethics is a moral theory in the Kantian tradition. (MCCA, 195; JA, 1, Rehg, 1994, 2, 114, 123; Forst, 2011: 154; Baynes, 2016: 102) Among

other components, Habermas's Discourse Ethics and Kant's moral philosophy share a proceduralist or formalist character. This means that both propose principles to assess the moral validity of norms. Nevertheless, an important difference arises between the statuses of these principles. Kant's categorical imperative has the structure of a basic or fundamental norm, whereas (U) just spells out the conditions of validity of a moral norm. In other words, Habermas's principle of universalization is not a specific norm but a procedure by which norms are selected as valid. Indeed, both principles work as criterions to assess the moral worth of norms of action.

To open this discussion, it is important to notice that the formal or proceduralist nature of these moral philosophies takes us into an historical argument. According to Habermas, the development of modern societies takes place in the course of an historical process in which at the post-conventional level the unity between morality [Moralität] and ethical life [Sittlichkeit] – that is immanent to the traditional order – crumbles. (TCA 2, BFN, 84) Thus, a critical attitude frames the practices of mutual understanding [*Verständigung*] and the social world and its norms are open to questions concerning their validity. More precisely, the social world is divided between norms in general and justified norms. By means of this historical argument, Habermas is not giving an account of morality per se, or a definition of it. Rather, he is offering a reconstruction of actually existing modern morality.

In this historical setting, moral theories in the Kantian tradition, Habermas maintains, aim to reconstruct moral objectivity in a posttraditional order in which is no longer tenable to defend an overarching moral authority agreeable to all. In a modern context, a proper moral theory has to come to terms with diverging views of the good life and the challenge of cultural pluralism. (Rehg, 1994: 33) Kantian moral theories in general and Habermas's Discourse Ethics in particular, acknowledge the fact of an increasingly pluralist and fragmented world. (Forst, 2007: 92)

As it is well known, Kant elaborates the categorical imperative as a formula of universalization of norms because he observes as a fact the impossibility to ground morality in a particularistic view. Otherwise, norms would not be universally binding. It is a consequence of the enlightenment the recognition of different positions and doctrines regarding the status of the good and the *reasonable* pluralism of traditions and beliefs that seek to give answers to these questions. Practical philosophy has to pursue universality in

the grounding of norms, otherwise, persons that do not share the same particularistic positions might not have reasons to recognize these norms. The categorical imperative is a procedure that neither defines norms a priori nor grounds them from a particularistic position. Rather, it works as a way to assess the universalizability of norms.

Kant's categorical imperative is stated in the following terms:

Act only according to that maxim whereby you can at the same time will that it should become a universal law without contradiction. (4: 34)

Habermas claims that Kant's Categorical Imperative bases the authority of morality on pure reason, and assumes that rational subjects (who are also finite and sensible) can work out individually whether a maxim can be a universal law, be seeing whether it can be universally applied consistently. In the face of Kant's formulation of the moral point of view, Habermas develops the principle of universalization of Discourse Ethics. In Kant's theory, morality is based on the philosophy of the subject. In other words, Kant's categorical imperative is grounded on a monological notion of pure practical reason and Habermas's principle of universalization in a dialogical conception of communicative reason. (RPR, 32) In Kant normativity depends on the autonomy of the moral agent to give a law to herself. And individual subjects are autonomous if they, using their reason, rationally prescribe a law to themselves, and act on the law solely because it is rational. Although Kant's notion of autonomy involves consideration of the relations to others (i.e. assessing the universality of the maxim) it is a matter of the rational self relation of the individual subject.

In contrast, in Habermas, normativity is embedded in social practices of mutual recognition, where agents ascribe the status of reason-giver to one another. (Baynes, 2004: 197) This practice supposes a Hegelian and Meadian insight which has been incorporated in Habermas's moral philosophy: only in social space and historical time the agents accept and recognize a moral norm as valid. (T&J, 258) Despite this difference, in Kant and Habermas autonomy is at the centre. (Habermas, 2011: 284; Forst, 2011: 154; Baynes, 2016: 108) In the former understood in a monological way, in the latter, as a competence of the agent that is performed in intersubjective practices of mutual understanding and in which participants in discourse as moral subjects are also authors of their moral norms.

The principle of universalization in Habermas is not exhausted by the requirement that moral norms must take the form of unconditionally universal statements. In Discourse Ethics, the idea of generalization of maxims intends that valid norms must deserve recognition by all those that are concerned. And Discourse Ethics appeals to the impartiality that takes place only in a standpoint where 'one can generalize precisely those norms that can count on universal assent because they perceptibly embody an interest common to all affected. It is these norms that deserve intersubjective recognition'. (MCCA, 65) In this regard, Habermas's theory supposes a modified version of Kant's notion of pure practical reason because the former grounds morality in a dialogical account of reason and the latter in a monological one. Here Habermas agrees with Apel, for whom argumentation is the reference point of unavoidable rules of the communicative practice. These rules of communication frame the moral point of view in Discourse Ethics.[16]

The intuition behind Discourse Ethics is that a monological application of the categorical imperative does not guarantee the intersubjective validity of moral judgements. Hence, in Habermas 'it does indeed seem an obvious step to formulate Kant's *assumption* in the form of a *postulation* of the sort, 'Act in such a way that your way of acting could be willed by *all* as a universal one''. (Wellmer, 1991: 154) To a certain extent, this understanding of the principle of universalization of Discourse Ethics shows that it has built an intersubjectivistic notion of practical reason. In Habermas, moral validity points to an intersubjective structure linguistically mediated which frames the unconditional character of the moral 'ought' and even shapes our identity as human beings. (Wellmer, 1991: 152)

This recognition of the fundamental place that intersubjectivity has in morality leads to a version of the universalizability of norms in which impartiality is expressed in a principle that constraints '*all* affected to adopt the perspectives of *all others* in the balancing of interest'. (MCCA, 65) This is G. H. Mead's concept of 'ideal role taking' which expresses a situation that 'requires that any morally judging subject put itself in the position of all who would be affected if a problematic plan of action were carried out or if a controversial norm were to take effect'. (MCCA, 198, See also T&J, 105)

---

[16] This needs more careful treatment because in Habermas's theory the rules of discourse have to be combined with the principle of discourse (D) to reconstruct the moral point of view (U). This discussion is developed in detail in the two next sections.

According to Rehg, without the introduction of this notion of moral perspective-taking it would be difficult to distinguish this principle of universalization from the strategic considerations underlying marriages of convenience, 'hardly a palatable result in any moral theory, least of all in a neo-Kantian one'. (Rehg, 1994: 40) However, according to Habermas Mead's ideal is still monological because it can be practiced by the agent privately. (MCCA, 198) Discourse Ethics transforms the notion of 'ideal role taking' into a principle that is performed communicatively *by all those involved*. Thus, the principle of universalization shifts Kant's emphasis from the individual what *I* ought to will to become a universal law to the collective what *we* ought to will.[17]

Another important feature of (U) is that it is not a first order moral norm resulting from an actual instance of successful moral discourse 'but a second order principle that captures the practice by which actual moral norms are selected'. (Finlayson, 2013: 524)[18] This is the chief respect in which Discourse Ethics is procedural and not substantive.

Moreover, this version of the principle of universalization explains an important difference between Habermas and Rawls. It also may help to understand what Habermas is expressing with the term 'dialogical'. He argues that Rawls's theory of justice is monological because it allows that a single individual, reasoning carefully, could arrive at a proper understanding of the requirements of morality. (MCCA, 66) In contrast, Habermas's moral theory claims that the identification of correct moral norms is a result of practical discourses actually carried out by the participants.

However, it has been argued that dialogicality admits of a weak and a strong interpretation and in that respect Habermas's notion of dialog opens an ambiguity. (McMahon, 2000) The first interpretation which McMahon calls 'weak dialogicality'

> Involves a simple requirement to consult everyone who might have evidence germane to the identification of the correct principles of morality. Strong dialogicality provides in addition that judgements identifying the correct principles

---

[17] This obligation to include all individuals as potential participants in discourse – as long as they are involved – 'presupposes a universalistic commitment to the equality, autonomy, and rationality of individuals'. (Benhabib, 1986: 319) Thus, Habermas's moral theory demands that the participants have to regard 'every being as equal'. (Benhabib, 1986: 320)

[18] As I have shown, this is clear with U but it is not clear with Kant's categorical imperative which is formulated as a command. Therefore, it looks like the categorical imperative is a basic norm or fundamental norm whereas U just expounds the conditions of moral validity of norms.

of morality must be made collectively by all those potentially affected. (McMahon, 2000: 514)

In the weak interpretation of Discourse Ethics according to McMahon, which is the only defensible one, there is no difference in principle between dialogical and monological approaches, and if he is right in this, Discourse Ethics is no different from other moral philosophies in the Kantian tradition – i.e., Rawls. The weak interpretation implies that every subject on its own decides what is morally justified and the communicative process is a mean to gather all the relevant information to reach an agreement.

On the contrary, in the strong interpretation, autonomy is now practiced collectively by all those possible affected by a norm and it implies a strong assumption about the capacity of persons for moral dialogue. According to McMahon (2000: 526) strong dialogicality is incoherent, because no individual has any basis for judging which norms are valid until the moment when everyone as it were casts their vote and agreement is unanimous. But that robs participants of all basis of judgement about whether a norm is valid nor not, which makes it impossible to see why discourse would arise in the first place.

Nevertheless, Habermas – and also Rehg – continues to endorse and defend the strong interpretation. (JA, 47-50; Rehg, 1994: 39-40, See also Baynes, 2016: 122-123) On his view, justified norms are shaped in a collective process of mutual understanding and here Habermas departs from Kant and Rawls concerning the distinction between monological and dialogical. In this way, Habermas's version of the moral point of view represents a modified and intersubjective version of Kant's concept of pure practical reason.

Henceforth, Habermas's moral theory belongs to the Kantian family because puts at the centre a Kantian notion of autonomy. And regardless of which reading is right the strong or the weak version of dialogicality, Habermas still defends a basically Kantian idea of autonomy. This is obvious in the first case since there is no difference in principle between Kant's categorical imperative and Habermas's principle (U). That said, if the strong interpretation is the correct one, then a Kantian version of autonomy is in play because in discourse each participant, as a decentred self, has made norms that are in the equal interest of all affected their own.

4. The derivation of the principle of universalization

In the previous sections – mainly in sections two and three –, I examined the essential components of Habermas's moral theory and its relationship with Kant. So far, I have discussed the intersubjective character of the principle of universalization of Discourse Ethics. This principle belongs to the Kantian family but certainly should be distinguished from Kant's categorical imperative. This is because the latter is based on an understanding of practical reason grounded on the philosophy of the subject and the principle of universalization depends on Habermas's notion of communicative reason which is based on the philosophy of language. In despite of this substantial difference, Habermas is still an heir of Kant because Discourse Ethics incorporates several Kantian features.

Nonetheless, the description of this relationship does not explain yet how Habermas grounds or derives his version of the moral point of view in the principle of universalization. The derivation had its first formulation in 'Discourse Ethics: Notes on a Program of Philosophical Justification' and elsewhere (MCCA, 43-109; IO, 43-45) and it has been commented and examined by more or less sympathetic critics of Habermas. (Rehg, 1994: 56-69; Benhabib, 1986: 306-309; Baynes, 2016: 114-120; Finlayson, 2013: 523-525) Habermas's *early* formulation suggested that one could 'derive' (U) by material implication from two not uncomplicated premises. The first refers to 'what it means to discuss hypothetically whether norms of action ought to be adopted. The second summarized the pragmatically unavoidable presuppositions of argumentation, as Alexy had spelled these out'. (Rehg, 1994: 40) Thus, the two premises from which (U) is derived are:

> (1) the normative, (but non-moral) preconditions of argumentation in general. (MCCA, 89)

> (2) a 'weak idea of normative justification' or 'the conception of normative justification in general as expressed in (D)'. (MCCA, 66, 93)

The first premise, the principle of discourse (D), guarantees normative validity in general. Albeit, (D) originally was introduced as the leading principle of Discourse Ethics, Habermas soon given it a broader (pre-moral) interpretation because 'it demands a post-conventional justification of norms of action in general, but without specifying a particular

respect in which the consensus-generating force of reasons is supposed to be mobilized'. (BNR, 109; Rehg, 1994: 30) In the more recent formulation, the principle of discourse can be described as 'what it means to discuss hypothetically whether norms of action ought to be adopted'. (BNR, 89; Rehg, 1994: 58) In other words, it pertains to the justification of a 'norm of action' in general. (MCCA, 198)

It is important to notice that in both the early and the more recent formulation of the nature and the scope of the principle of discourse, (D) belongs to the Kantian family. In this respect, in his very influential book *Insight and Solidarity*[19] Rehg points out that (D) is *deontological, formal, cognitivist* and *universalistic.* (Rehg, 1994: 31) Henceforth, it shares these features with Kant's reconstruction of the point of view of morality.[20] Nevertheless, this might raise the question whether Rehg is referring here only to the early account of (D) – which sets it as the leading principle of Discourse Ethics – or to both the early and the more recent formulation – the latter position gives a wider scope to (D) than the early one. If Rehg is bearing in mind only the early account it is possible to conclude that these Kantian features only frame the principle of discourse when it is specified in the moral domain of Discourse Ethics – when it becomes U –. This argument can be rejected because Rehg explicitly acknowledges that the principle of discourse has a broader scope in writings after MCCA and it goes beyond morality. (Rehg, 1994: 30)[21] Therefore, the principle of discourse both in the early and in the more recent formulation is *deontological, formal, cognitivist* and *universalistic.* Moreover, in BNR Habermas explicitly connects (D) with 'the postconventional level of justification of norms of action in general'. (BNR, 89) This level of normative justification systematically has been connected to a re-working of Kant's concept of autonomy and of practical reason. In this regard, Habermas claims that:

> I follow Kant in assuming that, with the concept of autonomy, the practical reason shared by all persons offers a reliable guide both for morally justifying individual actions and for the rational construction of a legitimate political constitution for society. (2011: 284)

---

[19] I say very influential because this book even helped Habermas to clarify his own theory concerning the status of his notion of dialogicality (See section 3 of this Chapter). Moreover, this contribution has been widely referred to explain some of the most difficult components of Habermas's theory, for example, the derivation of the principle of universalization. (See for example Baynes, 2016: 115)

[20] See section of this Chapter.

[21] Nevertheless, this is a matter of controversy because it has claimed that in Habermas's moral theory, the principle of discourse is a substantial principle (Benhabib, 1990: 345) based on an essential postulate of the respect for persons. (Larmore, 1995: 66-67)

This statement clearly shows that the justification of moral norms and of legitimate laws concerns the Habermasian Kantian notion of practical reason. Insofar as the principle of discourse is the normative kernel of both spheres of validity – and bearing in mind Habermas's passage referred above – this shows that (D) fundamentally belongs to the Kantian family and presupposes the Kantian idea of autonomy. Henceforth, here Habermas implies that the Kantian idea of autonomy is contained already in (D) and thus also in (U) and in the principle of democracy.[22]

Notwithstanding, the previous argument needs further clarification. In Kant, both moral validity and legal legitimacy stem from the moral principle.[23] In Habermas the argument is more complex because the principle of discourse has a broader scope than the moral principle or the principle of legal and political legitimacy. In this regard, Habermas has firmly rejected Karl Otto Apel's argument, which claims that the principle of discourse is already a moral principle and Benhabib and Larmore agree. (BNR, 77-97, Apel, 1998: 689-838; Benhabib 1990: 345; Larmore 1995: 66-67)[24]

He rejects it, as already mentioned, on the grounds that 'despite its normative content, it lies at a level of abstraction that is *still neutral* with respect to morality and law, for it refers to action norms in general'. (BFN, 107; See also BNR, 84, 89) The idea behind this claim is that not all forms of normative behaviour can be examined in moral terms – i.e. legal norms and political norms. Norms in general regulate the actions of the subjects that share a common life-world. Thus, Habermas's view is that (D) is more and less than a moral principle. According to Rehg, 'It is more in that it covers more areas of social action, less in that it does not tell us what distinguishes the validity of moral norms from that of other kinds of action norms'. (Rehg, 1994: 31) Then, the principle of discourse is more general than the principle of universalization. The latter results from (D) and the preconditions of argumentation in general and it is a specification of normative validity concerning questions of morality.

---

[22] The principle of democracy states that 'Only those statutes may claim legitimacy that can meet with the assent of all citizens in a discursive process of legislation that in turn has been legally constituted'. (BFN, 110) For a detailed discussion on this principle see Chapter Three, Part I, section 3.

[23] This discussion is going to be broached in detail in the second part of this thesis.

[24] In the face of this issue, up to this point I am going to assume only that the principle of discourse derives from a Kantian notion of practical reason. Nevertheless, in following chapters I will draw the conclusion that it is not really clear whether (D) is or is not a moral principle. Henceforth, it seems that some concessions need to be made to Apel's claim.

After this description of the principle of discourse (D) and its scope, it is necessary to examine the second set of norms that explain the derivation of the principle of universalization: the normative, (but non-moral) preconditions of argumentation in general. These conditions refer to the commitments that involve to a person to engage in an argumentative discourse. Among these commitments Robert Alexy claims that whoever gets involved in a discourse of justification accept others as speaking partners with equal rights, at least to what concerns the justificatory practice. Moreover, he adds that in discourses coercion is ruled out. (Alexy, 1990)

According to Habermas, these commitments or norms of discourse can be outlined in three pragmatic presuppositions which he regards as one of the components of the derivation of the principle of universalization (U):

Anyone who enters argumentation must take the following presuppositions:
  (a) Every subject with the competence to speak and act is allowed to take part in a discourse.
  (b) i. Everyone is allowed to question any assertion whatever.
      ii. Everyone is allowed to introduce any assertion whatever into the discourse.
      iii. Everyone is allowed to express his attitudes, desires, and needs.
  (c) No speaker may be prevented, by internal or external coercion, from exercising his rights as laid down in (a) and (b) above. (MCCA, 89)

Up to this point, I have described the elements that take place in the derivation of (U) but I have not explained the derivation as such. Whether the logical derivation of (U) was even possible was indeed a matter of controversy. In this regard, Benhabib broached the dilemma in which Discourse Ethics is trapped. Either the moral principle is redundant or inconsistent. (Benhabib, 1986, 307; Benhabib, 1990; Finlayson, 2000, 2013)

The first part of the criticism means that the principle of universalization does not go beyond the principle of discourse. In other words, (U) is redundant because it does not add anything relevant to what is already contained in (D). The second part of the criticism – and the one from which Benhabib develops his program of justification of Discourse Ethics – is that 'a derivation can be had, but it would require supplementary, normatively more robust premises'. (Finlayson, 2013: 524, See Benhabib, 1986: 298-309) In response to

this potentially devastating criticism, in a *later* formulation Habermas 'now claims that the argument for Principle U is best construed as an '"abduction" or "argument from the best explanation" rather than a strict logical derivation'. (Baynes, 2016: 114. See also Finlayson, 2000: 331)

In Habermas's revised formulation – which Rehg offers in a parallel account (Baynes, 2016: 115) – 'it is precisely the *combination* of the two premises, in a modern context of commitment to argument, that leads to a notion of universality exceeding the content of either premise by itself'. (Rehg, 1994: 66, See IO, 45) According to Rehg, Habermas successfully derives the principle of universalization of Discourse Ethics. And the principle is more than the sum of the two premises, it goes beyond them. As I have shown, on the background of this interpretation lies the early criticism made by Benhabib. (1986: 308) In the final section of this Chapter I discuss in detail this criticism. The issue that I want to examine in more detail in the rest of this section is Habermas's and Rehg's argument for the reconstruction of the principle of universalization.

According to Rehg, what makes the difference and allows Habermas to successfully build the moral point of view is 'the semantics of the moral 'ought'. (Rehg, 1994: 68) Then he continues:

> Following the results of our earlier analysis of norms, normative expectations must be seen as extending to anyone who could possibly come under the roles defined in the norm. If the suggestions regarding broader consequences and side effects are not mistaken, we can further extend this to include 'all those affected by its general observance. (Rehg, 1994: 68)

Hence, in Rehg's interpretation, (U) can be regarded as a principle of universalization and it is properly derived from the two premises. That said, it is also necessary to assume the existence of a pluralistic group that decides to resolve its conflicts of interests cooperatively by reaching argued agreement on a norm. Rehg recognizes that this argument is similar to Benhabib's program of justification of the moral point of view which is based on more robust premises based on social and historical arguments. (Rehg, 1994, Benhabib, 1986, Finlayson, 2013) In this respect, Benhabib develops a Hegelian justification to underpin the procedures of Discourse Ethics:

> The *interest* in rational discourse is itself one which *precedes* rational discourse, and it is *embedded* in the contingency of individual life histories and in collective patterns of memory, learning, and experience. (Benhabib, 1986: 319)

If this commitment – which is embedded in historical time and social space – is combined with the two premises – (D) and the preconditions of argumentation in general – then, every moral norm 'must rest on reasons all those subject to (and affected by) the expectation can accept in open debate, for otherwise the norm is not justified for those subject to it, and thus its observance may not be expected of them (nor may the non-interference of other affected parties be expected)'. (Rehg, 1994: 66-67)

After this exposition of the reconstruction of the moral point of view, in the next section I develop in more detail Benhabib's criticism of the derivation of (U). Up to this point and for the sake of my argument I have examined briefly Benhabib's criticism. In what follows, I argue that even though Discourse Ethics incorporates a Hegelian insight, namely, the social and historical commitment of a group to participate in discourse, I will show that Habermas's strategy of 'abduction' or 'argument for the best explanation' (IO, 43) of the moral point of view resembles Kant's deduction of the categorical imperative as a 'fact of reason'. (Baynes, 2016: 85, 91)

5. The Kantian strategies of justification of (U)

In this last section, I examine in more detail the derivation of the principle of universalization of Habermas's moral theory, which refers to the specification of the principle of discourse in the domain of morality. (BNR, 80-81, BFN, 109) To discuss this, I keep examining Benhabib's criticism of Habermas, which asserts that the moral point of view of Discourse Ethics is either redundant or inconsistent. I understand the first part of this claim in the sense that (U) does not add anything relevant to the principle of discourse (D). (Benhabib, 1986: 308) And the second, to mean that (U) is not redundant but it is inconsistent with the premises. In other words, it cannot be derived from them. Thus, in order to give a justification of the moral point of view, Habermas needs the support of other components.

According to the first part of Benhabib's criticism, insofar as the principle of universalization (U) is redundant (1986: 308), then is not necessary for discursive ethics. (Benhabib and Dallmayr, 1990; Wellmer, 1991) Thus, for Benhabib 'discourse ethics is sufficiently equipped with its basic principle (D) and the pragmatic rules of argument'. (Rehg, 1994: 65)

In a first reply to this charge someone might say that (U) adds to (D) the condition that 'the common interest of all concerned means taking into account the satisfaction of the interests of *each single* individual, such that a naïve *collectivist* interpretation of the common interest which would violate minority interests is precluded'. (Benhabib, 1986: 307) However, from the general premise which takes the form of the principle of discourse (D), that states, 'what it means to discuss hypothetically whether norms of action ought to be adopted', it is possible to deduce that the interest of each individual cannot be ruled out if a norm can claim validity. Then, the principle of universalization seems to be redundant because it is not adding anything that it was not already present in the principle of discourse.

This is important because if (U) is redundant, if (D) is already given, then it is not clear what the differences between diverse domains of validity are. Habermas argues in BFN that legitimacy derives from the interpenetration of the principle of discourse and the institutional features of modern law. So, if (U) is redundant, then the difference between legitimacy and morality in Habermas seems to be only a matter of institutional application of (D). In other words, if this part of Benhabib's criticism hits the mark, it is possible to draw the conclusion that legitimacy is equivalent to morality adding the functional features of modern law. From the reading of BFN and BNR, this is at odds with Habermas's understanding of moral validity and legal and political legitimacy. As a matter of fact Habermas has been criticized for giving an intra-legal notion of legitimacy, 'overly immanent to law' and for not taking into account the moral support that the 'system of rights' requires. (Forst, 2011, 173, See also Flynn, 2003, 2011) In this respect, Habermas asserts that

> Although legal norms may also be selected under the aspect of justice and must not contradict morality, the principle of democracy that empowers the citizens to create legitimate law is not subordinate to the moral principle. (BNR, 90)

For Habermas both levels of justification are different – moral validity and legal and political legitimacy. That said, bearing in mind Benhabib's criticism it is still not clear whether (D) is or is not a moral principle. Either way, in this thesis I show that both (D) and (U) share an essential structure which pertains to a modified Kantian notion of practical reason.

From a different angle, it is possible to argue that the moral point of view contained in the principle of universalization connect norms to consequences and interests. This is something that does not seem to be included in the principle of discourse (D). Thus, 'The moral principle first results when one specifies the general discourse principle for those norms that can be justified if and *only* if equal consideration is given to the interests of all those who are possibly involved'. (BFN, 108) However, the standpoint from which Habermas defines norms (in general) is that of social coordination. And according to Rehg:

> This point of departure already contains a certain semantics linking norms to consequences and interests, though not in an unduly consequentialist or utilitarian fashion. (Rehg, 1994: 45)

Thus, as the task of the principle of discourse (D) is to define the validity of norms of action in general, then when Habermas adds to the principle (U) the consideration of consequences, side effects and interests, he is adding something that was already implicit in (D).

The second part of Benhabib's criticism of the derivation of the Habermasian moral point of view makes a twist and consequently has different consequences. In this position, (U) is different from (D) and it goes beyond. However, the principle of universalization is inconsistent with the premises. In the same vain, Finlayson argues that neither Habermas nor any of his followers have ever managed to derive (U) formally or as he says 'immanently' from the premises. (Finlayson, 2013: 518; See also Rehg, 1994: 40) Although, Habermas continues to maintain that such derivation is possible. (Finlayson, 2013: 523) Therefore, (U) is different from (D) and it adds something relevant, but it does not follow from the premises. Instead, it depends on other components which require further justification.

One difference between (D) and (U), according to Finlayson, is that the principle of discourse establishes that amenability to rational consensus is a necessary – but not sufficient condition – for a norm to be valid. Then, bearing in mind (D) it is not the case that every norm on which participants in discourse can agree, is therefore valid. Hence, with (D) the question concerning the actual validity of a norm remains open. On the other hand, (U) defines a procedure that refers to a necessary condition for the validity of a norm but also a sufficient one. (Finlayson 2000: 329; Ingram 2010: 133) If a norm is accepted by all participants in discourse then is valid. Consequently, (U) goes beyond (D). But nothing in either of the premises shows why this is so.[25] As I have shown, neither Habermas nor any of his followers have not managed to derive (U) from the premises.[26]

Habermas has been aware of this problem and he has proposed a weaker justification of the principle of universalization that rests much more firmly on an historical argument based on considerations of the modernization theory. This specific component refers to the commitment of a group to rational argument. Rehg recognizes that this justification of the derivation shares much in common with Benhabib's Hegelian 'weak justification program'. (Rehg, 1994: 68) In this respect, morality is anchored on the development of an enlightened consciousness. And the grounding of morality, of discursive ethics and even of critical theory depends on a theory of modernity and not in the rules of discourse. Consequently, there is not a logical or transcendental-pragmatic justification of the principle of universalization because (D) as a premise in the argument for (U) presupposes modernization theory. Hence, the derivation is not itself a transcendental-pragmatic argument. Rather, what is taking place is the modern self-understanding of what it means to be rational and autonomous subjects in a posttraditional order of justification.

---

[25] However, William Rehg seems to claim that (D) is biconditional and amenability to consensus in discourse is necessary and sufficient condition for the validity of a norm. As I have shown above, U is a biconditional principle as well. Therefore, if both principles share this feature then this seems to undermine Rehg's defense of Habermas's position in the sense that (D) is not a moral principle. In this way, if (D) is so close to (U) – because both principles are biconditional – then this may as well amount to the fact that (D) is the moral principle of Discourse Ethics. (See Rehg, 1991)

[26] As a matter of fact, Habermas's latest considered position seems to be that only one premise (the rules of argumentation) can strictly speaking be given a transcendental-pragmatic justification. Of course, this is a slightly different issue than the one we are discussing at this point. Nevertheless, if the justification of the premises does not fully proceed in transcendental-pragmatic terms, then the idea that the principle of discourse can have this sort of derivation gets blurred. The notion of a pragmatic-transcendental justification proceeds by way of demonstrating that the denial of any of the rules of argumentation involves a performative self-contradiction. This is the case of the sceptic because she implicitly, by virtue of the performative act of making a statement, invokes rules that she explicitly by the propositional content of his claim, denies. According to Habermas the transcendental-pragmatic justification is based on those rules that every participant in a process towards reach understanding has to follow. Even the sceptic that denies them is making use of them when she is incurring in a performative contradiction. However, 'no one has actually succeeded in carrying out such a justification of the rules of discourse'. (Finlayson, 2013: 523)

In this interpretation, the principle of universalization is not superfluous or redundant and the combination of the two premises – in a modern context of commitment to argument – leads to a notion of universality exceeding the content of either premise by itself. (Rehg, 1994: 66) And the *derivation* succeeds only if it takes place in the background of a social space constituted by the commitment of a group towards the post-conventional level of normative justification embedded in the historical dimension of a posttraditional order.[27]

In light of this discussion I would like to state few conclusions. First, the actual measure and substance of the difference between the principle of discourse (D) and the principle (U) is problematic and constitutes an important issue that must be addressed by Habermas and his followers. Second, at the end Habermas's (IO, 43), Rehg's (1994), Benhabib's (1987) and Finlayson's (2013) interpretation conclude that Discourse Ethics is not based on a transcendental-pragmatic justification, but it is based on a theory of modernity. Third, it is not clear whether (D) or the rules of discourse could be justified by a transcendental-pragmatic argument. Habermas tried to give this justification in MCCA and he assumes that Alexy has done this concerning the rules of discourse. (MCCA, 88-89) Nonetheless, following Finlayson, it must be that some hidden premises of the derivation of (U) also are supplied by modernization theory. Four, even in the case that (U) goes beyond (D), then this contribution establishes not just a necessary condition for the validity of a norm but also a sufficient condition. By (U) the validity of norms can be actually assessed, whereas (D) contributes to stating the condition the norms that can claim validity must meet.

However, in either case, the question whether the principle of discourse is or is not a moral principle remains unanswered. In BFN and BNR Habermas strongly denies that (D) is a moral principle. I do not want to argue against Habermas in this respect, others have done so (too a certain extent Benhabib, 1986; and Larmore, 1995). Rather, for my purposes I claim that whether we look at (D) only, or also at (U), the Kantian component prevails in Habermas's concept of normative validity. Hence, Discourse Ethics is still basically Kantian, and wedded to a Kantian conception of autonomy and equality. As long as the principle of discourse is specified in the moral, the legal and the political domain, then, the

---

[27] In despite of this defence, it is interesting to notice that this element is the contribution that (D) gives to the derivation. The universalization of a norm happens when we include 'all those affected by its general observance'. If one recall (D) a norm is valid when it can meet 'with the approval of all affected in their capacity as participants in a practical discourse'. Then (D) itself includes the component 'all affected' and when in (U) is added the 'general observance' of a norm, it does not seem to be adding something new that was not already present in (D).

Kantian notion of practical reason concerns both the moral theory and the theory of legitimacy.

Once I have showed that Discourse Ethics endorses strong Kantian presuppositions, in Chapter Two I challenge this component from the point of view of Hegel's critique of Kant's moral philosophy. Concerning the political theory, in the second part of the thesis I examine the political Kantianism of Habermas – which he defines as Kantian republicanism (MW, 94) – in light of criticisms inspired by Hegel's critique of Kant's moral philosophy.

It is important to remark that here I am not saying that morality and politics are the same, but just that they are grounded in the same Kantian components. The difference comes from the side of the structural features of modern law and of the democratic state that does not furnish morality. Up to this point, certainly this discussion is not conclusive and it aims to illuminate a difficult issue in Habermas's theory. Although, the claim is that at the end normativity, morality and legitimacy in Habermas all rest on a modified Kantian concept of practical reason, which has a weak-transcendental status. It is transcendental inasmuch as refers to the factual counterfactual presuppositions of justification, but it is weak because it needs the support of an historical argument provided by the theory of modernity. In other words, it depends of a Hegelian component, namely, a set of institutions and practices of modern *Sittlichkeit*.

It has been pointed out that a similar argument takes a central place in Kant's practical philosophy in the doctrine of the *fact of reason*. (Baynes, 2016) Due to this stance, he abandons the attempt to provide a transcendental deduction of freedom:

> In the second *Critique* he instead treats it as a "fact" to which appeal can be made to help make explicit what is already implicitly known in practice or, in Kant's words, already known by "common human understanding". Thus, rather than a transcendental argument for freedom, the doctrine of the "fact of reason" helps us to better understand (and to resist naturalist or sceptical objections) what, from a practical point of view, is already familiar to ordinary humans'. (Baynes, 2016: 91)

I cannot analyse in detail if Kant was successful or not with this strategy. My focus is Habermas's Kantianism and the weaker alternative proposed by Discourse Ethics to give a justification to the moral point of view. As I have shown, Habermas appeals to a more modest or weaker type of argumentation to reconstruct the principle of universalization. He refers to 'what is (nearly) unimaginable from a practical point of view', and this 'is more clearly seen in his remark that we are all "children of modernity", that is, products of historical and thus contingent traditions which are, nonetheless, practically inescapable for us'. (Baynes, 2016: 91) Here, the contingent nature of these practices is remarked – their dependence on an historical context – but at the same time they become non-contingent because as we are within a life-world we cannot avoid these rules.

These practices give a central place to a notion of autonomy and it is here where the Kantian heritage echoes in Habermas's Discourse Ethics. This Kantian concept occurs at every stage in this theory because normative validity claims only have a binding force among individuals who consider themselves accountable – and to that extent are autonomous – persons. Roughly speaking, this accountability involves the possibility of giving reasons. Also, as I stated in the first section, rational agents start from the presupposition of the accountability of the other subjects, and they demand each other this capability to give reasons.

Consequently, if there is a commitment to argument, that implies treating all the competent speakers as equal dialogue partners. Henceforth, I read this commitment in terms of the Kantian notion of autonomy, which in both Kant and Habermas goes alongside the equality that the agents mutually grant to each other. Moreover, autonomy means that different members of a life-world may be expected to observe a norm precisely because its validity can be made evident to their reason. (Rehg, 1994: 35) This time a norm is evaluated and it obtains its validity in a concrete communicative practice of exchange of reasons. This actual carrying out of discourse is fleshed out through the validity claims raised by real participants in communication and it is not decided in a monological fashion as it is the case in Kant. This is how the Kantian concept of autonomy takes place in the discursive theory of morality. Moral agents must give reasons for their action and accept or reject justifications in a collaborative process of mutual understanding [*Verständigung*] and they can do this because they are autonomous. This intersubjective concept of autonomy grounds an orientation to resolve conflicts of interests by reaching argued agreement in

discourse. According to Discourse Ethics, autonomous subjects co-constitute the validity of the norms by which they act and they make these norms their own. This is because, they are decentred selves at the post-conventional level of normative justification which I have already discussed (stage 6).

2

Hegelian criticisms of Discourse Ethics

In Chapter One I have shown that Habermas's account of normativity and morality relies on a *modified* Kantian notion of pure practical reason and autonomy. As far as Habermas incorporates this component, it seems relevant to discuss whether and to what extent his moral theory can rebut Hegelian criticisms inspired by Hegel's critique of Kant. Habermas addresses this issue in his article 'Does Hegel's critique of Kant apply to discourse ethics?'. (MCCA, 195-215) There, he elaborates several arguments to show that his Kantian moral theory can respond, if not to all, then to most of the objections that stem from Hegel's critique. Now, before I examine Habermas's stance on this debate, in this introduction I need to begin describing, at least in broader brushstrokes, Hegel's critique.

In its early formulation, Hegel claims that the *empty formalism* of Kant's moral philosophy does not allow him to give a proper account of the phenomenology of the moral life. This objection was developed from his early essay on *Natural Law* (1802), his *Phenomenology of Spirit* from the middle period in Jena (1807) and in *Elements of the Philosophy of Right* in the late period (1820). In the latter book, Hegel not only criticizes the so-called *empty formalism* of Kant's moral philosophy, but also he develops an objection to the will as a tester of maxims. In this Chapter, I show that this charge does challenge Kant's moral theory and Habermas's Discourse Ethics as well. However, I maintain that with minor modifications, Habermas can successfully refute the criticism of his moral theory. In this Chapter I will spell out what these modifications of Habermas's theory are. Before I examine these problems, now I have to develop the criticism of the will as a tester of maxims. In a nutshell, Hegel's argument is that Kant cannot connect the universal will and the particular will because he rules out particular contents as susceptible of being universalized. Due to

this break Kant cannot explain one of the conditions that show how an empirical will can be motivated and have a reason to be moral. In this respect, in the *Philosophy of Right* (1820) Hegel asserts:

> If a duty is to be willed merely as a duty and not because of its content, it is a *formal identity* which necessarily excludes every content and determination. (§135)

The main reason that explains this problem is that Kant was trapped in the doctrine of the two realms. It is important to notice that Hegel interpreted Kant's transcendental idealism as positing two metaphysically distinct worlds, rather than as describing one and the same world under two aspects.[1] In this metaphysical view, the law giving self is fully *noumenal* (hence transcendentally free) while the law receiving and respecting self is also *phenomenal*. The issue is that this distinction might imply that the *noumenal* self lays down the law leaving unaffected the *phenomenal* self. Moreover, the distinction supposes that in Kant the contents that morality allows are strictly separated from the contents that the empirical agents can endorse (inclinations, interests). To my understanding, in light of this strict limit two questions arise that Kantians should address: on the hand, does Kant's view on morality includes any content and therefore can it tell us that we ought to do this or that, and not just tell us to respect the moral law? On the other hand, whether and to what extent in Kant's moral philosophy the empirical self can be motivated to act according to the demands of morality (in this Chapter I focus on giving an answer to the former question).[2] These questions are different but they are related. This is because it is reasonable to expect that morality has to include contents such as particular inclinations and interests to motivate particular agents. Of course, it is possible to maintain that good moral reasons might be enough to motivate empirical agents to perform moral actions. In this respect, Christine Korsgaard claims that 'the reasons why an action is right and why you do it are the same' (Korsgaard, 1986: 10; in Forst, 2007: 23) and in the same Kantian spirit Rainer Forst asserts that:

---

[1] For a detailed discussion on this issue see Allison 2004.

[2] Let us say that the former concerns to a question of 'moral phenomenology' or how morality can include proper content. And the latter refers to an issue of 'moral motivation' or how morality can motivate real people. They are related questions but they are different. I assume that to answer the latter first we have to give a complete answer to the former. In this Chapter I aim to provide that response.

> The reasons that normatively speak for the action also speak for it in a subjective motivational way, since within the given normative context no justifiable counterreasons have to be found. (2007: 23-24)

Nevertheless, that position might still be not convincing and leave the door open for the objection that the strict distinction between moral contents and particular contents supposes that Kant (and also the Kantian tradition) cannot tell moral agents what they ought to do. In this regard, the issue is why should the empirical *phenomenal* self accept and obey laws imposed by the rational *noumenal* self? Let us remember that in the Kantian view of morality, the rational *noumenal* self discards particular contents.

Habermas endorses a Kantian pragmatism (T&J, 83-130; Baynes, 2016: 82-97), which among other things discards the two realms distinction. In this way, in his theory the strict distinction between *noumenal* and *phenomenal* selves is abandoned and what takes its place is an intersubjective procedure of moral discourse. Additionally, in Discourse Ethics Habermas as built the view that the development of post-conventional moralities depend on processes of socialization in which the moral agents have learnt to respect others and their capability to perform validity claims in discourse. Henceforth, not only the intrinsic power of moral reasons from above, but also both the rejection of the two realms doctrine and the consideration of the relevance of processes of socialization and dispositions to act on good reasons from below, ensure that the agents are going to act according to the rules of morality. At first appearance, if Habermas can make good on his claim that moral reasons and adequate socialization, suffices to show how agents can respect the contents of morality then Discourse Ethics can successfully rebut Hegel's critique of Kant.

However, in Habermas's version of the universalizability of norms he claims that norms embody universalizable interests and that these are distinct from particular interests. He also claims that moral reasons are impartial, and hence that they are not agent-relative, and seems to imply that no agent-relative reasons whatsoever have any moral validity. (IO, 7, 43) In Discourse Ethics, the principle of universalization (U) works as a criterion of moral validity by distinguishing between generalizable and particular interests, and filtering out the latter from the process of reaching agreement in moral discourse. (MCCA, 64-65) Therefore, the question still arises of whether, given his conception of the impartiality of

morality, he has convincingly answered Hegel's later criticism of Kant that targeted the will as a tester of maxims; this in light of what Habermas proposes.

In this Chapter, I show how Habermas could respond to this objection by means of an analysis of Mead's concept of ideal role taking which has been incorporated in the reconstruction of the moral point of view offered by Discourse Ethics. That said, my reading supposes that Habermas needs to slacken his notion of impartiality, leaving room for generalizable interests (moral contents) that can be at the same time particular interests (particular contents), that is agent-relative interests. Yet to my mind, this is actually what takes place with his concept of ideal role taking.

Now, there is an additional issue in Habermas's theory that has to be addressed. It is connected with Hegel's strategy of answering the problems he found in Kant: the dialectical relationship that Hegel proposed between *Moralität and Sittlichkeit*. According to this perspective, the break between universality and particularity could be solved by blurring the distinction between moral norms and ethical values. Then, the argument goes, moral norms (moral contents) can be at the same time ethical values (particular contents).[3] Henceforth, throughout *Sittlichkeit* the individual can be at home with himself and he can find his own *particular* way of being universal. (Pinkard, 1999: 227) In this regard, the point of contention is whether is it necessary or not for Habermas in Discourse Ethics to blur the distinction between moral norms and values. I argue that there is an additional point in Discourse Ethics to solve the problem, which is neater and does not involve blurring the distinction. It stands in an interpretation of Kant's formula of humanity of the categorical imperative to respect others as not only means but always also as ends in themselves. (4: 429) In this formulation, one particular content, the interest of an agent to be respected as an end in herself, coincides with the universal content of everybody to be also respected as end in themselves.

In order to develop these issues, I divide this Chapter into five parts. In the first, I examine Hegel's objections to Kant. It is true that Hegel praises Kant's concept of morality and autonomy. (see §135) However, according to Hegel, Kant's account of pure practical reason needs to be supplemented with a consideration of the dimension of ethical life

---

[3] Habermas himself endorses the view that ethical values have a particularistic nature opposed to the universalistic scope of moral norms. (JA, 1-17) Ethical values pertain to a particular way of life of one individual and/or of her community and they cannot claim universal recognition. See section 5 of this Chapter.

[*Sittlichkeit*]. In Hegel's view this absence in Kant supposes an *empty formalism* and this resulted in a barrage of objections against Kant's reconstruction of the categorical imperative. I show that most of them can be successfully rebutted. Nonetheless, a later objection, which does not target the categorical imperative but the conception of the will as a tester of maxims, poses a real challenge to the Kantian position (1);

In the second section, I examine Habermas's response regarding the question of whether and to what extent Hegel's critique applies to Kant and to Habermas's moral theory. On some points Habermas argues that Hegel's criticisms of Kant are mistaken. On others, he argues that Hegel has good arguments that challenge Kant's moral philosophy. That said, Habermas claims that Discourse Ethics incorporates elements that fortify its position against these Hegelian criticisms. In this section, I reconstruct these argument and I begin to delineate the contours of a Hegelian objection that can apply to Kant and to Habermas as well (2);

In the third section, emerges in sharp relief the Hegelian criticism of the will as a tester of maxims this time applying to Discourse Ethics. In this regard, I examine Habermas's notion of the universalizability of maxims, and I discuss whether the strict conditions that he imposes to the contents that can claim moral worth leaves room to the inclusion of particular contents. In my interpretation, if that is not the case, then it is difficult to argue that Habermas has closed the gap between the moral will and the particular will (3);

In the fourth section, I develop a Habermasian response to Hegel's later criticisms of the will as a tester of maxim. The argument here is that the process of ideal role taking, properly understood, already contains a potential answer to the Hegelian objection. However, this demands that Habermas slacken the strict distinction that he has proposed between agent-relative and agent-neutral reasons (generalizable interests and particular interests). I maintain that this modification is coherent with Mead's concept of ideal role taking which has been built into Discourse Ethics (4);

Finally, I discuss a further Hegelian challenge, this time to the distinction that Habermas draws between morality and ethics. I claim that there are several reasons why Habermas should not blur this distinction. Additionally, I maintain that Habermas has a second argument to solve the Hegelian challenge, this time relying on a Kantian insight: the

recognition of persons as ends in themselves. Discourse Ethics incorporates this element by means of the idea of relations of mutual recognition and understanding. In this regard, Habermas can still endorse the distinction between moral norms and ethical values, and at the same time answer the Hegelian challenge (5).

1. Hegel's critique of Kant's Moral Philosophy

In this section, I examine Hegel's critique of Kant's moral philosophy, from his early essay on *Natural Law* (1802), his *Phenomenology of Spirit* from the middle period in Jena (1807) and the later formulation in *Philosophy of Right*. (1820) In this regard, Habermas rightly states that 'The criticisms Hegel levelled against Kant as a moral philosopher are many'. (MCCA, 195) In this section, I focus mainly in two sets of criticisms: the first set concerns the *empty formalism* of the moral point of view; the second the criticism of the will as a tester of maxims.[4]

Now, it is important to notice that according to a quite influential interpretation of Hegel, he can be regarded as a *post-Kantian* philosopher. This could just mean a philosopher after Kant, but also a *kind of Kantian philosopher after Kant*. In light of the analytical reading of Hegel's critique of Kant offered by Robert Pippin and Terry Pinkard, I claim that Hegel is of the latter type. (Pippin, 1989; 1991; Pinkard, 1994; 1999: 222) This is because Hegel did not reject Kant's concept of pure practical reason and autonomy in favour of something else. Rather, Hegel in fact is best seen as extending Kant's rationalist morality by critiquing it and supplementing it, but not rejecting it.

Hegel agrees with Kant that the moral will is autonomous. (See Freyenhagen, 2011: 164) For example in §133 he states, 'In doing my duty, I am with myself [*bei mir selbst*] and free. The merit and exalted viewpoint of Kant's moral philosophy are that it has emphasized this significance of duty'. In this way, Hegel argues that Kant has to be praised because in his moral philosophy, the knowledge of the will first gained a firm foundation and point of

---

[4] For some Kantians, the charge of *empty formalism* is misplaced. (See O'Neill, 1989: xi) In this regard, I do not take for granted that Kant's moral philosophy is *empty* and *formal*. (MCCA, 204-205. See also Freyenhagen, 2011 for a detailed discussion of the Empty Formalism Objection) Nevertheless, as I am going to show, Hegel's later criticisms of Kant of the will as a tester of maxims cannot be consistently rebutted if one remains inside Kant's philosophy. This is because Kant's notion of pure practical reason makes a metaphysical claim that suppose a strict distinction between the requirements of morality and the features of real agents. This claim is the doctrine of the two realms between the *noumenal* self and the *phenomenal* self. (For discussions of this criticism see MCCA, 207-208; See also Benhabib, 1986; and Finlayson, 1999)

departure through the thought of its infinite autonomy. (see §135) Moreover, Hegel's critique of Kant does not imply that the former pleads for a return to an ethical theory, like the one advanced by, say, David Hume in *A Treatise of Human Nature.* (1739-1740) That is to say, it does not amount to the claim that the will could only be shaped by particular inclinations, desires and feelings.[5] In this respect, Hegel asserts:

> What constitutes right and duty, as the rationality in and for itself of the will's determinations, is essentially neither the *particular* property of an individual, nor is its *form* that of feeling [*Empfindung*] or any other individual – i.e. sensuous – kind of knowledge, but essentially that of *universal determinations of thought*, i.e. the form of *laws* and *principles*. (§137)

For all these reasons, Hegel could be considered as a *post-Kantian* philosopher. However, in the *Philosophy of Right,* specifically in §135 he also claims 'to cling on to a merely moral point of view without making the transition to the concept of ethics reduces this gain to an *empty formalism,* and moral science to an empty rhetoric of *duty for duty's sake'*. In other words, the moral standpoint in Kant's philosophy does not recognize the dimension of ethical life and that is why it is charged with the *empty formalism* objection. Therefore, it is a theory based on formal reasoning that is divorced from the real life and its values, practices and institutions. In this regard, the relationship between *Moralität* and *Sittlichkeit* is going to be Hegel's main focus of the problem and also of his proposal of solution of the *empty formalism* in Kant's moral philosophy. (Pinkard, 1999: 227) To my mind, if this objection hits home, Kant's concept of pure practical reason and autonomy are put into question and they become problematic. According to Hegel, the philosophy of Kant needs to be supplemented because the former asserts that the moral theory of the latter cannot claim jurisdiction over the substantive problems of daily life. Consequently, the Kantian moral will does not seem to have concrete content.

Moreover, there is a second criticism which is connected with the charge of *empty formalism*, although, its focus now it is not the Kantian reconstruction of the categorical imperative. Rather, its target is the concept of the will as a tester of maxims. Specifically, this criticism targets the distinction that Kant proposes between the *noumenal* will and the *phenomenal* will.

---

[5] In this regard, Hume famously wrote that 'Reason is, and ought only to be the slave of the passions and can never pretend to any other office than to serve and obey them'. (1975: 415) For a contemporary account of Hume's moral theory see Bernard Williams. (1981) Of course, with important modifications of Hume's view.

This supposes a break between the moral will – which is the articulation and imposition of pure reason – and the empirical will – which is the real individual with her actual desires and interests. Due to this, it is not clear whether and to what extent the rational *noumenal* self allows the inclusion of particular contents.

In what follows, I will sketch Hegel's criticisms from the early formulations up to its later expression in the *Philosophy of Right*. I will maintain that there are good Kantian arguments to rebut the objection to the *empty formalism* of the categorical imperative, insofar this amounts to the charge that it cannot give any content. Nevertheless, Kant's metaphysics of the two realms between the *noumenal* self and the *phenomenal* self makes difficult to refute the charge against the will as a tester of maxims. If this criticism is correct, then the moral subject cannot achieve autonomy, because she cannot have her *own* reasons to be bind by pure practical reason. If pure practical reason discards the inclusion of particular contents then morality does not seem to be able to tell moral agents what they ought to do.

<center>***</center>

The objection to formalism starts with the claim that the categorical imperative[6] is a logical test of the universalizability which any maxim can pass. Thus, as Hegel points out in the *Phenomenology*, Kant's reconstruction of the moral point of view cannot perform its function as a proper test of universalizability. As he indicates, 'The criterion of law which Reason possesses within itself, fits every case equally well, and is thus in fact no criterion at all'. (1977a: 259) Since the moral standpoint is not a criterion at all, then the ground for autonomous agency is futile. On this point one needs to remember the basics of Kant's moral philosophy in which there has to be an equivalence or identity between the autonomous (moral) will and the particular will or as he asserts 'a free will and a will under moral laws are one and the same' (4: 447) – See footnote 1 in Chapter One. And this identity is problematic too, because if only the moral will is autonomous that might imply that the empirical will is heteronomous or unfree. This cannot be right because then agents would not be responsible for their empirical actions. Kant overcomes this issue drawing the distinction between freedom of choice [*Willkür*] and free will [*Wille*]. An agent might seem to be heteronomous insofar as its freedom of choice [*Willkür*] is governed by morality, but her autonomy depends on its free will [*Wille*] which is the legislative function

---

[6] Kant's first formulation of the Categorical Imperative states, 'Act only according to that maxim through which you can at the same time will that it become a universal law'. (4:421)

that Kant equates with pure practical reason. (See Allison, 1990: 129-36) Moreover, '*Wille* is the source of the laws that confront the human *Willkür* as imperatives [and] it seems clear that this must include both the categorical and hypothetical imperatives'. (Allison, 1990: 130) Hence, for Kant the categorical imperative is the self-given law of an autonomous will. Therefore, if this test is not criterion at all, then there is not a basis for morality and for autonomy.

In despite of this, it is possible to answer satisfactorily this charge. It is not difficult to find maxims that fail the test as Kant himself does in the *Groundwork of the Metaphysics of Morals*. (See also O'Neill, 1989: 158) There, Kant considers an example in which a person needs money and at the same time knows that she will not be able to pay it back. The maxim of the action for that person could be, 'when I believe myself to be in need of money I shall borrow money, and promise to repay it. Even though, I know that it will never happen'. (4:422) The question now is what would happen if this maxim becomes a universal law. Kant asserts that if this were then the practice of promise making itself would cease to exist. (Snare, 1990: 40) Hence, it would not be possible even to break a promise. Thus, the maxim fails as it contains what is often called a 'contradiction in conception'. (Korsgaard, 1985: 2, O'Neill, 1989: 89) Consequently, it is wrong to claim that any maxim can be made to past the test of universalizability. Moreover, the negation of an unsuccessful maxim expresses a strict duty. Following the example we are examining, the negation of the maxim 'I should make false promises' results in the strict duty 'I should not make false promises'.

However, there are a lot of counterexamples to the view that we have a strict duty not to act on whatever maxims fail the test of Universalizability. For example 'Always open doors for other people'. This is a plausible example of a maxim, however 'given the fact that two people cannot open the same door, the maxim clearly fails the test of universalizability'. (Finlayson, 1999: 38) Nevertheless, Onora O'Neill introduces a scope restriction on what can count as a candidate maxim. She distinguishes between 'underlying intentions' and 'ancillary intentions'. The former offers cases of morally relevant maxims – for example, 'I should not make a false promise' –, the latter refers to cases than are not relevant in the moral domain – for example, 'I should open doors for other people' or 'I have to offer tea to my guests. (O' Neill, 1989: 145-46) Ottfried Höffe offers a defence of Kant in this point as well. He draws a distinction between general maxims – those that pertain to the moral domain – and rules or precepts of action – that refer to one's way to order her particular

life. (Höffe, 1994: 149-151) From these two arguments I draw the conclusion that Kant still can rebut Hegel's criticism.

Hegel has a second more careful objection to the *empty formalism* of the categorical imperative, developed in the essay on *Natural Law* and in the *Phenomenology*. In this criticism, he argues that the maxims can succeed the test of universalizability because in Kant's moral theory there is presupposed the existence of substantive moral values that shape our social life. This means that any maxim is successful or failed depending on its coherence with certain institutions and the values and practices that underlie them. This challenges Kant's account of morality, because the doctrine of the two realms marks a clear distinction between the conditions of possibility of an autonomous will, which is *noumenal*, and the *phenomenal* world of social institutions. In other words, according to Hegel, Kant should have been aware that in his theory substantive contents were incorporated. In this criticism, Hegel discusses Kant's argument in the *Critique of Practical Reason* concerning the case of a person who wonders whether she should or not return a deposit 'the owner of which has died and left no record of it'. (5:27) The maxim that could guide this sort of action would be that it is morally right to try to 'increase my wealth by every safe means' (5:27), for example keeping a deposit that no one will claim as it is unrecorded and the owner passed away. Now, this person may wonder whether that maxim could also hold as a universal law. Kant argues that when that person begins to wonder whether that maxim could become an universal law she will realize that 'I at once become aware that such principle, as a law, would annihilate itself since it would bring it about that there would be no deposits at all'. (5:27)

According to Hegel, the maxim that we are discussing fails the test of the categorical imperative, not because it might have a self-contradictory character, but because it goes against substantive moral values and institutions, such as property. Thus, one could imagine a world in which those institutions are not present and therefore the maxim would succeed the test. In this way, Kant's argument shows that a system without deposits is contradicted by a system with deposits. As Hegel asserts, 'Property, simply as such, does not contradict itself it is an *isolated* determinateness, or is posited as merely self identical. Non property, the non ownership of things, or a common ownership of goods, is just as little self-contradictory'. (1977a: 258) Hegel's claim is that one element that plays a role when the moral worth of any maxim is assessed is the social context in which the agents

are placed. However, due to the theory of the two realms, in Kant's moral philosophy the deontological worth of any maxim has to be independent of the context and be based only in the autonomy of a subject that is rational and gives itself universal laws. The autonomy of morality would be undermined if it depends on the institutional setting in which the moral assessment takes place.

The conclusion that one has here is that Kant's conception of autonomy is failed. If Hegel's is right in this criticism, then it is impossible to conceive autonomy beyond the realm of the institutions and practices of a particular ethical life [*Sittlichkeit*]. Thus, what confers moral validity is not the logical structure of maxims (pure reason) but the coherence of those maxims and the underlying values, institutions and practices of a particular form of life. Therefore, pure practical reason and autonomy are always already social. However, there are arguments to counter Hegel's challenge in this respect, but to a certain extent the answer relies on a Hegelian insight.

Hegel has an interesting point as he already has the beginnings of a good argument that is improved in the *Philosophy of Right.* (1820) Nonetheless, as in the case of the false promise the problem with the deposit example is that it implies a 'contradiction in conception'. The contradiction does not depend on envisioning a society in which there are no deposits; the problem arises when someone wills a world in which, the appropriation of a deposit becomes a universal law and at the same time that person also wills the existence of a world, in which there are deposits. (Finlayson, 1999: 37) As a result, the contradiction is not generated by the reasons Hegel argues for. At this point, in his defence of Kant Christine Korsgaard asserts that the contradiction 'is generated when the agent tries to will his maxim and the universalization of his maxim at the same time, or tries to will it for a system of which he is to be a part'. (Korsgaard, 1985: 13)

Consequently, Korsgaard's defence agrees with Hegel at least in one respect: moral agency has to be understood in a *phenomenal* social context. The contradiction in conception is caused not because the maxim is contradictory with a set of institutions. The contradiction depends on the will of a particular subject that wills her maxim and at the same time wills institutions that are contradictory with that maxim. In this respect, Korsgaard's defence of Kant from Hegel's criticism is both based on an Hegelian insight and it already advances the later objection to the will as a tester of maxims which is examined in what follows.

In the later objection which is developed in the *Philosophy of Right*, specifically in the discussion of 'Moralität' (3rd Section, 'The Good and Conscience'), Hegel asserts:

> From this point of view, no immanent doctrine of duties [Pflichtenlehre] is possible. One might indeed bring in material from outside and thereby arrive at particular duties, but it is impossible to make the transition to the determination of particular duties from the above determination of duty as absence of contradiction, as formal correspondence with itself, which is no different from the specification of abstract indeterminacy; and even if such a particular content for action is taken into consideration, there is no criterion within the principle for deciding whether or not this content is a duty. On the contrary, it is possible to justify any wrong or immoral mode of action by this means. — Kant's further form — the capacity of an action to be envisaged as universal maxim — does yield a more concrete representation [Vorstellung] of the situation in question, but it does not in itself [für sich]. (Hegel, 1991: §135)

This passage has several elements that need to be examined. Here, Hegel claims that to cling on to the moral point of view without making the transition to the concept of ethics has the effect that 'no immanent theory of duties is possible'. (§135) This means that Kant cannot try to deduce a whole repertoire of maxims just from the categorical imperative procedure and from logic. Nonetheless, it is not difficult to respond to this objection. In the *Groundwork* and in the *Second Critique* Kant's principal task is to provide a justification for the moral law. It is in the *Metaphysics of Morals* (the 'Doctrine of Virtue') where Kant provides a doctrine of duties in the sense of a repertoire of concrete duties.

Moreover, Hegel maintains his early criticism that the categorical imperative is not a proper criterion to assess the universalizability of the maxims. Therefore, even wrong and immoral contents could successfully pass the test for consistency, and the mere lack of contradictoriness of a maxim does not yield enough information about what morality should demand. (See Freyenhagen, 2011: 165) Henceforth, it is the wrong kind of normativity as long as formal rationality does not rule out moral wrongness. This is because it is not possible to derive the normativity of morality from the normativity of rationality. As I have shown, however, this objection could be rebutted by Kant because there are

maxims that contain a 'contradiction in conception' and consequently that actually fail the test. For example, the maxim 'I will make lying promises when it achieves something I want', contradicts itself once made into a universal law. Therefore, the categorical imperative is a proper criterion to evaluate the validity of norms that as long as they fail the test they are morally prohibited.

Notwithstanding, it is still a worry here because it seems that there are not many of these maxims which actually fail the categorical imperative. Therefore, there is some, but not enough moral content to regulate our daily life. Also, there is no strict correlation between maxims that fail the categorical imperative test and strict duties – as I have shown the maxim 'Always open doors for other people' fails the test and there is surely no duty not to do it –. Certainly, O' Neill and Höffe might be right when they propose scope restrictions on what can count as 'morally relevant' contents and the aforementioned example should be discarded. Nonetheless, bearing in mind the issue that I am analysing now these restrictions are equally unrewarding: if O' Neill's and Höffe's arguments are sound then even less content can be susceptible of being tested by the categorical imperative. This discussion allows sharpening the reply to Hegel's criticism of the categorical imperative: the Kantian test is a proper criterion to reject norms that are not morally valid, but still Kant cannot guarantee that his moral philosophy includes enough moral content.

There is a further objection that only emerges in sharp relief in the later objection, namely, the criticism of the will as a tester of maxims. Roughly speaking, Hegel claims that Kant's philosophy supposes a gap between the *noumenal* self (moral will) and the *phenomenal* self (empirical will). The first refers to the *noumenal* rational being who follows the rules of pure practical reason, and the second is the *phenomenal* real agent with her inclinations, interests, and located in a particular space and time. Hegel arrives at this conclusion in the previous passage when he states that Kant's conception of the moral will is characterized by 'formal correspondence with itself, which is no different from the specification of abstract indeterminacy'. The moral will (*noumenal* self) is both only identical with itself and only incorporate maxims that have an abstract form. Henceforth, particular contents and interests remain alien to the moral will and Kant's concept of autonomy is incomplete. In other words, the empirical will, which has substantive contents and interests, however it be made up, cannot be united with the moral will. Thus, Kant's moral philosophy cannot give an answer to the question of how moral insights can be realized in practice. (MCCA, 196)

This is because, as I have shown in the introduction of this Chapter, due to this break, Kant's philosophy cannot explain how morality will motivate real agents if the contents that they might endorse are filtered out, due to the bounds that the moral point of view imposes. If the moral law must override or abstract from all particular interests there remains no reason for the empirical will to act in accordance with it. The moral self splits into a noumenal self that makes demands and an empirical self that is supposed to obey.

In this regard, one of the tasks, 'arguably the most important one' (Finlayson, 1999: 41), of Hegel's philosophy of the objective spirit, is to show how the particular interests acquire universal form or how the empirical will acquires a moral form. In §153 of the *Philosophy of Right* Hegel asserts 'the right of individuals to their subjective determination to freedom is fulfilled in so far as they belong to ethical actuality'. This means that through Sittlichkeit, freedom (or the moral will) and the right of the individual to be free for her own reasons (or the particular will) can be joined together. Later in that paragraph, Hegel explicitly argues that this can be achieved in a state with good laws:

> When a father asked him for advice about the best way of educating his son in ethical matters, a Pythagorean replied: 'Make him the citizen of a state of good laws'.

In summary, the point is that Hegel's early formulation of his criticism of Kant's empty formalism can be countered. However, the later and more mature argument shows a strong case against Kant's moral philosophy and it represents a challenge to the notions of pure practical reason and autonomy. The impossibility of reconciling the moral will and the particular will poses a clear difficulty for Kant's concept of moral autonomy. If Hegel is right, then in Kant's moral philosophy the empirical subjects are subjugated to a universal moral will. Therefore, they do not have their own reasons to obey the moral law and it would not be the moral law in Kant's sense either, as in his account the will has to be autonomous to be moral. A first step to give an answer to this Hegelian challenge consists in discarding the theory of the two realms. This is something that Hegel already did in his philosophy, through the dialectical relationship that he proposed between *Moralität* and *Sittlichkeit*. By means of this relationship, in the *Philosophy of Right* it was possible to conceive the reconciliation between the *noumenal* self and the *phenomenal* self. According to Hegel the projects and contents of the particular agent were embedded in its particular form of life.

Consequently, as far as *Moralität* was embedded in *Sittlichkeit,* then the real agent could be at home with herself and be moral at the same time.

Habermas too discards the doctrine of the two realms, however not by blurring the distinction between *Moralität* and *Sittlichkeit*, between moral norms and evaluative statements. That said, however, Habermas has a good argument to rebut this Hegelian criticisms, by means of the concept of ideal role taking. Nevertheless, before I move into that issue, I need to discuss Habermas's position concerning Hegel's critique of Kant.

2. Habermas on Hegel's critique of Kant

In what follows, I examine Habermas's account of Hegel's critique of Kant and I begin to delineate the Hegelian criticism of the will as a tester of maxims this time applying to Discourse Ethics. In the previous section, I argued that this criticism actually cannot be rebutted by Kant. According to Hegel, in Kant's moral philosophy there is a separation of universal form and particular content which supposes a break between the universal will and the particular will. Habermas replaces Kant's rigid division of universal form and particular content for the distinction between universal and particular content. Nevertheless, the result could be equally problematic. This is because, if as in Kant, in Discourse Ethics the universal will and the particular will cannot be joined together, then also Habermas theory can be charged of being formal. Consequently, if the objection hits the target, then it seems that Habermas is not able to give an account of how particular content might bleed into moral discourse. This Hegelian criticism of Habermas's moral theory is developed mainly on section 3. In response to this objection, in section 4, I show that it can be successfully rebutted by means of an analysis of the concept of ideal role taking. However, in order to do this Habermas has to slightly modify or slacken some components of his theory. First, though, more preparatory work needs to be made in order to set the stage to address that issue.

In Habermas's view, on some points Kant's moral philosophy is simply immune to Hegel's criticisms. On others, it is vulnerable to Hegel's criticisms. Now, by contrast, the neo-Kantianism that Habermas develops in Discourse Ethics has the resources to rebut these Hegelian criticisms. This is because Habermas has incorporated a modified version of

Kant's notion of pure practical reason and autonomy which fortifies and shields his theory from the contemporary version of Hegel's critique of Kant.

In his article, 'Morality and Ethical Life: Does Hegel's Critique of Kant Apply to Discourse Ethics?' Habermas gives a response to this question and also he examines the pertinence of Hegel's critique of Kant's moral philosophy. The criticisms that Habermas considers are (MCCA, 204-210): i) the formalism of the moral principle; ii) the abstract universalism; iii) the objection to the impotence of the 'Ought'; iv) the subject of 'Virtue and the Way of the World'. In what follows, I examine Habermas's stance on the first three (1, 2, 3) criticisms.[7] At the end of this section, I outline in a preliminary way what I think it is the centre of a Hegelian criticism that can challenge Habermas's moral theory (4).

1. Concerning the first criticism, that refers to the formalism of the moral principle Habermas argues that: 'Neither Kantian ethics nor discourse ethics lays itself open to the charge that since it defines the moral principle in formal or procedural terms, it can only make tautological statements about morality' (MCCA, 204). Kant's Categorical Imperative and Habermas's principle of universalization (U) are formal but not empty. (Finlayson, 1999: 34) In this respect, Hegel is wrong when he argues that Kant's formulation of the moral point of view just postulate logical and semantic consistency and nothing else. For Habermas, these principles refer to a substantive reconstruction of the inner structures of practical reasoning and consequently they are not purely formal. As we have seen, if at least one maxim passes the categorical imperative, then the test is not *empty* and merely *formal* (See section 1). There is another worry here because Habermas is supposed not to say anything about what the norms are that pass the test of (U), he leaves that up to participants themselves. This may mean that Habermas's principle of universalization does not include content. Nevertheless, I think that this is an advantage because the participants themselves are the ones that are in the best position to regulate their conflicts of action by means of moral discourse.

Notwithstanding, it has been argued that the principle of universalization imposes very demanding conditions to what can count as a moral norms. The reason that explains why Discourse Ethics sets these strict limits is because (U) aims to cohere with an historical context determined by a plurality of forms of life, in which moral norms appeal to be

---

[7] For the sake of my argument, I do not consider Habermas's fourth reply.

binding universally. Due to this, very few norms pass this test. Habermas admits this, for example when in JA he says that:

> To be sure, the sphere of questions that can be answered rationally from the moral point of view shrinks in the course of the development toward multiculturalism within particular societies and toward a world society at the international level. (JA, 91; See also IO, 3-46)

Baynes addresses the issue of the burden that is imposed to the norms to reach the status that is demanded by (U). Fundamentally, he refers to factical constraints of practical discourses – i.e., space and time. As a consequence of this, there would be very few norms that can be identified as the outcome of a practical discourse. In this way, Baynes correctly claims that 'At most, perhaps, basic moral norms will come to resemble a core set of basic human rights'. (2016: 104) Habermas does not reject this reading. Now, if very few norms pass the test of (U), can morality play the central social role that Habermas assigns to it? Let us remember that in Discourse Ethics and in TCA Habermas argues that morality is the basis of social integration.

Habermas appears to assume that, it is not a problem if very few norms pass the test of (U) in the context of a pluralism of forms of life and progressively greater individualization of life projects. (JA, 90-91, OI, 41) I agree with this position, because a basic set of moral norms can successfully regulate the essential conflicts of our lives in common. Particular issues can be regulated by specifications given by contextual features of the conflicts at hand – i.e., the scope of those who are affected, cultural particularities and so on. That said, however, these norms need to recognize and respect a core set of basic norms morally justified. Perhaps, this position is more tenable once Habermas has specified the role played by ethical discourses and by democratically legitimate law, because they take up the burden of social integration. Nevertheless, morality in Habermas never has lost his role for social integration and in this respect positive Law and morality have a complementary relation. (IO, 256-258)

Habermas has another argument to show that the principle of universalization of Discourse Ethics is not empty and purely formal. Moral norms emerge as normative and morally right answers to the conflicts of action that 'grow out of everyday life'. (MCCA,

204) Hence, moral discourse takes place as a way in which social agents solve conflicts in everyday (lifeworld) contexts. In other words, these disputes and the norms that regulate them are products of our social life and not the speculative inventions of the philosophers. Habermas's claim here is that moral content arises from the lifeworld in the form of conflicts of interest and the norms that allow their discursive resolution.

However, this response is a little misleading, because even if the norms (maxims) 'grow out of everyday life', Hegel's claim is that the categorical imperative fails as a criterion to evaluate the maxims, because it is a test to evaluate the logical and semantic consistency of the maxims and nothing else. The point is not really whether the maxims come from the everyday life or the philosophers create them. Rather, Hegel's criticism asserts that the categorical imperative is not an adequate criterion to evaluate the universalizability of the maxims, no matter if they come from the *noumenal* sphere of reasons or from the *phenomenal* realm of the everyday life. The response that Habermas is giving here is not relevant to answer this criticism but it is useful to respond to the charge that Kant's cannot provide an *immanent doctrine of duties*.[8] If at the end Kant can show that the maxims come from our *phenomenal* social life, then he does not need to provide a *doctrine of duties* at all.

Habermas also claims that it is not correct to state that the categorical imperative can only make tautological statements. I agree with this position, but not for the reasons that Habermas gives. In this respect, it is enough to show that there are maxims that can fail the test, and for Kant they are crucial because they yield strict negative duties not to do them. In the case of Discourse Ethics, it is the norms that pass the test of U that are important because they tell us what we ought to do. Nevertheless, according to Habermas there is a sense in which Hegel's charge of formalism could apply to any procedural ethics. These sort of moral theories need to segregate 'from among the general mass of practical issues precisely those that lend themselves to rational debate'. (MCCA, 204) As a result, there is a strict distinction between normative statements and evaluative statements, norms and values. On the one hand, norms are the objects of morality because they are universalizable. On the other hand, values refer to particular ways to orient the existence of a community or a person's particular projects towards a specific conception of the good

---

[8] Freyenhagen (2011) discusses three Kantian replies to this charge. The first concerns an interpretation in which Kant is a moral realist. This means that he does not need to provide a doctrine of duties because they come from the social world. The second Kantian reply argues that Kant provides this doctrine in the *Metaphysics of Morals*. The third argues that the second formulation of the categorical imperative (the formula of humanity) properly interpreted allows a doctrine of duties.

life.[9] According to Hegel, this distinction results in the abstraction of moral norms, in the sense that they become detached from the good life and 'that made impossible for morality to claim jurisdiction over the substantive problems of daily life'. (MCCA, 204)

Habermas argues that Hegel has an interesting point here. However, the former asserts that Human Rights are an example in which normative statements have at the same time an evaluative character. Human rights can be thought dialectically as a group of moral norms that both convey generalizable interests and also are ethical values. Therefore, they are part of the dimension of *Moralität* and are part of the *Sittlichkeit* of modern life. As Habermas asserts:

> Human rights obviously embody generalizable interests. As such they can be morally grounded in terms of what all could will. And yet nobody would argue that these rights, which represent the moral substance of our legal system, are irrelevant for the ethics (*Sittlichkeit*) of modern life. (MCCA, 205)

Then, human rights are both moral entities and ethical entities. Moreover, I read this passage in the sense that in Discourse Ethics moral norms (as embodied in human rights) are not so abstract that they have no relation to the legal and ethical framework of modern society. Therefore, they are not completely unrelated to the substance of our social life. So, in this way Habermas can rebut Hegel's critique that Kantian morality cannot claim 'jurisdiction over the substantive problems of daily life'. (MCCA, 204)

2. Regarding the objection to the abstract universalism, for Habermas neither Kant nor Discourse Ethics can be charged with the criticism that 'the generalizability of norms necessarily leads to the neglect, if not the repression, of existing conditions and interests in a pluralistic society'. (MCCA, 205) Modern societies have more differentiated interests and value orientations; therefore moral norms necessarily become more general and abstract. However, this does not mean that these interests and values are discarded, at least not by Habermas's moral philosophy. Nevertheless, Habermas's position here could be challenged

---

[9] This insight is part of the *Eudemonist* tradition since Aristotle. (Aristotle, 2009) In this perspective, the aim is not to address the questions concerning the deontological worth of moral norms. Rather, the issue is how to guide the human existence towards the good life which certainly appeals to substantive contents. In the *Nicomaquean Ethics* Aristotle argues that the *Telos* or function of the human being is to live rationally well. Moreover, to achieve this human function it is necessary for the agent to lead a life according to the virtues of ethics.

because there is a specification of the sort of interests that can be part of a rationally motivated consensus. He asserts that these interests need to be universalizable and agent-neutral. (IO, 7, 43)

Thus, the same problem arises here concerning the later Hegel's criticism of Kant. Does Habermas's moral theory close the gap between the moral will (universalizable interests) and the empirical will (particular interests)? There are many cases in which a particular agent has an interest that is not universalizable. For instance, the free rider wants to get a benefit of the collective and at the same time she does not contribute to the creation of those goods. If Discourse Ethics cannot answer this case, this seems to be detrimental to its position.[10] This point is going to be developed further in the next section and here is where I think Hegel's objections still have purchase on Discourse Ethics. However, in section 4 I argue that Habermas's moral philosophy has arguments to rebut this objection.

Hegel's objection to abstract universalism sometimes takes the form of a criticism of rigorism. As he construes the charge, procedural ethics fails to take account of the consequences and side effects 'that may flow from the generalized observance of a justified norm'. (MCCA, 206) In his lecture *Politics as Vocation* (1919) Max Weber was moved by this objection to counterpoise an ethics of responsibility for consequences and side effects against Kant's ethics of conviction. For Habermas, this criticism applies to Kant but not to Discourse Ethics, since the latter breaks with Kant's idealism and monologism (MCCA, 206). The procedure of Discourse Ethics is formulated in the principle of universalization (U) that explicitly requires sensitivity 'to the results and consequences of the general observance of a norm for every individual'. (MCCA, 206) In this way, Habermas builds a consequentialist consideration into the heart of his deontological moral theory.[11]

According to Habermas, Hegel is right in another respect too: 'Moral theories of the Kantian type are specialized. They focus on questions of *justification*, leaving questions of

---

[10] However, from early on Habermas has traced a distinction between instrumental, strategic and communicative modes of action. (CES, TCA) The case of the free rider refers to the instrumental or strategic type of action.

[11] I think that it is possible to connect the consideration of consequences and the contents that are relevant for the agents. If that connection is correct then Habermas has another argument to respond to Hegel's later charge. Consequences may be the contents of the empirical will. In other words, they may be the particular reasons that explain whether the agent follows or rejects moral norms. Nevertheless, again one might argue that these considerations need to recognize contents that have to be universalizable and agent-neutral. Therefore, although Discourse Ethics incorporates the consideration of consequences, does still Habermas's position can close the gap between the moral and empirical will?

application unanswered'. (MCCA, 206) Due to the abstractions that any procedural ethics have to make to justify norms the questions concerning the application of these norms are dismissed. In this regard, the application of general norms to particular cases seems to require an Aristotelian faculty of *prudence.* This would tend to weaken the universalistic claim of pure practical reason because it would be tied to a particular form of life. For Habermas, his moral philosophy does not need to go back to a philosophical position prior to Kant, and in Discourse Ethics 'even in the prudent application of norms, principles of practical reason take effect'. (MCCA, 206-207) These principles of application are not based on a faculty of *prudence* but are part of the moral point of view. In Habermas's concept of autonomy not only the questions of *justification* but also of *application* are attended from the point of view of practical reason. (JA, 35-39)

Thus, the guidance that any moral theory has to provide for the autonomy of the subject is considered in Discourse Ethics. To achieve moral autonomy, the question not only concerns the *justification* but also the principles of *application* of a norm. Otherwise, the consequences and side effects of the observance of a norm are not taking into account and the agent's actions cannot reach moral autonomy. This distinction is another reason to argue that Habermas's concept of autonomy cannot be charged of being certain form of abstract universalism. According to Habermas, 'Kant neglects the problem of *application*'. (JA, 35) In contrast, Discourse Ethics makes a careful distinction between the validity of norms and the correctness of singular judgments that prescribe some particular action on the basis of a valid norm. (JA, 35-36) Henceforth, the inclusion of this element guarantees that the autonomy of the moral agent in discourse is not a form of abstract universalism. It is still Universalist, but not abstract, in the sense that takes into consideration the norms of *application* of a valid norm, and therefore the results are more likely to be moral.

3. Concerning the criticism that refers to the impotence of the 'Ought', Habermas argues that Kant is actually vulnerable 'to the objection that his ethics lacks practical impact because it dichotomizes duty and inclination, reason and sense experience'. (MCCA, 207) In the view built in Discourse Ethics, the participants in communication are real people who are not split between *noumenal* selves and *phenomenal* selves. Therefore, in Habermas's theory there is not a rupture between duty and inclination. To accomplish this result, Habermas discards the doctrine of the two realms. This refers to the distinction between things in themselves and appearances. Instead of this, Habermas adheres to a Kantian

pragmatism in which reason is desublimated and it is embedded in the everyday communicative practice. (T&J, 84) The rejection of this doctrine allows that the particular interests of the empirical agents can be included in the justification of a moral norm. Henceforth, the contradiction between the universal will and the particular will that Hegel encountered in Kant's moral philosophy is rebutted in Discourse Ethics. Moreover, there is not a break between the *phenomenal* self and the *noumenal* self and Discourse Ethics can include inclinations and interests. That said, however, these particular contents need to be universalizable. A good example in this point is that everyone has reason to avoid pain to themselves. Therefore, a particular inclination (avoid pain to me) can be the object of a moral discourse because it embodies a generalizable interest (everyone has an interest to avoid pain to themselves). However, this is more problematic because this content is not agent-neutral, but agent-relative. As I discuss in the next section, Habermas discards agent-relative contents as susceptible to justify moral norms, even if they are universal.

According to Habermas, in another respect Hegel is also right, 'Practical discourse does disengage problematic actions and norms from the substantive ethics (Sittlichkeit) of their lived contexts'. (MCCA, 207) This is because in the process of justification it is necessary to make abstractions from existing motives and institutions. Otherwise, morality would not be a critical inquiry.

Notwithstanding, Habermas claims Hegel correctly argues that moral insights should become part of the concrete duties of everyday life. In this respect, Habermas asserts that:

> There has to be modicum congruence between morality and the practices of socialization and education […] between morality and socio-political institutions. (MCCA, 207)

For Habermas, morality occurs in a social context in which post-conventional ideas of law and morality have been already institutionalized at least to a certain extent. Moreover, the moral agents are socialized in the practices of reaching agreement through discourses which are shaped by principles and procedures that aim to reach universalizable contents. In Kohlberg's terms, they have attained level six of a post-conventional moral judgment.[12]

---

[12] However, in Habermas's reformulation of Kohlberg's model level six is related with the procedures that the former advances in Discourse Ethics (MCCA, 116-194) and not with Kant and Rawls's moral philosophies as Kohlberg argues. (Kohlberg, 1981; 1984)

In this way, to confront the impotence of the 'ought' Habermas argues that ethical life – through socialization – has to meet post-conventional moralities halfway. (MCCA, 207)[13] As I have shown, people need to be in a social context in which the post-conventional level of morality (in Kohlberg's sense) has been achieved and then they take to heart, and identify with universal moral norms, and are committed to resolve conflicts of interest by means of discourse. Certainly, in these conditions the odds that morality can be effective increases.

Furthermore, Habermas argues that this is only possible in the context of a modern society in which law and morality are organized by formal procedures that do not suppose the preference of particular substantive contents – i.e., principle (U), principle (D), the principle of democracy and even the legal form.[14] Morality is an historical result connected with the Enlightenment and therefore it refers in Habermas's Discourse Ethics to a theory of modernity. However, it is important to notice that while Habermas endorses this Hegelian insight in his theory, the weight of the validity of moral norms is based on the test of universalizability. Post-conventional moralities need to be met halfway by accommodating life-worlds. (MCCA, 207) Nevertheless, the normative validity of moral norms does not depend on these forms of life. Rather, that worth is grounded on the fact that they embody universalizable interests.

4. In this final part of this section, I want to assess whether and to what extent Hegel's critique of Kant can apply to Habermas's moral theory. As a general conclusion, it seems that Habermas rebuts if not all, then most of these criticisms. In this way, there are several points in which he has advanced good arguments to respond to these objections. That said, the objection that still can challenge Discourse Ethics is the criticism of the will as a tester of maxims. This issue needs to be addressed because every moral theory has to explain how the agents can be bind by morality. This is what Hegel describes in §153 of the *Philosophy of Right* as 'The *right of individuals* to their *subjective determination to freedom*'.

---

[13] Interestingly enough, this argument is also made in *Between Facts and Norms*. In this respect, Habermas asserts, 'What is more, deliberative politics is internally connected with contexts of a rationalized lifeworld that meets it halfway'. (BFN, 302; see also 358; 461; 471; 487; IO; 252)

[14] The principle of democracy states that 'Only those statutes may claim legitimacy that can meet with the assent of all citizens in a discursive process of legislation that in turn has been legally constituted' (BFN, 110). For a detailed discussion on this principle see Chapter Three, Part I, section 3. The legal form refers to the institutional framework of modern law. It comprises these three features: i) it is positivistic (depends on the will of the sovereign; ii) it is enforceable (its norms are backed by sanctions); iii) it protects subjective rights. (IO, 254) In Chapter Three, part I, section 2 I examine in detail the legal form.

As I have shown, this criticism applied to Discourse Ethics could be convincingly answered bearing in mind that Habermas has discarded the doctrine of the two realms. From the point of view of Habermas's Kantian pragmatism, reason is not opposed to the empirical world but it is located in the everyday practices of *giving reasons*. For Habermas morality is not abstract and *noumenal* as for Kant is. Henceforth, morality is concrete and part of the *phenomenal* and it is not necessarily opposed to the empirical will of the moral agent. Moral reasoning is not performed, as it is for Kant, by *noumenal* selves in the kingdom of ends, but by real social agents. Additionally, Habermas argues that his moral philosophy incorporates in questions of morality particular interests, inclinations and consequences and side effects. Therefore, apparently Habermas's concept communicative reason has incorporated several components that fortify and shield his moral theory in front of the Hegelian criticisms of the will as a tester of maxims.

In spite of this, in the next section I maintain that the problem arises again because Discourse Ethics set strict conditions on the features that moral norms need to meet to be universalizable. Habermas argues that the particular interests that can pass the principle of Universalization (U) need to be universalizable contents and never particular contents. In this way, the limits that moral discourse establishes open again the question whether and to what extent particular contents can be included in practical discourse. As I will show in the fourth section of this Chapter, in light of the concept of ideal role taking, which is built into the procedure of universalization of norms in Discourse Ethics, it is possible to rebut this Hegelian challenge. However, before I develop that, in the next section I examine in detail Habermas's notion of the universalizability of norms.

3. The universalizability of norms and the return of Hegel

So far, Habermas has advanced good arguments to rebut Hegel's critique of Kant's moral philosophy this time applied to Discourse Ethics. Nonetheless, the question whether and to what extent Habermas can refute Hegel's later criticism of Kant concerning the break between the moral will and the particular will needs further elaboration. Habermas's version of the universalizability of moral norms set strict limits in what can count as the proper contents of a practical discourse and it filters out particular contents and interests. Considering this notion of universalizability, the issue if Habermas can convincingly (or not) respond to the Hegelian criticisms of the will as a tester of maxims still remains

nebulous. In what follows, I begin developing this issue examining Habermas's version of the universalizability and I conclude that Hegel's criticism of the will as a tester of maxims applies to Discourse Ethics.

<p style="text-align:center">***</p>

Throughout the concept of universalizable or generalizable interests, Habermas has included in the contents that could claim moral worth, the interests and inclinations of the moral agents. In this way, it is possible to envision the union between the *rational* self and the *empirical* self. However, Habermas rules out agent-relative interests. He appears to disqualify agent-relative interests as reasons that are apt to justify moral norms, even if they are universal. He claims rather that the contents that are susceptible of being universalized have to be solely agent-neutral. (IO, 7, 43)[15] Additionally, a moral norm has to be in 'everyone's interest'. (MCCA, 65) Therefore, it aims to the universalizability of the moral contents and rejects particular interests.[16] This is explicitly stated by Habermas in the principle of universalization, in which valid norms have to achieve the following condition:

> (U) *All* affected can accept the consequences and the side effects its *general* observance can be anticipated to have for the satisfaction of *everyone's* interests. (MCCA, 65)

In order to expound Habermas's position, it is possible to distinguish some aspects of his version of the universalizability of norms. One weak interpretation could state that validity 'is conferred on a norm only if everyone has *an* interest in its general observance, but not necessarily the same interest' (Finlayson, 1999: 42), so this would be agent-relative but not universalizable. In that case, such interest could be particular, not common to all, and different people could agree to a norm for different reasons.[17] From the point of view of

---

[15] In this respect, Habermas is following Thomas Nagel's distinction between agent neutral and agent relative reasons for action. (1986: 153) Agent-relative interest is one that refers back to the subject and her on interest. In general terms, in the Kantian tradition agent-relative interests are explicitly ruled out. Although, endorsing Kantianism, Christine Korsgaard criticizes the distinction because for her it does not cohere with an intersubjectivistic account of morality. (1993) Nevertheless, Habermas upholds the distinction.

[16] There is also in *Moral Consciousness and Communicative action* another formulation of the principle U, which reads: 'For a norm to be valid, the consequences and side effects of its general observance for the satisfaction of *each person's particular interests* must be acceptable to all'. (MCCA, 197) This formulation captures an ambiguity, because it implies the idea that Discourse Ethics could include *particular interests*. Nonetheless, this translation of U from the original German version is plausible (Finlayson, 1999: 51), as I am going to show is not coherent with what Habermas has been arguing systematically.

[17] To understand the weak condition of universalizability, I draw attention in the following example. In a particular community some people wants to protect a building for its historical worth and others because it is good to have that building as many tourists visit the area to see that heritage and they spend money in the

Hegel's criticisms, this version seems to have a clear advantage. In this case, the agents have their own particular interests to follow a moral norm. The particular will and the universal will are joined weakly through a contingent overlap of particular interests, what Rehg refers as strategic 'marriages of convenience'. (Rehg, 1994: 40) Then, it looks as Hegel's criticism can be answered. Notwithstanding, neither Hegel's account of the will nor of *Sittlichkeit* could be read in this loose sense.

Moreover, it is not difficult to show that Habermas does not endorse this version in his understanding of the universalizability of norms. This is because, as said in previous paragraphs, a norm that can claim to have moral worth should be based on agent-neutral reasons and never on agent-relative reasons. Additionally, it has to be underpinned by the same reason. Conversely, in the weak version a norm could be in everyone's interest but different agents might have different interests and agent-relative reasons to uphold it. As I have shown, in Habermas's view of the universalizability, norms have to be in everyone's interest but at the same time they have to be based on agent-neutral reasons. Thus, they cannot be grounded on particular interests, as the weak version allows. In this regard, Habermas asserts,

> Particular interests are those that prove on the basis of discourse testing not to be susceptible of generalization and thus to require compromise. (Habermas, 1982: 258)[18]

Secondly, a weak version could mean that any content is agreed by a compromise or by the contingent overlap of particular interests. This description is similar to a contractualist position, which reduces 'the validity of moral norms to a conventional agreement [*Vereinbarung*] among rational egoists, that is, to a happy coincidence of their respective interests'. (T&J, 241) According to Habermas, this type of agreement does not possess the requisite stringency that moral validity demands. Hence, in Discourse Ethics, the deontological worth that moral norms claim it is not coherent with this version of the universalizability. In this weak version, everyone one has some reason to agree to a norm,

---

local shops and restaurants. The different members share a content, which can be stated as protecting that particular building, but for different reasons. Here, more than a rational consensus one faces a merely de facto consensus or compromise.

[18] This formulation raises the following question: Is the norm on which people compromise in everyone's interest to some extent? In other words, can compromise rest on everyone's interest? If Habermas means everyone's interest with generalization, then compromise does not rest in everyone's interest.

but each person does not have the same one. So everyone has some interest in abiding by the norm, but not the same one in each case.

Henceforth, Habermas's version of the universalizability is stronger than the weak conception stated before. In his account (U) means that the moral validity of a norm depends on everyone's accepting it for the same reason on the basis of the *same interest*. As he states in *Moral Consciousness and Communicative Action*:

> True impartiality pertains only to that standpoint from which one can generalize precisely those norms that can count on universal assent because they perceptible embody an interest common to all affected. (MCCA, 65)

Norms are impartial and that means that agent-relative interests are not susceptible of being universalized, because they refer back to the particular subjects whose interests they represent. A norm is universalizable, Habermas claims, when it embodies 'everyone's interest' in a stronger sense. In other words, it is valid if the interest that is served by the norm, and hence the reason why different people agree to it, is one and the same for all. Habermas's moral theory establishes its aim as a rational consensus and not a compromise. In Discourse Ethics, Habermas holds, normative rightness is analogous with epistemic truth, and this implies that, in both cases a consensus is based on the agents sharing or having the same reasons.

However, the stronger version of the universalizability can have two interpretations because it is possible to distinguish more than one form of sharing or *having the same interest.* (Finlayson, 1999: 43) In this regard, consensus 'might be understood in either of two senses, the one a distributive sense, as I shall say, the other a collective one'. (Pettit, 1982: 215) Therefore, an interest can be collectively or distributively universalizable. Interestingly enough, in both cases different moral agents share the same interest in some sense, and therefore both versions could count as Habermas's notion of the universalizability of norms.

Now, in order to understand both versions I rely on the following example. It is possible to argue that everyone has an interest in avoiding suffering. However, this is still an ambiguous formulation because someone might have an interest in her avoiding her own

suffering but not in everyone avoiding their suffering. This is the distributive or relaxed case. In contrast, everyone might share the interest in avoiding not only his or her own suffering but in preventing suffering per se, her and everybody else's. This is the collective or more stringent version of the universalizability of norms.

In light of Hegel's criticism of the will as a tester of maxims the relaxed version seems to provide a straightforward response, because the agent-relative interest of the agent can be included in practical discourse. Thus, apparently in the distributive interpretation it is relatively easy to attain moral autonomy and reconcile the particular will and the universal will. Nevertheless, there are good reasons to argue that Habermas endorses the second more stringent formulation. First, as I have shown, he rules out agent-relative interests as being germane to the moral domain, and in the relaxed version a consensus can rest (partly or wholly) on the basis of universally distributed agent-relative interests. Habermas appear to assume that all agent-relative interests are particular, and that all moral norms are impartial and agent-neutral. In this respect, he asserts that, 'in contrast with empiricist varieties of contractualism, this view holds that these reasons are not conceived as agent-relative motives, thereby leaving the epistemic core of moral validity intact'. (IO, 7) Secondly, the distributive more relaxed version could be coherent with a monological notion of moral autonomy which is at odds with Habermas's dialogical approach. In the former case, whether the participants engage in discourse or not, and whether they actually know what others' interest are, is a contingent matter, provided it is true that the norms serves an interest that each individual has, and that this interest is the same in each case. (Pettit, 1982: 215)

In summary, Habermas neither endorses the view that a moral norm could rest on a compromise of particular reasons, nor the more relaxed version of the universalizability of norms in which they rest on agent-relative but universal interests. Rather, Discourse Ethics presuppose the collective more stringent version of the universalizability of norms in which they rest only on agent-neutral reasons. Consequently, bearing in mind that Habermas endorses this version it is still necessary to find in Discourse Ethics a proper answer, or at least the beginnings of a plausible response to the Hegelian criticism of the will as a tester of maxims. That is going to be the focus on the next section, by means of an analysis of the concept of ideal role taking, which is at the centre of the procedure of universalization of Habermas's moral theory.

4. The process of ideal role taking

This section develops a Habermasian answer to Hegel's later criticisms of Kant's moral philosophy by analysing Mead's process of ideal role taking. In this concept one can find an additional reason to argue that a rational consensus in Discourse Ethics is different from compromises or strategic 'marriages of convenience'. (Rehg, 1994: 40) Both the stringent and the relaxed version of universalizability are stronger than compromise. Nonetheless, in light of the process of ideal role taking I will show that Habermas endorses the stringent version of the universalizability of norms. Moreover, through this concept it is possible to find the beginnings of an answer to the Hegelian challenge, namely, ideal role taking is an adequate account of how particular wills, in discourse, become moral wills. Consequently, it accomplishes what Kant's moral theory does not seem to be able to do. Thus, on the one hand, the process is evidence that Habermas uphold the stringent version of the universalizability (See the previous section), and on the other hand, could help to answer the Hegelian challenge.[19]

This process refers to the post-conventional level of morality, in which the participants in discourse take a universalistic perspective. Moreover, it rules out any attempt to ground this procedure on the philosophy of consciousness, because it is a collective practice that has to be performed in public discourse. (JA, 49; BFN, 109-110) In Discourse Ethics, specifically in the principle of universalization (U), Habermas 'picks up Mead's notion of ideal perspective-taking and demands that participants take an interest in each other's interests'. (Rehg, 1994: 39) Here the participants both recognize their own interests and the interests of everybody else, throughout 'the expansion of a reversible exchange of interpretative perspectives'. (T&J, 105) Elsewhere, Habermas has defined this process as the 'complete reversibility of participant perspectives that *unleashes* the higher-level intersubjectivity of the deliberating collectivity'. (BFN, 228)

Here I maintain that in this concept it is possible to find an answer to Hegel's later objection of Kant. In the process of ideal role taking, particular wills are channelled into the *concept* of the will – in Hegel's *parlance*. In other words, the interests and inclinations of the participants are shaped so as to take into consideration everybody else's particular

---

[19] As it is well known, Mead is indebted to the Hegelian tradition. Therefore, it is understandable that Habermas could solve some Hegelian challenges by means of his appropriation of Mead's concepts within the frame of his Kantian pragmatism. (See Mead, 1967)

perspectives. This means that in discourse the agent is brought to the insight that her own interest matters equally to the interest of everybody else. Hence, bearing in mind this process it is possible to claim that in Habermas's version of the moral point of view throughout public discourse agent-relative interests can be combined, coincide and be shaped in terms of agent-neutral interests or generalizable interests. In other words, agent-relative interests are shaped so people take an interest in the agent-relative interests of all others.

In this way, Habermas can successfully answer Hegel's later criticism of Kant by means of the introduction of the process of ideal role taking. That said, although, this process implies to slacken the strict distinction between agent-relative and agent-neutral reasons that is the view of impartiality that Discourse Ethics seems to build. (See section 3) As I have shown, in the process of ideal role taking, agent-neutral reasons can be at the same time agent-relative reasons. Now, it is important to remark that it is not the case that agent-relative interests merely overlap with agent-neutral (moral) reasons. Rather, the concept of agent-neutral interests needs to be slackened and leave room for the inclusion of agent-relative interests within its formulation. To my mind, this is what happens when we examine the dialectical concept of ideal role taking.

Kant was not able to accomplish this. His version of autonomy was monological and it was stymied by his two world metaphysics. These are the main reasons why he could not solve the later criticism of the will as a tester of maxims. It is not possible, Habermas claims, that the expansion of interchangeable perspectives – which is what the process of ideal role taking performs – can be made by every individual privately. Only another real participant in discourse can make me realize that her interest matters. Moreover, in Kant's moral theory it was difficult to envision the inclusion of inclinations and interests into the shaping of the moral will. These contents for Kant were ruled out from the domain of pure practical reason, because they were part of the *phenomenal* self and not of the *noumenal* self. After all, this was the condition to arrive to universal and unconditionally binding norms. In Kant's conception of pure practical reason, '*if* there is such an unconditional, categorical imperative, *then* it must be one that binds all rational agents necessarily independently of what particular purposes they will'. (Pinkard, 2002: 50) In Discourse Ethics, this two realms distinction is ruled out. Due to this, practical discourse can go to work on contents of the empirical will – particular interests – that on Kant's conception have to be excluded or

overridden. In Habermas, the contents that could be part of a rational consensus include inclinations and interests. They can be binding as long as they are shaped by means of a procedure of universalization that reconstructs generalizable interests.

5. Moral norms, values and the formula of humanity

The previous section showed that throughout the process of ideal role taking – which is forced on participants by discourse – Habermas has a good argument to rebut the Hegelian criticism of Kant's moral philosophy which concerns the break between the moral will and the particular will. There is another distinction that Habermas sets out in Discourse Ethics, namely, between moral norms and values (MCCA, 103-104; JA, 1-17; T&J, 213-235; Baynes, 2016: 104; Finlayson, 1999: 47), which can be the object of a further Hegelian criticism. This criticism states that Habermas needs to blur the distinction (McCarthy; 1991; Putnam; 2002), and due to this he would be in a better position to refute Hegel's criticism of the will as a tester of maxims. In despite of this, I agree with those who claim that the distinction is necessary and can be convincingly defended. (See for example Forst, 1994: 44; 2007: 62-78)

In what follows, I maintain that the distinction between moral norms and values does not have to be blurred as Hegel and Hegelians assert. Still, the interplay between them may be more complex than Habermas allows (1). Then, I examine a Habermasian answer to Hegel's criticism of the will as a tester of maxims, which does not imply that the distinction between moral norms and values should be overcome, as Hegel suggests. Rather, this reply is connected with a Kantian insight, the recognition of others as ends in themselves in Kant's second formulation of the categorical imperative (4:429) (2).

1. Hegel's strategy to complete Kant's moral philosophy, considering the problems that he found in the latter, namely, the break between the empirical will and the concept of the will, was throughout the dialectical relationship that he proposed between *Moralität* and *Sittlichkeit*. In light of this, the issue is whether and to what extent Habermas can rebut the objection to the will as a tester of maxims, while at the same time endorsing the distinction between moral norms and ethical values. To understand Hegel's solution it is important to explain what role he grants to *Sittlichkeit*. According to Pinkard, in Hegel's philosophy:

> What gives legitimacy […] to each of these forms of *Sittlichkeit* is that in each, the individual can be 'at home with himself' in that he can find his own *particular* way of being ethical ('universal'), his own particular *way* of orienting himself in 'social space' in the light of a determinately structured 'whole' that nonetheless embodies within itself reasons that can be shared by all. (Pinkard, 1999: 227)

Hence, in order to adopt reasons that can be shared by all, that is to say, to endorse the moral standpoint, it is necessary to recognize the point of view of ethical life. In Hegel's philosophy *Sittlichkeit* refers to the institutions and practices that surround *Moralität*, which allows the agents to be at home with themselves and at the same time respect the moral point of view. In other words, through the tension between *Moralität* and *Sittlichkeit* the agents can be moral for their own reasons and this was Hegel's solution to the deficits that he found in Kant's moral philosophy. By means of this dialectic, it is possible to envision the reconciliation between the universal will and the particular will because 'ethical life' is the domain and the medium of socialization, which enables the agents to be bound by morality.

Against Hegel's alternative, Habermas supports the strict distinction between moral norms and ethical values. For example, he asserts that 'particular values are ultimately discarded as being not susceptible to consensus'. (MCCA, 103) To put it briefly, ethical values are not the objects of morality because they are not universalizable. Ethical values are of a different nature. Habermas distinguishes at least two types of ethical values. First, the ethical-existential justification of life decisions that one can vindicate in relation to oneself and 'significant others'. (JA, 1-17) This level already has an intersubjective dimension, since such values are justified in the context of ethical communities. Second, the ethical in his theory can have a political meaning. The form of life of the political community 'that is in each case our own' constitutes the reference system for justifying decisions that are supposed to express an authentic, collective self-understanding. (BFN, 108) Both types are connected because, 'The "existential" question of who I am and would like to be, which is posed in the singular, is repeated in the plural — and is thus given a different meaning'. (BFN, 160)

In his Hegelian reading of Discourse Ethics, McCarthy asserts that implicitly Habermas blurs the distinction because 'the articulation of needs in practical discourse will draw upon

existing standards of value; as interpreted, needs are internally related to, and thus inseparable from, cultural values'. (1991: 183) Thus, he proposes that Habermas should be clearer and slacken the distinction between moral norms and ethical values. To my mind, Habermas can accept the premises of McCarthy's argument but not the consequence that the latter draws from them.

Taking into account the unambiguous version of the contents that are susceptible of rational consensus[20], it is clear that ethical values are not the expected outcomes of the procedure of universalization (U) of practical discourses. From the definitions given above, ethical values have a particular dimension which is connected with the person and her community. Henceforth, they cannot reach the universal level of all those possible affected by a moral norm which characterizes morality in Habermas. Values might bleed into the process by which interests are shaped in discourse, but they are then tested throughout the principle of universalization of norms. As long as they can claim moral validity these elements can successfully pass the principle of universalization (U) and underpin the justification of norms. In order to have the moral status demanded by Discourse Ethics these considerations have to rest on universalizable and agent-neutral interest.[21] Otherwise, if their content cannot be universalizable, they are filtered out.[22] In my interpretation, moral norms can come in the language and the form of values, but that does mean that they can have deontological worth if they are not at bottom moral norms. In this way, certainly conceptions of the good and values could contribute to carry out moral discourses but as Baynes correctly asserts:

> Insofar as it is not reasonable to expect that a given conception of the good will meet with the agreement of all in a discourse, that conception itself can no longer serve as a moral norm or be invoked as the basis for a moral norm. (Baynes, 2016: 120)

Consequently, even if it is true that interests (qua need interpretations) are initially formed in the light of values, even these are subject to further interpretation in intersubjective discourse. So it is wrong of McCarthy to assume that because the values that feed into

---

[20] See section three of this Chapter.
[21] See section three of this Chapter.
[22] For instance, certain values like generosity or respect for others could be seen as both moral norms and they also can be justified in light of a modern ethical life *Sittlichkeit*. Nonetheless, if they can be universalizable this depends on the fact that they convey a moral content and not because they are ethical values.

discourse are individual or group specific, that so will the norms that are the outputs of discourse. Hence, in my view, Habermas does not endorse the solution proposed by Hegel and McCarthy and does not need to do it. Notwithstanding, this analysis have shown that the interplay between norms and values seems to be more complex than Habermas allows.

In summary, in this section I have expounded and defended Habermas's distinction between moral norms and ethical values. However, I have not answered yet if this strict condition allows Habermas to respond to Hegel criticism of the will as a tester of maxims. In the next section, I examine a further reply which relies on a Kantian inside, namely, the recognition of *me* and others as end in themselves.

2. In what follows, I develop an answer to the Hegelian challenge from a Kantian interpretation of Discourse Ethics. In Habermas's justification of the principle of universalization (see Chapter One), one of the elements is a claim of what it means to 'justify a norm of action' or what he names a weak '*concept of normative justification*' (IO, 45). In simple terms, this refers to the obligation to give an account on validity claims. According to Baynes, 'It is a thin, but non-negligible, idea of norm-justification in which individuals are viewed as mutually accountable agents'. (Baynes, 2016: 115) To my understanding, if in an interpretation of this component it is possible to find a response to Hegel's charge, then Habermas's moral philosophy in its Kantian reading also stands.

In his interpretation of practical reason and autonomy, Rainer Forst has argued that, 'whenever an ethical or instrumental reason for being moral is asked for, the moral standpoint simply cannot be found'. (2007: 76) Of course, Habermas agrees with this definition, namely, when he draws the distinction between norms and values and earlier when he distinguishes between communicative and instrumental reason. Thus, morality is based in itself and it does not need to find its ground anywhere else.

This autonomous foundation of morality needs the support of a moral conception of the selves. If that is the case, then there is another alternative to close the gap between the universal will and the particular will: 'communicative subjects have a moral interest in recognizing universalizable interests'. (Finlayson, 1999: 48-49) In a first reading, this seems to be circular. For some reason, the subjects are moral and it just happens that they have a particular interest to recognize universalizable interests. Henceforth, morality is effective

because the moral agents have their own reasons to follow and respect moral norms that are in 'everyone's interest'.

I maintain that this argument is not circular in a negative way and it is justifiable. The reason is that in order to be a moral agent at all it is necessary to recognize others as valid agents too. This insight can be found in Kant's second formulation of the categorical imperative, the formula of humanity to respect others as not only means but always also as ends in themselves. Also in Fichte's theory of recognition (2000), in Hegel's struggle for recognition in the *Phenomenology* (1977a), and also in Habermas weak '*concept of normative justification*' (IO, 45) broached in the principle of discourse (D). Here, it is an interest of the empirical will to recognize everybody else's interest. In other words, it is impossible for an agent to be moral and to be autonomous means to recognize the right of every other to participate in a communicative process of giving reasons. In this respect, Forst asserts:

> Perceiving and *cognizing* others as *human beings* also means *recognizing* them as *moral persons* with a right to reciprocal and general justification, and knowing that no further reason for this recognition is required aside from this reference to the shared characteristic of being human. This is the fundamental insight of an autonomous morality of autonomy. (Forst, 2007: 77)

Thus, practical reason is intersubjective. Strictly speaking, it is a common practice of giving and sharing reasons. Then, it implies the recognition of the interests of the others. As a moral agent I know that my autonomy can only arise in the company of others. Therefore, I have an interest to recognize everyone else's interests because I am also one of them. Somehow, not recognizing their interests means that my own interest cannot be recognized as well. Hence, when I am not moral I am not recognizing myself. Kant's second formulation of the categorical imperative and the respect of persons as end in themselves (4:429), goes in this direction. I recognize the autonomy of the other as equal to me and in this movement I also recognize myself.[23]

In this regard, Rainer Forst supports an autonomous conception of morality and in this point he certainly agrees with Habermas's account of practical reason and autonomy. Put

---

[23] Freyenhagen (2011) also replies to the *empty formalism* objection adducing the formula of humanity. In that formulation of the categorical imperative, 'we can derive some content from the mere idea of duty for duty's sake, specifically the objective end of humanity'. (2011: 168)

simply, this conception means that morality finds its ground in itself. Therefore, the general will and the particular will can be joined together for moral reasons and it is not necessary, as Hegel argues, to blur the distinction between moral norms and values. Accordingly, the duty of giving and accepting reasons is the root from which morality is grounded. Moreover, justification is the 'best possible way to philosophically reconstruct the Kantian categorical imperative to respect other persons as 'ends in themselves''. (Forst, 2007: 2)

As I have shown, in Discourse Ethics autonomy is understood in intersubjective terms (see particularly Chapter One). This pertains to the insight that the recognition of the autonomy of the others is at the same time the recognition of *my* own autonomy. Therefore, in the idea of mutual recognition (and mutual understanding) the gap between universalizable content and particular content can be closed. This is because it is in my own interest to recognize the agent-neutral interest of everyone being recognized in their autonomy. If that argument is tenable, then moral agents can be both moral and autonomous for her own reasons. Hence, mutual recognition means that a real person makes a generalizable interest – in the sense settled by the stringent condition of the universalizability of moral norms (See section three of this Chapter) – his or her own. Of course, this justification proceeds from the assumption that 'the participants do not wish to resolve their conflicts through violence, or even compromise, but through communication'. (IO, 39) In communication, a particular person might endorse as hers the agent-neutral interest of everyone else to be recognized as a proper participant in practical discourse. To be sure, still this argument might sound unappealing. If both, the argument broached in section four concerning ideal role taking and the one developed in this section fail, then it seems to me that a Habermasian solution to Hegel's criticism of the will as a tester of maxims it is simple unworkable. In this Chapter I have tried to give reasons to justify the contrary.

Part II

Habermas's Political Kantianism and Hegelian criticisms

3

The Kantian Foundations of Habermas's Political Theory

In this Chapter I examine the inner structures of Habermas's theory of legal and political legitimacy and I conclude that they are framed in Kantian terms.[1] On the one hand, this statement does not seem to be problematic because Habermas openly recognizes that his political theory has a Kantian pedigree. For example, he has argued that his notion of legitimacy is based on a Kantian Republicanism. (MW, 94, 113; Forst, 2011: 180) Moreover, in the debate with Rawls on *Political Liberalism*, Habermas claims that the latter 'proposed an intersubjectivistic version of Kant's principle of autonomy' (RPR, 25), and at the same time he asserts concerning Rawls's *Theory of Justice* that 'I share its intentions and regard its essentials results as correct'. (RPR, 25) Due to this, Habermas's claims his critique of Rawls remains within 'the bounds of a familial dispute' (RPR, 25), a family that finds its affiliation in the Kantian tradition. (Forst, 2007: 80; Laden, 2011: 135)

---

[1] Habermas's theory of legal and political legitimacy gains its concrete shape and substance in works that were written mainly during the 90's. This topic is developed in *Between Facts and Norms*, in the *Inclusion of the Other* and in the debate with Rawls. (Habermas's RPR and MW) Here I agree with Finlayson (2011, 2016) for whom the relevant issue in the dispute is the question for legal and political legitimacy. Otherwise, I would not be discussing this quarrel when I am broaching Habermas's Kantianism in his legal and political theory. In the final Chapter of this thesis I focus more on the debate than I do in this Chapter. Habermas's theory of legal and political legitimacy is also discussed in further articles and in Habermas's replies to critics. See Rozenfeld and Arato (1998) and Finlayson. (2011) See also Habermas. (2001) In what follows, I do not explore Habermas's contributions on international law. Although, the analysis of this theme is helpful to understand the relationship between morality and law in Habermas (See Flynn, 2003, 2011), I cannot discuss this issue in this work. Habermas's position concerning this theme is set out in his 'Kant's Idea of Perpetual Peace: At Two Hundred Years' Historical Remove' in IO (165-202) and the Postnational Constellation (PC). Finally and for the sake of my argument I do not discuss Habermas's position concerning the project of the European Union.

On the other hand, the aforementioned claim is controversial because Habermas connects Kant with the tradition of natural law (and rational law) which he wants to avoid. (BFN, 44, 101 & 103; BNR, 102; Baynes, 2015: 135) This is, he claims, because these reconstructions lost contact with the reality of contemporary societies (BFN, 42-45; Finlayson, 2011: 7; Flynn, 2011: 252) and they have a paternalistic understanding of political philosophy in which the political philosopher offers to the citizens the basic frame and the limits of what can count as legitimate. (Habermas, 2011: 296; Forst, 2011: 165) Therefore, they set arbitrary and external constrains on the political autonomy of the demos. Then, the question whether and to what extent Habermas is developing a Kantian political theory cannot be easily answered. Here, I argue that despite what Habermas says to the contrary – especially what he wrote in BFN – his political philosophy is essentially framed in a modified version of Kant's notion of pure practical reason and autonomy (See Chapter One). In other words, the inner architecture of Habermas's theory of legitimacy is grounded on this philosophical element, despite the fact that he includes components which supposedly go beyond, and differ from, the Kantian tradition – i.e., the legal form as an historical input. (BFN, 82-131; Flynn, 2003: 438-439) In this respect, I side with some sympathetic critics of Habermas who argue that a fully immanent – or intra-legal – reconstruction of constitutional democracy and the 'system of rights' is not possible. (Flynn, 2003; 2011; Forst, 2007)

The Kantian framework is, I argue, necessary to understand Habermas's position properly, but he tends to leave it to one side in BFN when he argues that the principle of discourse is an immanent reconstruction of a social practice which stands below the threshold of morality (BFN, 107, 121; BNR, 84, 89), and by the emphasis that Habermas gives to the legal form as a historical and functional input. (BFN, 83, 111-113; 117-130, Finlayson, 2011: 7-11; Hedrick, 2010: 105-106) In this Chapter I start from one of the conclusions of Chapter One, namely, that despite the fact that the principle of discourse is an immanent social practice – and in that sense here Habermas is taking on board Hegel's critique of Kant – it also has a Kantian pedigree, because it is based on a Kantian account of practical reason. In light of this claim, I examine Habermas's principle of democracy which unfolds from the principle of discourse and the legal form. Moreover, in this Chapter I show that it is possible to give a Kantian interpretation to the legal form. At the same time, in the debate with Rawls, Habermas's Kantianism emerges in sharp relief in the justification of

legal and political legitimacy.[2] The underlying Kantianism of Habermas's position as it figures in his debate with Rawls is examined in more detail in Chapter Four. In this Chapter I set the stage for that discussion (and also for the discussion of Hegelian criticisms of Habermas's political theory) and I shed light on the question whether and to what extent the normative dimension of Habermas's theory of legitimacy demands relies on the reconstruction of the Kantian notion of pure practical reason and autonomy.

As I have shown in Chapter One, this Kantian element is a *modified* and *attenuated* version of Kant's concept of autonomy. The normative component in Habermas – the principle of discourse – refers to the intersubjective account of the post-conventional level of justification (MCCA, JA) whereas in Kant, autonomy is grounded on a metaphysical view based on the philosophy of the subject. Conversely, the principle of discourse has a *weak-transcendental* status. (MCCA, 75-76) It is *weak* because it refers to institutions and practices of modern *Sittlichkeit*. Moreover, it is not linked with rational subjectivity, but to the specifically human form of life mediated through communication and discourse. That said, it also seems to apply universally whenever human beings aim to reach a mutual understanding about something.[3] Therefore, it is *transcendental* but at the same time is *weak*. According to Habermas, in the theory of legitimacy these normative components – fundamentally the principle of discourse – are clad in the medium of law:

> Once moral principles must be embodied in the medium of coercive and positive law, the freedom of the moral person splits into the public autonomy of co-legislators and the private autonomy of addressees of the law, in such a way that they reciprocally presuppose one another. (MW, 113)

Now, private and public autonomy are protected by the 'system of rights' which arises from the relationship of the normative component, the principle of Discourse (D), and the factical component, the legal form. (BFN, 82-131) This is the co-originality thesis which aims to give equal weight to the liberties of the moderns – whom stressed individual

---

[2] See specially MW. See Forst. (2011) In his reply to Rainer Forst's article (2011), Habermas rejects a moral argument for the reconstruction of legal and political legitimacy. (Habermas, 2011: 295-298) Habermas asserts that the thesis that basic rights 'are conceptualized in intrinsically juridical terms as individual rights that absolve agents from the duty to provide justifications speaks against a monistic construction of law and morality'. (Habermas, 2011: 298) In order to avoid this controversial and difficult issue, I have preferred to state that this support is a normative Kantian component which not necessarily as a moral tinge.

[3] In the *Groundwork* Kant develops a transcendental argument to justify the moral law (the categorical imperative) on the transcendental idea of freedom. However, as I have shown in Chapter One, in the *Second Critique* Kant also develops a second strategy to justify the moral law, namely, the doctrine of the fact of reason. (See specially Section 5 of Chapter One)

freedom – and the liberties of the ancients – whom stressed collective self-determination (Constant, 1988). Habermas systematically claims that no one before him has been able to find the balance between both principles (BFN, 84; 100-101; IO, 258). However, in BFN and elsewhere Habermas endorses the view that his theory is a 'two-stage' construction. (RH, 63, 76; Habermas, 2001: 766-782; Habermas, 2011: 295; Forst, 2011: 171) This methodology makes that one of the criticism that Habermas develops against Rawls's *Liberalism*, as he understands it, namely, that this doctrine 'generates a priority of liberal rights which demotes the democratic process to an inferior status' (RPR, 41), could apply also to his own critical theory. Therefore, if this interpretation of Habermas's co-originality is correct, then he would not be immune to Ingeborg Maus's criticism of Kant and Kantianism, namely, that this way to reconstruct the co-originality is *detrimental* to *popular sovereignty*. (Maus, 1995: 825-882) Moreover, as I have shown, according to some critics the 'system of rights' needs moral support. (Flynn, 2003; 2011; Forst, 2007; Cohen, 1999) In light of the co-originality thesis, this means that normative content is included at the beginning, before the practice of public autonomy takes place.

Before I continue, a particular aspect of the idea of a 'two-stage' construction needs clarification to avoid misunderstandings. This methodology does not grant in a first step the priority only to liberal rights, but also to republican rights.[4] Habermas asserts that the theory of legitimacy 'consists of a "system of rights" that at first only serves to constitute the procedure of democratic law making' (Habermas, 2011: 295). In the first stage, abstract rights or 'unsaturated placeholders' (BFN, 125) which define private and public autonomy are framed. Then, in the second stage, they gain their concrete shape by means of the 'exercise, the actual carrying out' (Habermas, 2001: 778) of public autonomy. From this, in this Chapter I will argue that Habermas relies on a Kantian methodology that gives priority to a normative frame of rights (the 'system of rights') which gains its actual shape and substance in a second step by means of its embodiment in the institutions and practices of modern societies.

In order to develop these issues this Chapter is divided into two parts. In the first, I expound the basic structure of Habermas's political theory and I show that it is conceptually[5] related to Kant's *Rechtslehre*. This part develops the relationship between Kant

---

[4] Kenneth Baynes opens a similar interpretation when he states that 'the system of rights (including the rights of public autonomy) must be institutionalized'. (Baynes, 2016: 142)

[5] In this work, I use the idea of a 'conceptual relationship' or also 'internal relationship' in the sense that if two components are conceptually or internally related it means that their relationship is necessary and not

and Habermas concerning the tension between [*Faktizität und Geltung*]; then, bearing in mind this distinction, I examine the legal form. Finally I discuss the principle of democracy and its relationship with Kant's principle of Universal Right (I).

In the second part, I examine Habermas's co-originality between private and public autonomy and I argue that at bottom it relies on a modified Kantian notion of pure practical reason and autonomy. At the beginning, I introduce in broader bush-strokes the issue of the co-originality in political philosophy and in Kant; then, I discuss the Habermasian version of the co-originality thesis and the 'system of rights'; and finally I assess Habermas's thesis from the point of view of some influential critics. (Rawls's RH, Larmore, 1995, Flynn, 2003, 2011, Forst, 2011, Cohen, 1999) (II).

I. Habermas's Kantianism in the theory of legitimacy

The main book in which Habermas develops his political philosophy is BFN. This extremely complex and rich text ranges among several branches of philosophical and social-theoretical investigation. (Michelmann, 1996: 307) Despite this wide scope, at the core of Habermas's theory of legal and political legitimacy lies a *modified* and *attenuated* Kantian element which pertains to the post-conventional notion of normative justification. This Kantian component is broached in the principle of discourse (D) which is embedded in the legal medium. By means of this arises the principle of democracy and the 'system of rights' (BFN, 82-131) and both together confer legitimacy to the law-making practice. Before I examine in detail Habermas's 'system of rights' (See second part of this Chapter), in this first part I elucidate the basic distinctions that Habermas uses to develop his concept of legitimacy and I show that they are fundamentally Kantian.

I divide this part of this Chapter into three sections. In the first, I show that the distinction between *Facticity* and *Validity* frames Habermas's theory of Constitutional Democracy and at the same time has a conceptual relationship with Kant's *Rechtslehre*. In this respect, the distinction not only shapes the theories of legitimacy of both philosophers. Rather, the link is stronger and refers to the fact that both Kant's notion of pure practical reason and Habermas's concept of communicative reason share in their inner workings the distinction between *Facticity* and *Validity*. Moreover – continuing the argument I developed in Chapter One – Habermas's notion of communicative reason has a Kantian pedigree because is

---

merely possible. For example, legitimate law is conceptually tied to the notion of political autonomy; otherwise it is simply not legitimate law. (See IO, 254)

connected to Kant's 'ideas of pure reason'. Bearing in mind these connections, I conclude that the very tension between *Facticity* and *Validity* is conceptually related to Kant's notion of pure practical reason (1).

In the second section, I examine in light of the tension between *Facticity* and *Validity* Habermas's Kantianism with a focus on the legal form. Here I argue that at bottom, the legal medium has an instrumental role in the protection of the autonomy of the subjects and the actual carrying out of the practice of exchanging reasons in legal and political discourses. In other words, the legal form has a functional place in the theory protecting the normative content, namely, either the Kantian notion of pure practical reason or the Habermasian notion of communicative reason (2).

Finally in the third section, having described the essential features of the principle of discourse (D)[6] and of the legal form[7], I study the chief principle of legal and political legitimacy in Habermas's theory, namely, the principle of democracy. The latter results from the interpenetration [*Verschränkung*] of the principle of discourse and the legal form. On the one hand, (following the argument developed in Chapter One) (D) has a *weak-transcendental* status and it is fundamentally Kantian. On the other hand, the legal form is a functional and historical input. Here I compare Habermas's principle of democracy with Kant's main principle of legitimacy: the principle of Universal Right (3).

1. *Facticity, Validity* and Communicative Reason

The *Theory of Communicative Action* addresses the sociological question of *how social order is possible*. Habermas argues that the conceptions of reason and agency that have been at the centre of the answers proposed to this issue have been incomplete. Neither the model of instrumental rationality – familiar in rational choice approaches of agency – nor the notion of functional rationality – i.e., the Marxist tradition and Luhmann's systems theory – 'can account for the contribution of the normative self-understanding of social actors'. (Baynes, 2016: 131) Following Baynes's account of Habermas's theory, these approaches disregard both the normative component and the perspective of the agents from the sociological analysis.

---

[6] See Chapter One.
[7] Section two of this Part.

That said, however, Habermas recognizes that significant contributions have been made in this tradition to integrate both the normative dimension and the point of view of the participants – i.e., Durkheim and Parsons. Habermas continues these theoretical trends because, as I have shown in Chapter One, the *Theory of Communicative Action* puts at the centre the question of its own normative foundations. (TCA 1, xli) Axel Honneth argues that Discourse Ethics gives account of the normative standards of Habermas's critical theory. (1991: 282) Probably that is one reason (perhaps the main reason) why right after Habermas wrote TCA in the 80's then he focused on developing the program of Discourse Ethics. (Finlayson, 2013: 522-23) Moreover, TCA focuses on the normative question of social integration of modern societies. At the same time, TCA incorporates the perspective of the participants for at least two reasons: first, the reconstructive methodology that is central in Habermas's theory refers to the knowledge that the agents have to perform in communicative processes (TCA 1, 103; PT, 23, 36); second, the centrality that TCA gives to the concept of the life-world which is reproduced communicatively and by the agency of the individuals.[8]

These elements of Habermas's social theory play a central role in the theory of legal and political legitimacy broached in BFN.[9] In this book Habermas develops from the very beginning the connections and the place that the constitutional state and democracy have into the more general sociological theory. Law is a central mechanism to guarantee the integration of modern societies and it relieves other mechanisms that perform these functions but by other means – i.e., morality, ethics and religion as well. Furthermore, the normative component that is broached in Discourse Ethics, namely, the principle of discourse (D), has a central place in the reconstruction of the concept of legitimacy in Habermas – the principle of democracy.[10] (BFN, 82-131) Additionally, legal and political legitimacy are inherent to the perspective of the participants. This is because in a post-

---

[8] The life-world is a key concept of the sociological theory developed in TCA. According to Habermas, society is divided into the life-world and the systems of the State bureaucracy and the market. The latter systems are steered by political power and money. On the contrary, the life-world is a communicative medium and it is integrated by culture (symbolic systems), society (norms that regulate interpersonal relationships) and personality (the individual's life-story and identity). (TCA 2)

[9] As a matter of fact, in TAC Habermas develops a more ambivalent view of law. On the one hand, law is an institution responsible for integration. On the other hand, it is also a medium through which social systems colonize the life world. (TAC 2, 365) The latter interpretation is more coherent with the traditional view of the first generation of the Frankfurt School. Adorno and Horkheimer combined the Weberian theory of the rationalization of modern societies with the Marxist critique of the capitalist society and in this way law was assessed as a formal system that contributed to the reification and alienation of humanity and social relationships. In BFN law has a more positive and prominent role in the legitimation process and as a mechanism of integration. In TAC, that place was fundamentally taken by morality.

[10] I develop the principle of democracy in the third section of this part of the Chapter.

traditional context only from them comes the legitimacy of the constitutional state. In BFN the rule of law is conceptually bound to democracy. This means that legitimate law ought to protect private autonomy, by legally recognizing basic rights, which are immanent components of modern legal[11] orders and the principle of *Popular Sovereignty* (public autonomy). Henceforth, this priority of private and public autonomy shows that only from the point of view of the participants a modern legal order obtains its legitimacy and its *Validity* [*Geltung*].

Nevertheless, legal orders not only refer to this normative dimension and the point of view of the participants. There is an additional feature of law which is its enforceability. This means that the rights of a legal order are enforced by the state. Habermas endorses this position and in *The Tanner Lectures on Human Values* he asserts that 'Legal norms borrow their binding force from the government's potential for sanctions'. (246) Of course, it is true that the subjects can respect law in general and particular rights due to their substantive rationality and because this citizens might recognize themselves as the authors of the legal norms that rule their political life. However, law functions in the 'stabilization of behavioural expectations' and when the recognition of its rationality does not motivate external actions (for example a person can recognize law's rationality but that does not necessarily motivate her) or someone does not recognize law's rationality (for example, someone who feels excluded from the political process) the participants normally assume the perspective of the observer and obey law because there is a threat of sanctions. In other words, the subject adopts an instrumental perspective towards the legal order and she calculates the cons and pros of respecting a norm. This aspect refers to the *Facticity* [*Faktizität*] of law. The latter concept and the notion of V*alidity* shape the tension that frames the discussion on BFN.

If one follows the origins of the tension between *Facticity* and *Validity* Kant's name comes to the fore. According to Rehg in the introduction of the English translation of BFN, the distinction between *Facticity* and *Validity* is heavily 'indebted to Immanuel Kant's concept of legitimacy'. (1996: xi) To my mind, Rehg's assertion is completely correct. For instance, in his political and legal theory Kant establishes an internal relation between the coercive character of law (*Facticity*) and its function guaranteeing freedom (V*alidity*). This is because the definition of right entails 'an authorization to coerce someone who infringes upon it'.

---

[11] The recognition of fundamental rights is an immanent feature of modern law. This thesis does not only apply to Habermas's normative theory of legal and political legitimacy. Rather, as I show in the following section of this part, even it can be found in more positivistic approaches.

(6: 231) In the *Rechtslehre* this coercive character of law is a necessary condition for civilization (Reiss, 1991: 26) since without this enforceability the citizens' freedom would not be protected. This enforceability is combined with the moral nature of legal norms or as Kant asserts in the *Feyerabend Lectures* 'right is the subset of moral norms that are also coercible norms'. (27: 1327) This clearly shows that Kant's political theory can be examined in light of the tension between *Validity* and *Facticity*. With some modifications, Habermas endorses this view of legitimacy as well. (IO, 255)[12] He claims that basic rights – or human rights – have a *'Janus-Faced'* nature. On the one hand, they have the same scope – they apply to all human beings – and demand the post-conventional justification which moral norms do. On the other hand, they are enforceable by the state – or by international institutions. (IO, 118; Flynn, 2003: 433) Henceforth, both Habermas and Kant accept the view that law is divided into the normative dimension of *Validity* and the functional dimension of *Facticity* (enforceability).

Now, I want to show that this connection between these authors is deeper and subtler. This is because the tension is conceptually related to the notion of communicative reason which is developed in TAC and also in the program of Discourse Ethics. In this respect, Habermas asserts that 'The theory of communicative action already absorbs the tension between facticity and validity into its fundamental concepts'. (BFN, 8) Following Habermas's terminology, the concept of communicative reason has a *'Janus-faced'* nature. On the one hand, it refers to counterfactual idealizations that ground the notion of rational understanding [*Verständigung*]. On the other hand, it has a sociological role because it is actually effective for processes of social integration. These features of the concept of communicative action show that it 'is the origin of the tension between "facticity" and "validity" that structures Habermas's BFN'. (Baynes, 2016: 132) Now, as I have shown in Chapter One, the concept of communicative reason has an internal relationship with Kant's account of pure practical reason. Combining both, the claim that the notion of communicative reason is the proximate ground of the tension between *Facticity* and *Validity*, and that communicative reason has a close kinship with Kant's concept of moral reason, I draw the conclusion that the tension is connected with Kant's notion of pure practical

---

[12] I say with some modifications because Habermas gives a more legalistic account of the 'system of rights' which is detrimental to the moral interpretation of private and public autonomy. (Larmore, 1995; Cohen, 1999; Forst, 2007; Flynn, 2003) Moreover he includes the principle of discourse in his reconstruction of legal and political legitimacy. This principle is a normative procedure and not a moral procedure. I examine these issues in more detail especially in the following sections and in the second part of this Chapter.

reason. This connection bears out my thesis that Habermas endorses a Kantian view concerning legitimacy.

Having described this connection, now it is important to show that the tension between *Facticity* and *Validity* appears in two forms: one 'internal' and another 'external' to law. The 'internal' tension is reflected in the intrinsic rationality of law (validity) and in the fact that legitimate orders can be coercible enforced (facticity). As I have shown, Kant establishes this tension in his *Rechtslehre*. The 'external' tension pertains to the claim of a political order to be legitimate (validity) *versus* the *de facto* recognition of its legitimacy by the citizens (facticity). In *An Answer to the Question: 'What is Enlightenment'?* And elsewhere (for example, see 6:313–14) Kant endorses this view. There he develops the idea that the only source of legitimacy resides in the autonomy of the subjects and cannot be determined by any external authority. As Rawls asserts regarding Kant's notion of autonomy 'Kant's main aim is to deepen and to justify Rousseau's idea that liberty is acting in accordance with a law that we give to ourselves'. (TJ, 225) Henceforth, in Kant's moral and political philosophy it is present the 'external' tension between the claim of a political order to be legitimate and the *de facto* acceptance of its members.

Up to now, I have developed the 'internal' tension in Habermas's theory which refers to the relationship between the normative rationality of law and its enforceability. In order to examine the 'external' tension, I follow Baynes who suggests that we look at 'Rawls's recent account of public reason' to illuminate this point. (Baynes, 2016: 133) Rawls proposes that the ground of legitimacy is a public notion of reason. This notion is based on the concepts of freedom and autonomy that pertain to the relationship between the citizens of a political community. The conception of public reason needs to be the focus of an *Overlapping Consensus* of comprehensive doctrines within the political culture, in order to guarantee the stability of a well-ordered society. (PL, 141)[13] In this regard, in the debate with Rawls, Habermas charges that the former collapses the distinction between justified acceptability

---

[13] According to Rawls, the political conception of *Political Liberalism* 'can be the focus of an overlapping consensus' (PL, 141) of comprehensive doctrines. This means that the citizens are going to endorse the political conception by means of their comprehensive doctrines. Before I continue developing the argument I need to describe Rawls's notion of 'comprehensive doctrine'. This concept has three different meanings: a) world views, religious and secular; b) actually existing moralities, or conceptions of the good; c) philosophical theories of one kind of another. (PL 2005: xviii, 12, 144-145) Now, Rawls's proposal is that the political conception (the conception of public reason) of *Political Liberalism* is a module that fits into the citizens' comprehensive doctrines and finds its support there.

and actual acceptance. (RPR, 122)[14] Moreover, Habermas argues that Rawls gives a purely instrumental role to the concept of *Overlapping Consensus* that instead of guaranteeing the intrinsic legitimacy of a legal order only warrants its acceptance by the citizens. The 'political ideals and values' that Rawls claims to be so merely in virtue of being shared, are, according to Habermas, only instrumentally valuable, because they contribute to stability.

In this way, for Habermas, Rawls is more concerned with the actual acceptance of a legal order than with its justified acceptability. Baynes agrees with this interpretation, and he states that Rawls seems to assume that the tension between justification and acceptability 'is sufficiently overcome within a liberal political culture'. (Baynes, 2016: 134) In other words, the demand for justification of the political conception – *Validity* in the external tension – and the issue of the actual acceptance of the citizens of it – *Facticity* in the external tension – seems to be taken for granted and solved beforehand in Rawls's theory.

Conversely, Habermas makes the 'external' tension explicit. (Baynes, 2016: 134) Specifically, he argues that the public reasons generated in processes of free and informal discourses of opinion-formation are channelled and transformed into legal and political decisions. In this way, the tension between the validity that a political and legal order demands and the *de facto* validity that it has from the point of view of the citizens it is something that needs to be achieved (more precisely a fact of the social world) and not an element to be assumed within the theory (as in Rawls). Therefore, in Habermas's theory the 'external' tension emerges in sharp relief.

In despite of this difference, it is important to mention that both Rawls's appeal to 'public reason' and Habermas's concept of 'communicative reason' as responses to the question of legitimacy rely on an intersubjectivistic account of Kant's notion of autonomy. (See Forst, 2011; See also Baynes 2016) This notion of agency in Rawls refers to the fundamental moral powers of the citizens and in Habermas to the presuppositions shared by those who act communicatively. In what follows, I continue examining Habermas's political theory and I focus on the *less* Kantian component of his theory, namely, the legal form. I study this component in light of the tension between *Validity* and *Facticity* that I have expounded in this section.

---

[14] For the sake of my argument, here I am going to follow Habermas's argument. That said, it is possible to claim that he might be mistaken in his interpretation of Rawls. The latter does not say that the only politically relevant feature of the ideas that are the content of public reason is that they are shared by all citizens. Habermas seems to think that in Rawls's theory the political ideas are constituted as reasonable because they are shared. That is not exactly what Rawls says in PL.

2. The legal form

In this section, I examine Habermas's Kantianism with a focus on the *less* Kantian component of his theory of legitimacy, namely, the legal form. (BFN, 82-131) The latter refers to an historical and functional input (BFN, 83, 111-113; 117-130, Finlayson, 2011: 7-11; Hedrick, 2010: 105-106) which is the result of two hundred years of Constitutional development. (BFN, 129; Hedrick, 2010: 111) According to Finlayson and Freyenhagen, the legal form has these three features: (I) it has been passed by a legally recognized body; (II) its observance is enforceable by various legitimate means; and (III) it has some intrinsic rationale or point independent of II (and of I) which pertains to the common good of the members of the legal community. (2011: 10) For these authors:

> This is what makes the legal system consistent, on the one hand, with the autonomy of individual citizens who can obey the law out of insight into its intrinsic rationale and, on the other, with the functional requirements of positive law, for when insight fails there remains the credible threat of prosecution and punishment to motivate compliance. (2011: 10)

Hence, the legal form combines the autonomy of the citizens (*Validity*) and the enforceability of legal norms (*Facticity*). Now, in light of the distinction between *Facticity* and *Validity*[15] I want to show that the normative element – i.e., the categorical imperative or the principle of discourse – has priority over the legal form. More precisely, my claim is that in a Kantian political theory, the legal form has an instrumental role protecting the normative notion of autonomy which in the political domain splits into private and public autonomy. (MW, 113) To examine this thesis, I focus on the tension between *Facticity* and *Validity* 'internal' to law that distinguishes between the immanent rationality and the coercive aspect of legal norms.[16] A good starting point to elucidate this issue is the tradition of Jurisprudence. This is because the legal medium has been the main subject of this discipline, where the point of view of the observer tends to prevail – i.e., Kelsen and Luhmann.[17] In what follows, I reconstruct some of the milestones of this tradition.

---

[15] See section 1 on this part of the Chapter.

[16] In section 1 of this part of the Chapter I discussed the 'internal' and 'external' tension between *Validity* and *Facticity* in Habermas.

[17] Luhmann's theory of law relies strongly on Kelsen's positivistic account and the former advances forward some theses that were developed in the latter. I develop this relationship in this section.

In its origins Jurisprudence endorses a Kantian notion of the legitimacy of law. Friedrich Carl Von Savigny develops the theory of 'subjective right' (private law) which establishes a conceptual relationship between the legitimacy of law and the protection of private autonomy. In his *System des heutigen Römischen Rechts*[18] the inviolability of the person grounds the legitimacy of law and establishes an area of independent rule for the free exercise of the individual will. Of course, this theory is supported and built from the point of view of Kant's political philosophy because in the *Rechtslehre* Kant asserts that the coercive dimension of legality is only justified and legitimate as long as it guarantees the protection of the freedom of the legal person. As a matter of fact, that is one of the intrinsic features of modern law: the protection of individual rights. In this respect, Habermas agrees and asserts that the legal form can be understood not only as positive and coercive law but also – in the Continental tradition – as 'subjective rights'. (Habermas, 2011: 285)

Nevertheless, in the development of the tradition of Jurisprudence other authors claim that the legitimacy of law does not depend on normative components understood in a deontological sense – i.e., individual freedom or the principle of discourse. Rather, they build the view that legitimacy is immanent to the legal system and they endorse Max Weber's thesis that legitimacy is based in its legality. (1978)[19] For instance, Hans Kelsen argues that the legitimacy of a legal order is based on 'the actual validity that political lawgivers confer on their decisions by coupling enacted law with penal norms'. (BFN, 86) Thus, legitimacy depends on the power that the *de facto* legislator poses and this power is an internal and self-referential feature of the legal order. Thus, in Kelsen's theory the connection of the moral person and her freedom vanishes. The ground of legitimacy is the power of the legislator – granted by the legal system – to enact norms backed by penal sanctions. Accordingly, Kelsen collapses the distinction between *Facticity* and *Validity*: the *validity* of law only depends on its *facticity*, namely, the will of the lawgiver. The focus on this aspect of law pertains to the notion of positive law. The latter and the coercive character of law define the main features of the *Facticity* of legal orders.

According to Habermas, the sociologist Niklas Luhmann radicalizes Kelsen's theory of law concerning the legal form. Luhmann argues, as Kelsen does, that the legitimacy of the legal order depends on its own internal features. (Luhmann, 2004: 76-141) Likewise, Luhmann

---

[18] The English translation of this book: *System of the Modern Roman Law.* (1980)
[19] Weber's thesis is that the legitimacy of a legal norm depends on the legality or the intra-legal rationality of the procedures that were performed to produce that norm. This view does not take into account the notion of self-legislation. Rather, it relies on a formalist notion of rationality.

goes beyond Kelsen because in the theory of the former the individuals are not part of the social system and of the legal system, rather, they are part of their environment. (Luhmann, 1996: 212) One of the consequences of this sociological *disenchantment* of law (BFN, 43) is that 'there is nothing to stop jurisprudence from conceiving rights along purely functionalist lines'. (BFN, 87) Therefore, by means of this methodology, law is completely uncoupled of all normative considerations. As far as the person is on the environment of the system, then it is not possible to base legitimacy in a Kantian concept of autonomy that certainly refers back to the moral or the legal person.

Habermas confronts the evolution that takes place in this tradition which ends in Kelsen's legal positivism and in Luhmann's sociology of systems. In this respect, the former argues that the coercive character of positive law 'is bound up with the demand for legitimation'. (IO, 254) This demand is satisfied by means of the Kantian notion of autonomy according to which the citizens are both *addressees* and *authors* of their legal order. Therefore, private autonomy (the legal persons as *addressees*) and public autonomy (the citizens as *authors*) are the grounds left that underpin legal and political legitimacy. On the one hand, the legal medium has a conceptual relationship with individual rights which protect private autonomy. (BFN, 128; IO, 256) And this is not only a conceptual relationship but also an historical one because 'Modern legal systems are constructed on the basis of individual rights'. (IO, 256) On the other hand, the legitimacy of legal orders is conceptually related to public autonomy (BFN, 89; IO, 253) because 'To be sure, the source of all legitimacy lies in the democratic law-making process, and this in turn calls on the principle of popular sovereignty'. (BFN, 89)

Now, let me briefly recall a reference from Habermas's MW which was stated in the introduction of this Chapter. (MW, 113) There, Habermas's argument is that when moral principles are embedded in the medium of law then moral freedom splits into public and private autonomy. (MW, 113) If we read this backwards, then it seems that behind private and public autonomy lays the normative principle as the root. Moreover, as I have shown in the previous paragraph, the legitimacy of law depends on the Kantian idea of self-legislation which splits into private and public autonomy. Finally, in Habermas's theory the normative component is the principle of discourse (D). Putting all these arguments together we see that the legitimacy of the legal form depends on private and public autonomy which refer back to the principle of discourse (D). Thus in this respect, even the coercive character of positive law is grounded in a normative component.

Thus, I end up this section claiming that despite the fact that the positivistic account of the legal form has tended to prevail in Jurisprudence and that in BFN Habermas aims to endorse the immanent approach to law – the idea that legitimacy is an intra-legal concept – if we distil the main contents of his concept of legitimacy what we have at bottom is the *modified* and *attenuated* Kantian notion of autonomy which has priority over the legal medium. In this respect, Habermas has to rule out the positivistic understanding of modern legal orders – which support an intra-legal notion of legitimacy – to give a proper account of his own position. (Larmore, 1995; Cohen, 1999; Forst, 2007; Flynn, 2003, 2011)

As a matter of fact, Habermas aims to avoid the shortcomings that affect this tradition which among other things put people in the environment of the legal order and uncouple law from normative components. (BFN, 86-87) Nevertheless, BFN definitely shows a more ambivalent relationship towards the developments in the tradition of Jurisprudence outlined above: on the one hand, legitimacy is related to normativity (an argument against positivism); on the other hand, Habermas aims to develop an intra-legal notion of legitimacy (an argument that coheres with positivism). That said, however, in works that were written right after BFN – i.e., fundamentally IO and MW – Habermas seems to admit the priority of normativity over the historical and functional input on his theory.

Having described the legal form and as far as the principle of discourse (D) was already developed in Chapter One, now we are in a position to present the main principle of political legitimacy in Habermas's theory, namely, the principle of democracy.

3. The principle of democracy

After the principle of discourse (D)[20] and the legal form[21] have been examined now I can introduce the chief principle of legitimacy in Habermas's theory, namely, the principle of democracy (BFN, 110) the role of which is to confer legitimating force on the legislative process. (BFN, 121) As I have shown, it arises from the interpenetration of the principle of discourse and the legal form. This principle states that:

> Only those statutes may claim legitimacy that can meet with the assent of all citizens in a discursive process of legislation that in turn has been legally constituted. (BFN, 110)

---

[20] See Chapter One.
[21] See the previous section.

The conceptual background for this principle is the idea that in a post-traditional order legal and political legitimacy can only be grounded when the *modified* and *attenuated* Kantian notion of autonomy – broached in (D) – is embedded in the legal medium of modern societies. (BFN, 128) Legitimacy cannot be based on a metaphysical world-view based on a parochial set of values and principles as a contextualist might argue. It cannot be founded on any sort of *Comprehensive Doctrine* in the Rawlsian sense of the term – i.e., world views, religious and secular; actually existing moralities, or conceptions of the good; philosophical theories of one kind of another. (PL, 12, 144-145) Moreover, it cannot be based on the self-interest of particular agents as the utilitarian approach proposes. Rather, theories of morality and legitimacy developed in the Kantian tradition are the most reasonable alternative to answer this issue because they 'hold out the promise of an *impartial* procedure for the justification and assessment of principles'. (TL, 241) In a post-traditional order this is the alternative left because only the ideas of autonomy and equality embedded in the medium of law furnish the procedures that ground the legitimacy of Constitutional Democracy. Kant and Habermas share the view that legitimacy is connected with normative principles that guarantee, due to their procedural character, *fair* and *impartial* results. This impartiality is connected to the Kantian notion of autonomy, because only the mutual recognition of the freedom of everybody guarantees *fair* and *impartial* decisions.

Certainly, these principles demand a social space in which they unfold. In §45 of *The Metaphysics of Morals,* Kant states that the political society [*Societas Civilis*] is 'a union of an aggregate of men under rightful laws'. (6: 313) Certainly, Habermas has a similar notion of the political community because his theory of Constitutional Democracy defines contents and procedures that guarantee the existence of an association ruled under legitimate laws. As I have shown, in a post-traditional order this legitimacy can only be ground on a Kantian notion of autonomy which pertains to the mutual recognition of equal rights among citizens. On §46 of MM Kant describes three rightful attributes of the citizens [*Cives*]. In light of them it is possible to understand Kant's concept of legal and political autonomy:

> Firstly, lawful *freedom* to obey no law other than that to which he has given his consent; secondly, civil *equality* in recognizing no-one among the people as superior to himself, unless it be someone whom he is just morally entitled to bind by law as the other is to bind him; and thirdly, the attribute of civil *independence* which allows him to owe his existence and sustenance not to the arbitrary will of anyone else

among the people, but purely to his own rights and powers as a member of the commonwealth (so that he may not, as a civil personality, be represented by anyone else in matters of right). (6: 314)

Habermas's endorses these Kantian components in his political theory. Above all, the concept of autonomy is at the core of Habermas's moral and political theories.[22] Now, this Kantian notion applied to the legal and political realm splits into private and public autonomy. (MW, 113) This idea is expressed in the thesis – which is the milestone of the project of radical democracy that connects Kant with Rousseau and Marx – which asserts that the normative weight of a legal order depends on the fact that the *addressees* of the norms are at the same time their *authors*. (BFN, 104, 126) In terms of the terminology of the modern enlightenment, democracy is based on the recognition of the private autonomy of the *bourgeois* and the public autonomy of the *citizen* (Maus, 1992: 216).

Moreover, the principle of democracy has some similarities with the chief principle of legitimacy developed in Kant's *Rechtslehre,* namely, the principle of Universal Right. This is because in both cases, a normative (or moral) principle is embedded in the medium of social space and historical time. In the *Metaphysics of Morals* Kant enunciates this principle which says:

> Every action which by itself or by its maxim enables the freedom of each individual's will to co-exist with the freedom of everyone else in accordance with a universal law is *right*. (6: 230)

On the principle of right Kant is applying the categorical imperative to the relationships between different freedoms. Accordingly, the move is from a moral principle to the social domain. (BFN, 83) Thus, Kant begins from the notion of autonomy that in Habermas becomes the core of the Constitutional state as well. In Habermas the place of Kant's principle of Right is occupied by the principle of democracy. In terms of the architectonic,

---

[22] In this thesis I have focused on the concept of autonomy. That said, however, the notion of equality also has a central place in Habermas's theory. This is because the notion of equality is at the centre of the legal medium. As Habermas asserts, basic rights are immanent to the legal form. (BFN, 83; IO, 256) In another formulation he states that 'To be sure, the establishment of the legal code as such already implies liberty rights that beget the status of legal persons and guarantee their integrity'. (BFN, 128) Now, immanent to the definition of basic rights is the idea of equality. These basic rights 'provide all persons with equal legal protection, an equal claim to a legal hearing, equality in the application of law, and thus equal treatment before the law. (BFN, 125) Thus, the existence of equal rights is an internal feature of the legal code. I think that the idea of equality can also be seen from the point of view of the principle of discourse because according to (D) everyone counts equally as a participant in discourses of normative justification.

in both theories a normative component is embedded in the medium of social space and historical time. Nonetheless, Kant's and Habermas's principles of legitimacy unfold from different normative components: on the one hand, Kant's principle of law derives from the categorical imperative which is a moral principle; on the other hand, Habermas's principle of democracy depends on the principle of discourse (D) that is different from (U) which is 'the central principle of the discourse theory of morality'. (Finlayson, 2011: 9) By means of this strategy, Habermas claims that he has avoided the fully Kantian strategy in which there is a monistic reconstruction of law and morality. (Habermas, 2011: 298)

Due to this distinction proposed by Habermas, it seems that the justification of the democratic principle is immanent to the constitutional state and does not need external moral support. Despite differences of emphasis in their interpretations of Habermas's theory several sympathetic critics have nonetheless claimed that he cannot give this justification and legitimacy needs to be underpinned by a moral content. (Larmore, 1995; Cohen, 1999; Forst, 2007; Flynn, 2003) In general terms I agree with them. However, I prefer to limit my claim to state that the normative component of Habermas's theory is a reworking of Kant's notion of pure practical reason, which not necessarily has a moral nature. That said, however, Habermas has tried to avoid this interpretation of his political theory and he has stressed that he offers an immanent reconstruction of legal and political legitimacy, at least in BFN.[23] One of the advantages of this strategy is that Habermas has good arguments against the Hegelian inspired charge that his theory lacks the institutional dimension that any political theory needs to include. Also, he can rebut the criticism that his approach is more or less a renewed version of natural law because it would be proposing a monistic construction of law and morality. Likewise, as long as it includes a normative component, Habermas avoids the shortcomings that in his eyes the positivistic approach has, namely, the collapsing of the distinction between *Facticity* and *Validity* and

---

[23] There are at least two reasons why Habermas avoids developing a Kantian argument in the justification of legitimacy. On the one hand, this is explained by his strategy to give an equal weight to the principle of popular sovereignty in front of the rule of law. (Flynn, 2013) If there were a strong Kantian argument, then that would suppose the primacy of a liberal 'system of rights' prior to popular sovereignty. Interestingly enough, at the end that seems to be the case in Habermas's theory. On the other hand, to certain extent Habermas *has not been allowed* to admit the degree of his Kantianism in his notion of legitimacy because of a particular context of ideas. Not only in the sociological tradition – i.e., Weber, Luhmann, Teubner, Willke – but also in jurisprudence – i.e., Kelsen – and in philosophy – i.e., Foucault, Derrida, Agamben – there is a strong rejection to articulate Kantian (or moral) arguments as the justificatory grounds of legitimacy.

the exclusion of the point of view of the participants by means of putting them in the environment of the legal system – i.e., Kelsen and Luhmann.[24]

Consequently, Habermas accomplishes the philosophical *Aufhebung* of Kant. On the one hand, the political theory of the former incorporates the post-conventional level of justification which pertains to a modified Kantian notion of autonomy at the centre of his concept of legitimacy. However, I say 'modified' because this notion of normative justification relies on a *dialogical* notion of reason, whereas Kant's concept of autonomy is framed in a *monological* notion of pure practical reason.[25] On the other hand, Habermas includes the legal form which refers to an historical and functional input that is not elaborated in Kant's *Rechtslehre*. At least, it is not broadly elaborated in Kant as it is in BFN. In summary, I have broached the essential stage in which Kant's and Habermas's theories are located and the relationship that they establish between normative components and the social dimension of institutions and practices. In part two, I discuss whether and to what extent Kant and Habermas succeed or fail in their aims to find the proper balance between individual freedom and collective self-determination.

II. The co-originality between private and public autonomy

The previous part shows that the essential components of Habermas's political theory have a Kantian pedigree. In the case of the principle of democracy, the principle of discourse is embedded in the legal medium of modern societies. This normative component pertains to the idea that the persons grant to each other the equal status of free citizens in discourse. In this regard, the private autonomy of the particular subject is co-original with the private autonomy of everybody else. Moreover, these autonomies are actualized in processes of mutual understanding which refer to the notion of public autonomy. Habermas claims that in his theory the legitimacy depends ultimately on both the private and the public autonomy of the consociates of a legal and political order. As he asserts,

> Human rights and the principle of popular sovereignty still constitute the sole ideas that can justify modern law. These two ideas represent the precipitate left behind, so to speak, once the normative substance of an ethos embedded in religious and

---

[24] See section 2 of this part.

[25] As I have shown in Chapter One, here I side with Habermas against McMahon, in the sense that effectively the dialogical approach of the former differs in important ways to the monological ontology of Kant.

metaphysical traditions has been forced through the filter of posttraditional justification. (BFN, 99)

The development of Habermas's argument shows that at the core of legitimacy there is a complementary notion of private and public autonomy. In what follows, I focus on the concrete relationship between these components. According to Habermas, he gives equal weight to both principles in the co-originality thesis. In his eyes, this is something that no one has successfully done before him. (BFN, 84, 100-101; IO, 258) In this part of the Chapter, I examine the co-originality thesis in detail. At the beginning I describe in broader brush-strokes the stage that in terms of the history of ideas explains the question of the co-originality. Following Constants's distinction (1988), the quarrel between the ancients – who gave priority to the political autonomy of the citizens – and the moderns – who gave priority to the basic rights of the individual – still now shapes the current debates concerning the basic foundations of constitutional democracy. Kant has been regarded as a prominent figure that represents the side of the moderns which gives priority to private autonomy and it is detrimental to popular sovereignty. Nevertheless, in what follows I argue that Kant gives a more balanced weight to both principles (or at least he recognizes the significance of public autonomy). Once this preparatory stage is settled I examine Habermas's thesis of the co-originality.

The argument unfolds in three steps. First, I describe the essential elements of the debate between ancients and moderns – or republicans and liberals – which in modern political philosophy is represented by Rousseau and Kant. On the one hand, Rousseau seems to grant more weight to the principle of popular sovereignty or public autonomy. According to the view endorsed by Habermas, Kant would be emphasizing the rule of law and the principle of Universal Right. Here, I cannot examine Rousseau's views on this issue. Rather, for the sake of my argument I focus the attention on Kant. I argue that it is a plausible interpretation of Kant's *Rechtslehre* that he not only gives priority to the principle of Universal Right but also to political rights, by means of his concept of self-legislation. Habermas argues in BFN that Kant gives priority to liberal rights (BFN, 100), but elsewhere he also claims that it is possible to read Kant from a republican point of view. (MW, 113) If it is possible to read Kant's political theory as granting a more balanced relationship between private autonomy and public autonomy, then it is more plausible to claim that Habermas endorses a Kantian view in his version of the co-originality (1).

Secondly, I describe Habermas's co-originality thesis and I examine the 'system of rights' that comprises private and public autonomy. At the beginning, this section is more descriptive than argumentative. Later, once I have described the 'system of rights', then I expound Habermas's co-originality thesis. Finally, I briefly discuss some initial arguments that could challenge Habermas's thesis (2).

In the last section, I examine in detail Habermas's co-originality thesis in light of the assessments that some influential critics develop of his thesis. At the end I argue that despite what Habermas says on the contrary, the thesis of the co-originality in his theory is based on a Kantian methodology – i.e., the two stage construction of the system of rights – and on the priority of a Kantian component – i.e., the idea of citizens that grant to each other equal rights to freedom (3).

1. Private and public autonomy in Kant's *Rechtslehre*

The names of Rousseau and Kant are the core of the modern self-understanding of the principles that frame Constitutional Democracy. Both philosophers discussed in detail the central notions that until now broach the question for legal and political legitimacy, namely, the rule of law and popular sovereignty; private and public autonomy; individual freedom and collective self-determination. However, their contributions go beyond because they tried to ground the intuition that both principles have an equal weight:

> That the idea of human rights, which is expressed in the right to equal individual liberties, must neither be merely imposed on the sovereign legislator as an external barrier, nor be instrumentalized as a functional requisite for legislative goals. (IO, 259)

Henceforth, Kant and Rousseau aim to give a foundation to the idea that Habermas's has labeled as the co-originality between private and public autonomy. In simple terms, this notion means that both principles have to be in a balanced relationship, both having an equal weight in the theory of legitimacy. However, Habermas claims that neither Kant nor Rousseau were able to find the balance between these principles. According to the widely accepted interpretation (Baynes, 2016: 136) – which Habermas endorses –, Kant gives priority to individual rights over popular sovereignty and Rousseau does the contrary. (BFN, 100: IO, 259) To certain extent, this dilemma can be seen in the difference between the approaches of natural law and legal positivism. (Baynes, 2016: 136) The first tradition

gives emphasis to individual freedom – i.e., Locke and Kant –, the second locates the weight of legitimacy on the will of the lawgiver – i.e., Kelsen (albeit, in the positivistic interpretation it does matter whether the will is democratic or not).

In light of the distinction between private and public autonomy it is possible to distinguish the following doctrines as well: Liberalism and Republicanism. On the one hand, the latter 'goes back to Aristotle and the political humanism of the Renaissance', and 'has always given the public autonomy of citizens priority over the prepolitical liberties of private persons'. (IO, 258) On the other hand, Liberalism is related to philosophers like John Locke and this doctrine 'has invoked the danger of tyrannical majorities and postulated the priority of human rights'. (IO, 258) Habermas argues that Kant should be seen as a liberal and Rousseau as a republican. (BFN, 100)[26] Habermas, though, claims that he endorses a Kantian Republicanism. (MW, 113) To my mind this hints that there are at least two possible readings of Kant's *Rechtslehre*: it can be seen as a classical liberal doctrine or a liberal-republican doctrine. In the first interpretation, Kant has been seen belonging to the tradition of natural law. After all, he is the philosopher of the modern enlightenment which advocates for the primacy of the autonomy of the subject over any form of tradition and perhaps even beyond the collective self-determination of the political community (let us remember that the emphasis on traditions and popular sovereignty pertains to the self-understanding of the *Republican* tradition). In this regard:

> Insofar as Kant's argument for the establishment of civil society (or the state) relies solely on the Universal Principle of Right, which guarantees equal subjective liberty for all, the notion of collective self-determination is subordinated to a moral principle. (Baynes, 2016: 136)

Thus, in Kant's theory the notion of private autonomy which is detrimental to public autonomy has the upper hand. Nevertheless, in this section I want to show that Kant introduces the idea of public autonomy and that there is at least something of a balance between the latter and private autonomy. In this respect, it is also possible to read Kant's engagement with the modern enlightenment in the sense that he not only incorporates the principle of private autonomy but also the principle of public autonomy. Now, before I

---

[26] In this section, I do not discuss Habermas's claim concerning Rousseau, mainly because the republicanism of the latter is not at the focus on the discussion that I address here. Rather, I examine the notion of the co-originality that can be found in Kant's Rechtslehre, because it is the Kantian component in Habermas the issue at stake in this Chapter.

keep developing my claim that Kant gives a balanced weight to both principles I want to summarize the ways in which the different traditions of political philosophy have emphasized private and public autonomy. I think that these distinctions can be helpful to assess Kant's and Habermas's reconstructions of the co-originality. For the sake of my argument, I distinguish only three main political doctrines:[27]

> a. *Classical Liberalism:* in this tradition the constitutional expert gives primacy to private autonomy and public autonomy gets overshadowed. When this expert frames the 'system of rights' she focuses on individual rights. Among other things, this solves the danger of the tyrannical majorities.

> b. *Republicanism:* in this tradition the sovereign (the united will of the people) decides to frame and shape the constitution according to its own point of view. In other words, the people are the only source that establishes the 'system of rights'. Among other things, this precludes the paternalism of natural law.

> c. *Kantian Republicanism*: in this tradition the constitutional expert gives primacy to public and private autonomy when she is framing the 'system of rights'. In other words, this expert offers a 'system of rights' which guarantee both individual and political rights which in a second moment gain its actual shape by means of public autonomy.

Of course, Rousseau belongs to the Republican tradition, and Kant has been seen belonging to *Classical Liberalism.* That said, however, I think that this is a mistaken reading of Kant's political philosophy because he takes both private and public autonomy into account. As I have shown, Habermas defines his theory as Kantian-Republican. In the same spirit, Kant says that 'The civil constitution in every state shall be republican'. (8:349) This supports my claim that Habermas endorses a Kantian view of legitimacy because Kant himself already incorporated a Republican component into his theory. Nonetheless, before I move into the discussion of the balance between private and public autonomy in Habermas's theory I need to justify why Kant could be seen as giving a balanced weight to private and public autonomy.

I want to start my argument by claiming that Kant's political philosophy can be understood in light of the claim that 'Nobody can be free at the expense of anybody else's freedom'.

---

[27] Of course, others might distinguish other doctrines.

(MW, 113) I read this passage as saying that the only source of legitimacy is the autonomy of the subjects in terms of the protection of their private rights but also in terms of their public autonomy as citizens of the political association. This interpretation coheres with Kant's theory. For example in the *Metaphysics of Morals* he says,

> The legislative authority can belong only to the united will of the people […] Therefore only the concurring and united will of all, insofar as each decides the same thing for all and for each, and so only the general united will of the people, can be legislative. (6:313–14)

It is still necessary to clarify the meaning that the ideas of popular sovereignty and of the general will have in Kant. He argues that there is a priority of the political sovereign which it is not everybody that belongs to the *demos* but normally a particular person or a group of people whom are able to reconstruct the collective will. As a matter of fact, Kant thinks that an elected legislator is the best form of government. (8: 353) In this way, his political philosophy has an ambivalent nature because this idea of representation seems to betray its concept of autonomy. To certain extent, this is a consequence of Kant's broader notion of pure practical reason. In his theory, pure practical reason is not based on the medium of social institutions and practices (as it is in Habermas). Rather, it is based on the monological structure of subjectivity. Henceforth, it is possible to understand that in this political theory an elected representative can successfully embody the collective will of the people. In this rationalist view, if this representative is correctly appointed then she is the best candidate to reconstruct the collective will (probably better even than the *demos* itself). In this respect, the representative is constrained to:

> Give his laws in such a way that they could have arisen from the united will of a whole people and to regard each subject, insofar as he wants to be a citizen, as if he has joined in voting for such a will. (8:297)

In Habermas's view of legitimacy the introduction of the principle of discourse demands that the practices of self-legislation cannot be performed in isolation in which one agent would reconstruct the collective will. Rather, (D) reconstructs the practice of argumentation that takes place between agents and this eminently refers to a cooperative process. Even if one agent is sufficiently altruistic, pluralist and moved by a post-conventional morality – in Kohlberg's sense – that allow her – and others – to expect that

she can properly reconstruct the collective will, Discourse Ethics claims that a rational agreement only can be reached communicatively and intersubjectively and not in isolation.

In general, I agree with this interpretation of the main difference between Kant and Habermas. In Chapter One, I side with the latter in the sense that Discourse Ethics proposes a dialogical principle that differs in important ways to the one proposed by Kant in the categorical imperative. Conversely, as far as these principles furnish their political theories, then the differences in the moral theories pertain to their notions of legitimacy. Roughly speaking, Kant seems to develop a *monological* account of legitimacy – the collective will can be reconstructed by a particular subject – and Habermas broaches a *dialogical* one – the collective will can (only) be reconstructed in processes mutual understanding.

Bearing in mind that it is true that Kant broaches this view of the representation of the collective will, I would like to end this section emphasizing that in one way or another he introduces the idea of public autonomy. Therefore, it is mistaken to take Kant as simply giving priority to the idea of individual freedom and dismissing collective self-determination. If I am correct, then Kant at least allocates a balanced weight to both private and public autonomy. Kenneth Baynes leaves rooms for this interpretation of Kant's theory because he argues that Kant has been seen connected to the tradition of the social contract or to the tradition of natural law. (Baynes, 2016: 135-136) The first version of Kant recognizes a place to collective self-determination whereas the second only to the idea of individual freedom. I think that he grants – certainly from a monological methodology – a place in his political theory both to private rights and a form of political rights. In this respect, Kant explicitly argues that 'A person is subject to no laws other than those that he (either alone or at least jointly with others) gives to himself'. (Kant, 1991: 50, In Baynes, 2016: 136) Thus, it is implicit that as long as the political community is shaped by a plurality of persons, then it self-legislation implies a dimension of public autonomy. The quite simple but powerful idea of the political autonomy of the *demos* which pertains to the identity of the citizens as *authors* and *addresses* of the legal order is essential in the architecture of BFN and also in Kant's concept of legitimacy. (BFN, 39, 120, 126) From this interpretation, it is not fair to argue that Kant gives only an emphasis to the idea of individual freedom, not taking into account at all the idea of collective self-determination. As I will show in the following section, individual freedom and collective self-legislation are

essential components that explain Habermas's discourse theory of law and democracy as well.

To sum up, Kant's emphasis in the idea of self-legislation, implies the recognition of both private and public autonomy in the same token. Thus, the Kantian theory of legitimacy could be read in the sense that private and public autonomy are framed in a 'system of rights' that guarantees equal freedoms and that gains concrete shape in a second moment throughout the collective self-determination of the *demos*. As I am going to show, this is the idea of a 'two-stage' methodology on the reconstruction of the principles of the constitutional state. To my mind, this is a reasonable reading of Kant's *Rechtslehre* that goes together with other claims that he could not make at his historical time. For example, his notions of human dignity and self-legislation imply the idea that women and minorities have individual and civic rights. Albeit, Kant did not grant these rights to these groups, certainly for more historical than conceptual reasons. In this regard, in philosophy it is a necessary practice to confront the spirit of a particular philosopher with its letter.[28] After I have broached this interpretation of Kant's notion of legitimacy, in what follows I examine Habermas's co-originality thesis (Section 2). At the end (Section 3), I conclude that Habermas's thesis is Kantian not only in terms of its contents – i.e., the priority of the idea of autonomy – but also in terms of its methodology – i.e., the 'two-stage' reconstruction.

2. Habermas's co-originality thesis

Habermas argues that he endorses Kantian-Republicanism. (MW, 113) The two components of this label pertain to a theory of legitimacy that aims to grant to individual freedom and collective self-determination equal weight. In this section, I examine the interpenetration [*Verschränkung*] between the principle of discourse (D) and the legal form not from the point of view of its logic and its connections with Kant which I did in the previous part of this Chapter. Rather, I analyse it from the point of view of its outcome: the 'system of rights' which protects and shape private and public autonomy. Thus, this section is more descriptive at the beginning. After that, I introduce the dilemma that is going to be the issue at the final section, namely, whether and to what extent Habermas successfully gives equal weight to private and public autonomy.

---

[28] According to Pinkard 'In the hothouse intellectual ferment of Germany in the 1790s and the first seven years of the 1800s […] a dominant metaphor emerged of 'completing the Kantian philosophy by reconstructing it in terms of its 'spirit', not its 'letter''. (1999)

To begin his thesis of co-originality Habermas argues that in the interpenetration [*Verschränkung*]:

> One begins by applying the discourse principle to the general right to liberties – a right constitutive for the legal form as such – and ends by legally institutionalizing the conditions for a discursive exercise of political autonomy. By means of this political autonomy, the private autonomy that was first abstractly posited can retroactively assume an elaborated legal shape. (BFN, 121)

From the interpenetration, a 'system of rights' arises. In Habermas's terms, in a first moment, these rights can be divided into three categories. (BFN, 122)

> 1. Basic rights that result from the politically autonomous elaboration of the right to the greatest possible measure of equal individual liberties

> 2. Basic rights that result from the politically autonomous elaboration of the status of a member in a voluntary association of consociates under law.

> 3. Basic rights that result immediately from the actionability of rights and from the politically autonomous elaboration of individual legal protection.

These first three categories of rights guarantee the private autonomy of consociates under the rule of law, or the recognition of citizens as equal *addresses* of laws. (BFN, 122) In this respect, these are the rights that the liberal point of view emphasizes. These sets of rights define a domain of autonomy in which the citizens do not have to justify their preferences of action and as in Kant's *Rechtslehre* they can act freely. From the point of view of TAC the significance of these rights is that a person could participate in the process of mutual understanding [*Verständigung*]. Nevertheless, it also pertains to the possibility of someone withdrawing from this process. (BFN, 119-120) After these rights are framed, a fourth category of rights which protects public autonomy arises:

> 4. Basic rights to equal opportunities to participate in processes of opinion-and will- formation in which citizens exercise their political autonomy and through which they generate legitimate law.

In light of these rights, the subjects become *authors* of their legal order. Therefore, this set guarantees public autonomy. This step is when the principle of democracy comes fully into

place, as the core of the system of rights. This category of rights which broaches the idea of citizens as authors of law, is reflexively applied to the basic rights identified in (1) through (3) and certainly also to itself – this is because it is always possible to legislate concerning the rules of political participation. (IO, 255) Thus, the 'system of rights' gains its full legitimacy.

Therefore, Habermas's theory grants a central place to the liberties of the ancients (public autonomy) and to the liberties of the moderns (private autonomy). In BFN private rights are conceptually related to political rights. Let us recall here that for Habermas's there is a conceptual and internal, and not a merely contingent, relation between the rule of law and democracy. (IO, 253) Thus, there is not democracy without rights and not rights without democracy.

Finally, a fifth category of rights is introduced which guarantees living conditions that make possible the exercise of private and public autonomy:

> 5. Basic rights to the provision of living conditions that are socially, technologically, and ecologically safeguarded, insofar as the current circumstances make this necessary if citizens are to have equal opportunities to utilize the civil rights listen in (1) through (4).

This set of rights shows that Habermas is not only recognizing private autonomy and public autonomy but also social rights. It is relevant to notice that in historical terms, in the model of the welfare state this component has come to the fore. Although Habermas recognizes this group of rights, it does not seem that they are central to his concept of legitimacy. In this respect, and bearing in mind that he always takes into account the point of view of the participants, this last group of rights should have also a central place in the concept of legitimacy. There are at least two reasons to claim that a proper concept of legitimacy should pay more attention to this component. First, it is not possible to protect private and public autonomy if it is not also guaranteed the *social autonomy* of the agents. Secondly, the 'external' tension between *Facticity* and *Validity*, namely, the claim of legitimacy of a legal order and the de facto recognition of its members can only be assured if certain minimum level of welfare and even of social equality is achieved.

Nevertheless, I cannot devote more attention to this issue here. Rather, for the sake of my argument the question that requires further elucidation is on which grounds Habermas can

convincingly argue that 'the logical genesis of rights comprises a circular process in which the legal code, or legal form, and the mechanism for producing legitimate law – hence the democratic principle – are co-originally constituted'. (BFN, 122)It is important to carefully discuss this claim because Rawls and others (RH, 81; Larmore, 1995: 612; Cohen, 1999; Forst, 2011) have argued that Habermas fails in his attempt to give an equal weight to the principles of private and public autonomy.

In this way, the rights from (1) to (4) gain their legitimacy when (4) is applied. The 'system of rights' is initially introduced abstractly, and then it gains complete shape and substance throughout the democratic process. As Rawls's reply to Habermas's critique (RH) shows, the relationship between private and public autonomy in the latter can have two interpretations. (Forst, 2007: 109) On the one hand, it seems that the genesis of legitimacy stems from public autonomy: the legitimacy of the whole system depends on the democratic practice. Then, what is the place of private autonomy? In this interpretation private autonomy looks as though it has an instrumental role to generate the conditions of possibility of public autonomy and the latter appears as the final ground of legitimacy. On the other hand, the introduction of an abstract 'system of rights' at the first stage pertains to the recognition of basic rights that protect autonomy and that have moral priority over the political self-determination of the *demos*. Thus, we are confronted with two possible readings of Habermas's theory.

That said, however, Habermas maintains that individual autonomy depends on collective self-determination in the same degree that the latter depends on the former. This claim can be examined from the normative point view of the principle of discourse (D) which refers to both private and public autonomy: the autonomy of the subject is conjoined with the autonomy of everybody else as participants in rational discourses. This balance can also be seen in light of the legal form. This is because 'The legal medium establishes a logical relation between rights-bearing individuals and constitutional law-making'. (Hedrick, 2010: 115) In other words, the internal structure of law can be explained in two ways: rights bearing individuals constitute a practice of legitimate law-making – then we go from private autonomy to public autonomy –; and citizens mutually grant to one another symmetrical rights throughout their law-making practice – then we go from public autonomy to private autonomy. In what follows, I keep framing and sharpening this discussion relying on the work of some more or less sympathetic critics of Habermas. I conclude that at bottom Habermas's theory grants a normative priority to the 'system of rights' that guarantee

autonomy over the actual carrying out of the practice of collective self-determination of the political association.

3. Critical assessments of Habermas's co-originality thesis

According to Habermas the tradition of political philosophy has been unable to give a proper balance to the principles of private and popular sovereignty. (BFN, 84, 100-101) Not only in BFN but also in IO he claims that:

> To be sure, political philosophy has never really been able to strike a balance between popular sovereignty and human rights, or between the "freedom of the ancients" and the "freedom of the moderns". (IO, 258)

After centuries of discussion of this issue, and bearing in mind that the most important political philosophers devoted the best of their efforts to elucidate this question Habermas's claim is at least provocative. The hidden implication of this statement is that only in Habermas's co-originality thesis it is possible to find the correct balance between both principles. In this final section, I aim to assess his thesis of co-originality from the point of view of his critics. Here I begin from a claim made by Habermas which contains the key move in his theory. In BFN he asserts:

> Nothing is given prior to the citizen's practice of self-determination other than the discourse principle, which is built into the conditions of communicative association in general, and the legal medium as such'. (BFN, 128)

The heuristic value of this statement is that although nothing is giving prior to public autonomy, there are two elements that are given prior to the collective self-determination of the citizens: the principle of discourse and the legal form. In the first part of this Chapter I have shown that both components are framed on a Kantian notion of autonomy. Therefore, it seems reasonable to argue that the collective self-determination of the citizens can only be conceived within this Kantian frame, namely, the Kantian conception of pure practical reason.

Some of the most relevant critics of Habermas draw similar conclusions: namely that there is a priority of the Kantian component in Habermas's co-originality thesis. Charles Larmore argues that Habermas mistakenly gives priority to public autonomy over private autonomy. According to Larmore, this is problematic because private autonomy is actually

the ultimate basis on which our political life is organized. Larmore asserts that Habermas gives priority to public autonomy, by means of the emphasis that the latter gives to the notion of communicative reason and discourse, and what Habermas should do instead is to acknowledge that private autonomy has priority. For Larmore public autonomy rests on a more fundamental Kantian-liberal principle of respects for persons. (1995) In the terms of Habermas's theory, the principle of discourse (D) presupposes a prior Kantian principle of recognition of persons as *free* and *equal*.

Rainer Forst also argues for the priority of a normative component. In his view, a normative pre-political core acts as a foundation which has more weight than popular sovereignty. This is because the framers of constitutions and the citizens are moral persons. (Habermas, 2011: 298) In his own theoretical work Forst proposes *The Right to Justification* which solves the moral deficit of Habermas's notion of legitimacy. (the moral deficit especially in BFN, See Forst, 2007) Discussing the Habermas–Rawls debate Forst asserts that: '[Habermas] still does not get around the problem of assigning these rights an antecent status, as he appears to do in the reference to the two-stage reconstruction'. (Forst, 2011: 173) In this regard, Forst correctly illuminates an ambiguity in Habermas's theory. On the one hand, BFN builds an immanent argument for the justification of a 'system of rights'. On the other hand, in the debate with Rawls, Habermas gives priority – an antecedent status – to a moral Kantian content in the justification of the 'system of rights'.

Joshua Cohen criticizes Habermas for the moral deficit of his theory as well. He suggest that 'the discourse principle […] appears to rely on a highly generic account of reasons – not an account restricted to political argument in a democracy of equal members'. (Cohen, 1999: 395) This principle does not give priority to equal liberties which is the essential ground of democracy. According to Cohen, democracy cannot be shaped without the recognition of substantive Kantian principles as the existence of citizens as *free* and *equal*. (Cohen, 1994) This is not guaranteed in Habermas's discourse theory of Constitutional Democracy. From the point of view of moral foundationalism the justification of the principle of discourse demands the recognition of moral substantive elements at the beginning.

In summary, these critics on the one hand claim that Habermas is not giving the proper place to the normative component which should be set at the centre of the co-originality

thesis. In other words, Habermas does not seem to acknowledge the inner structures of his own theory – i.e., Larmore and Forst. On the other hand, a stronger criticism is that Habermas's reconstruction does not guarantee the protection of this Kantian core, which is essential for Constitutional Democracy – i.e., Cohen.

The idea of a 'two-stage' reconstruction is developed in BFN when Habermas maintains that by means of the interpenetration between (D) and the legal form an abstract 'system of rights' of 'unsaturated placeholders' (BFN, 125) unfolds in a first stage. In the second stage the 'system of rights' gains its actual shape by means of the exercise of *Popular Sovereignty*. This idea is also central in the debate with Rawls (Habermas, 2011: 296; RH, 63, 64, 76)[29] and in the article *Constitutional Democracy, A paradoxical Union of Contradictory Principles* (2001). In the latter Habermas asserts:

> The first stage involves the conceptual explication of the language of individual rights in which the shared practice of a self-determining association of free and equal citizens can express itself – rights, thus, in which alone the principle of popular sovereignty can be embodied. The second stage involves the realization of this principle through the exercise, the actual carrying out, of this practice. (Habermas, 2001: 778)

In this way, it is possible to observe that the recognition of citizens as *free* and *equal*, by means of a conceptual reconstruction at the first stage of the 'system of rights', has priority. Nevertheless, these rights leave room for the idea of the collective self-determination of the *demos* which becomes a concrete practice in the second stage. In that further step, the principle of *Popular Sovereignty* is performed throughout the exercise, the actual carrying out, of this practice. According to Forst and Flynn, the second stage is not possible unless citizens bring valid moral norms to bear into legal and political practical discourses, and this shows that the justification of basic rights cannot be fully immanent to law.

Another point which deserves attention is the fact that Habermas's political theory gives primacy to moral reasons over other considerations – pragmatic and ethical-political reasons – whenever popular sovereignty is performed. (BFN, 103, 108, 113; IO, 42-43; Flynn, 2003: 434) Furthermore, legitimatite laws must 'harmonize with the moral principles of justice and solidarity'. (BFN, 99, 155) According to Finlayson this refers to the *Moral*

---

[29] In RH Rawls asserts: 'The idea of a two-stage construction is implicit in the summary argument. (RPR–43–4)' (76)

Permissibility Constrain in legal and political discourses. (Finlayson, 2011: 12, 2016a: 12, 14) In other words, no legitimate law may violate a valid normal norm. Once again, I draw the conclusion that not only in terms of the basic frame of the theory, but also in terms of its contents there is a priority of a Kantian element, namely, the notion of moral reason understood this time in terms of the post-conventional structures of justification contained in the principle of discourse of Habermas's moral philosophy.

4

Hegelian criticisms of Habermas's Political Theory

Hegel's criticisms of Kant's moral philosophy are elaborated from the vantage point of social and political institutions which in the *Philosophy of Right* are defined as the components of ethical life [*Sittlichkeit*]. In this respect, Hegel asserts in §142 that *Sittlichkeit* 'is accordingly the *concept of freedom which has become the existing* [*vorhandenen*] world and the nature of self-consciousness'. I read this passage in light of the analytical interpretation that Robert Pippin (1989, 1991) and Terry Pinkard (1994, 1999) offer of Hegel as a *Post-Kantian* thinker, according to which, Hegel is not just a philosopher after Kant, but rather, he is a *Kantian philosopher after Kant.* (See the introduction of Chapter Two) Hegel's critique does not intend to rule out the Kantian concept of moral autonomy. Rather, the issue for Hegel is that pure reason cannot define the autonomy of the will independently of institutions and practices. Henceforth, freedom is always embedded in a social and political context. That is why Hegel states that *Sittlichkeit* is the concept of freedom which has become the existing world and the nature of self-consciousness. (§142) Due to this emphasis on a situated concept of autonomy within social and political institutions, it is not difficult to present these criticisms as objections to political theories constructed in the Kantian tradition in the sense that they may lack this institutional dimension.

In Habermas's theory of constitutional democracy both Kantian and Hegelian components are given a place. In this respect, I claim that Habermas's version of moral validity and legal and political legitimacy takes into account the analytical reading of Hegel's critique of Kant. Habermas has built into the theory the view that the Kantian conception of autonomy must be embedded in, and supported by, social institutions and cultural traditions. The

incorporation of Hegelian elements in Habermas's notion of legal and political legitimacy results in the fortification of the theory.

As I have shown, in *Between Facts and Norms* at the centre of legitimacy stands the principle of democracy which derives from the interpenetration [*Verschränkung*] of the principle of discourse and the legal form. (BFN, 121) In crude terms, the first can be thought of as the expression the Habermasian concept of normative justification. (MCCA, 66, 161; BFN, 107) The second is the result of social and historical processes of evolution. (BFN, 63, 129) In this respect, the Kantian element (D) is embedded in the legal medium – the legal form.

Nonetheless, Habermas asserts that legitimacy is not grounded on morality, because the principle of discourse is below the threshold of this domain. (BFN, 107-109, 121; BNR, 84, 89) Additionally, as already stated, the legal form is a social and historical input. (BFN, 63, 83, 111-113; Finlayson, 2011: 7-11; Hedrick, 2010: 105-106) Moreover, Habermas reinforces the Hegelian side of its theory by incorporating further historical and social components. For example, he claims that the constitution is a 'learning process' (BFN, 421; 2001: 768); democracy demands an active citizenry that embraces a liberal culture (BFN, 59, 132-133); and post-conventional morality and legal and political legitimacy require the support of a life-world context that meets them halfway. (MCCA, 207; BFN, 302, 358, 461, 471, 487; IO, 252) Certainly, all these components strengthen the claim that Habermas is a Hegelian (Gledhill, 2011: 182) and that his concept of legitimacy departs from the Kantian tradition. If that interpretation is correct, then the question whether Hegel's critique of Kant applies to Habermas's concept of legitimacy is answered from the beginning because it seems that he has already left the Kantian family.

However, as I have shown in Chapter Three, an analysis of other parts of BFN among other texts gives rise to a different picture and it comes to seem that the weight of legal and political legitimacy in fact rests not on a rejection of Kantianism but on a modified notion of Kant's concept of pure practical reason (Chapter One). Among other things, this kernel implies the post-conventional notion of normative justification – contained in the principle of discourse – (MCCA, 66; BFN, 107; BNR, 80); the Kantian idea of self-legislation (BFN, 39, 120, 126); and the mutual recognition[1] that citizens grant to each other as free and equal. (MW, 113)

---

[1] Not an idea exclusively owned by Kant, of course – it may be found in Fichte (2000) and in Hegel. (1977a, 1991)

Thus, on this view Habermas thinks legitimacy is based on a Kantian normative ground. Now although that requires the support of social institutions, practices and political cultures, and to this extent Habermas recognizes Hegelian components in his theory, nevertheless, I maintain that Habermas's conception of legitimacy rests upon the Kantian concept of practical reason and autonomy. Consequently, legal and political legitimacy has a normative core that it is embedded in historical time and in social space. In other words, and paraphrasing Hegel, by means of ethical life, the Kantian concept of autonomy becomes *objective spirit*.

In the 1995 debate between Rawls and Habermas the Kantian component of both philosophers emerges in sharp relief. Therefore, this would be the appropriate place to discuss criticisms of Habermas´s theory of legitimacy inspired by Hegel's critique of Kant. In this Chapter, I discuss within the dispute the thesis of the co-originality of private and public autonomy (RPR, 40-45; RH, 63-81); and the relationship between principles of justice and the fact of pluralism. (RPR, 34-40; RH, 55-63; MW, 92-113) These two issues will help me set the stage for the Hegelian criticisms that I examine in this Chapter.

Concerning the co-originality thesis, I develop the argument I made in Chapter Three examining it now in light of the Habermas-Rawls debate. I conclude that the 'system of rights' is prior to the actual exercise of *popular sovereignty*. The notion of a 'two-stage' methodology supports this interpretation. In the first stage, a 'system of rights' composed of 'unsaturated placeholders' (BFN, 125) that protect private and public autonomy, is framed. In the second stage, the system is fleshed out through the 'exercise, the actual carrying out' (Habermas, 2001: 778) of *popular sovereignty*.

I understand this view in the sense that the 'system of rights' needs the support of a normative content which refers to the Kantian concept of autonomy (private autonomy and public autonomy) and that has priority over the exercise of collective self-determination. (Forst develops a similar interpretation of the co-originality, 2011: 164-180) This priority is expressed in Habermas when he asserts that 'the freedom of the moral person splits into the public autonomy of co-legislators and the private autonomy of addressees of the law'. (MW, 113) This means that both forms of autonomy, private and public, derive from the normative concept of justification. And as I have shown, the latter concept is a re-working of Kant's notion of practical reason.

This return of Kant in Habermas's co-originality thesis implies the return of Hegelian criticisms that target the priority of an abstract 'system of rights' over the democratic process. In this Chapter I examine two criticisms. The first arises in light of the Habermas-Rawls debate in which both authors argue that in the theory of the other there is a priority of the 'system of rights' over democratic self-determination (public autonomy). (RPR, 42; RH, 76) This charge can be seen as a Hegelian criticism because it challenges the priority of a normative core over institutions and practices. The second criticism is developed by Frank Michelman, (1998) and to a certain extent also by Richard Bernstein (1998). It asserts that substantive components should be included in the first stage. For example Michelman asserts that 'Constitutional law is institutional stuff from the word go'. (1998: 320) In other words, Michelman's claim is that from the beginning the essential norms of the 'system of rights' need to be shaped in substantive terms.

I believe, however that this charge can be successfully answered by Habermas, given a proper understanding of his Kantian Republicanism. (MW, 113; BNR, 102) Concerning the first criticism – the priority of the 'system of rights' over democratic self-determination – Habermas's notion of practical reason incorporates an element of mutual recognition: 'Nobody can be free at the expense of anybody else's freedom'. (MW, 113) Therefore, it must be that the idea of private autonomy is co-original with the idea of public autonomy. Thus, the inclusion of a Kantian notion of practical reason at the beginning is not detrimental to democracy. Regarding the second criticism, since 'pragmatic, ethical-political and moral reasons' are all included in discourses of legitimation (BFN, 108), then, it is not true – as Michelman asserts (1998: 321) – to say that constitutional democracy is immune to the inclusion of substantive contents. That said, even though it is true that the moral substance enters into the picture at the second stage of justification, when citizens themselves undertake to fill out and justify particular rights, still there is a priority of morality here.

Nonetheless, there is another issue that arises from a particular interpretation of Habermas's political theory, namely, the priority of morality for the justification of legitimate laws. This reading is at hand in Baynes who claims that in Habermas, legitimacy depends on a moral core. (Baynes, 2016: 170, 179) If this account of Habermas is correct then the question is how do particular agents endorse the moral core of legitimate laws? This is because, in this version, legal and political norms not only have to be coherent with moral norms but all these different types of norms need to be justified in a similar way: the

validity requirement of norms implies that they need to be justified in the basis of impartial, agent-neutral reasons. Due to this particular understanding of the centrality of morality in legitimacy and to the stringent condition of the validity requirement, then Hegel's criticism of the will as a tester of maxims not only challenges Habermas's moral theory but his theory of legitimacy as well.

To a certain extent, Habermas's second criticism of John Rawls deals with this issue. For that reason, it will help us to set the stage for this Hegelian criticism. Habermas charges Rawls with neglecting to give a clear account of the validity claim of his political theory. (RPR, 34) For Habermas, the condition of neutrality of the concept of legitimacy determines the epistemic status of the political conception, and sets at the centre of the theory a Kantian notion of practical reason.[2]

Although, in order to ensure that the agents would endorse the political conception, Rawls avoids appealing to comprehensive notions (of practical reason), and legitimacy is instead made to rely on the *overlapping consensus* of the comprehensive doctrines that the citizens endorse.[3] In this way, Rawls's alternative seems to cohere better with the Hegelian question concerning the motivation of the participants because 'moral norms may not be imposed in an abstract manner on the life-histories of individual persons'. (MW, 112) In contrast, Habermas's strong claim concerning the normative kernel of legal and political legitimacy leaves enough room for Hegel's criticisms of the will as a tester of maxims.

Moreover, the Kantian concept of practical reason in Habermas is related to the reconstruction of the idea of normative justification and sets strict limits to what can count as rational. As I have shown, the procedure that derives from this Kantianism permits the justification of norms on the basis of impersonal or agent-neutral reasons, and not at all on the basis of agent-relative reasons. (IO, 7, 43; MW, 94) Therefore, it poses once again the question whether and to what extent the substance that Habermas includes in his notion of normativity can appeal to the particular agents who might be oriented by their agent-relative reasons.

---

[2] At this point and for the sake of my argument, I am not going to examine the difficult issue whether Habermas's charge is misdirected or not.

[3] The notion of *overlapping consensus* is mainly developed in *Political Liberalism*. (133-167, see also Chapter Three) As I have shown (See Chapter One and Three) Rawls's notion of 'comprehensive doctrine' has three different meanings: a) world views, religious and secular; b) actually existing moralities, or conceptions of the good; c) philosophical theories of one kind of another. (PL 2005: xviii, 12, 144-145)

Hegel's criticism of the will as a tester of maxims now becomes relevant to the question of legitimacy. According to this criticism, Kant was not able to combine the particular will of the agent with the universal will of pure practical reason. Therefore, he failed to show that the agents would be motivated by the abstract principles of morality. (See Hegel, 1991: § 135, 153) Relying also on Habermas's Kantian Republicanism, I offer a solution to the problem. The idea of mutual recognition points us to the way in which the agent-relative reasons can coincide with agent-neutral reasons. This is because the agent-relative reason of the citizen for being the author of her political and legal order coincides with the agent-neutral reason that this autonomy should be also enjoyed by everybody else. Nonetheless, this dialectic supposes that Habermas needs to relinquish – at least in this case – the strict distinction that he draws between agreement on the basis of agent-neutral and agent-relative reasons and the stipulation that moral agreement must rest only on the former – agent-neutral reasons. (Finlayson, 1999)

Nevertheless, once again, the fact of pluralism clouds this alternative and cast some doubt on this Kantian solution. In this historical context, it is possible that this notion of mutual recognition and practical reason might still not be appealing for the agents. At the end of the Chapter, I examine a further Kantian alternative to solve this problem. There I examine the particular content that the Habermasian notion of autonomy demands, namely, the principle of discourse which refers to the thin and weak '*concept of normative justification*'. (IO, 45) I conclude that the nature of the principle that legitimacy demands at its core removes most if not all of the difficulties that stood in the way of its endorsement by the citizens of a modern political order.

In order to develop these issues, I divide the Chapter in three sections. In the first, I show that Habermas incorporates Hegelian elements in his notion of legal and political legitimacy. Nevertheless, I maintain that the normative Kantian part has priority over the aforementioned Hegelian components. I disentangle the dialectic between Kant and Hegel in Habermas's political theory examining two Hegelian criticisms inspired by Hegel's critique of Kant's moral philosophy (I).

In the second, I examine the co-originality between private and public autonomy from the point of view of the Habermas-Rawls debate. I show that the Kantianism prevails and that there are good reasons why Habermas still needs to defend this version of legitimacy from criticisms inspired by Hegel's critique of Kant (II).

In the third, I discuss whether the normative core that constitutional democracy demands can find the support from the particular citizens. As far as deontological validity demands agent-neutral reasons it seems difficult to take for granted the allegiance of the citizens to this kernel. Nevertheless, I discuss a solution which implies that Habermas needs to slacken the distinction that he draws between agent-relative and agent-neutral reasons. (III).

I. Hegelian and Kantian components in Habermas's political philosophy

In what follows, I show that Habermas incorporates Hegelian components in his political theory. That said, however, I claim also that a Kantian element has priority. This dialectic between Kant and Hegel in Habermas's theory unfolds in light of two Hegelian criticisms of Habermas's theory. (Honneth, 2014: 1-11; Bernstein, 1998: 287-305; Wellmer, 1998: 39-62) I respond to these criticisms from the point of view of Habermas's Kantian Republicanism which seems to me the best candidate to answer the question of legitimacy in a context of social and ideological pluralism.

To develop these criticisms I divide this part of the Chapter in two sections. In the first I develop the criticism which states that Kantian political philosophies are misleading because they are not performed as a fully immanent reconstruction of principles of justice. In this regard, Honneth asserts 'we should follow Hegel in abstaining from presenting free-standing, constructive justification of norms of justice prior to immanent analysis'. (Honneth, 2014: 5) In the face of this criticism it is important to argue that Habermas develops immanent components in his theory of constitutional democracy, but not in the fully immanent fashion that Honneth proposes. Honneth problematizes the '*freestanding*' character of political theories built on the Kantian tradition.[4] Nevertheless, there are good reasons that justify putting at the centre a Kantian core which goes beyond historical practices. In my interpretation, a proper theory of legal and political legitimacy needs to recognize a kernel – which is the idea of public reason and the recognition of citizens as free and equal – that it is not justified by means of an immanent analysis (1).

---

[4] Honneth understands the concept *freestanding* in a different sense than Rawls. For the former, *freestanding* refers to a political theory that is dissociated from substantive institutions and practices. (See Honneth, 2014: 5) For the latter, *freestanding* means that *political liberalism* is independent from comprehensive doctrines i.e., religious or metaphysical theories. For Rawls, Political Liberalism is circumscribed within the realm of the political and it cannot go beyond that due to the fact of social and ideological pluralism. (PL) It cannot rely on particular conceptions of the good, religions, or philosophical theories. This point is one of the focuses of Habermas's criticism of Rawls. (RPR, 34-40) In order to avoid misunderstanding, in what follows I will use other concepts to refer to political theories that are dissociated from substantive institutions and practices than *freestanding*.

In the second section, I examine another criticism. According to this objection, Kantian political theories do not put at the centre a democratic *ethos*, which is essential for democracy. (See Bernstein, 1998: 289) Instead, the ethical commitments that the citizens should embrace towards democracy and liberalism – or what Habermas has defined as a culture that *meets* deliberative politics and law *halfway* (BFN, 302, 358, 461) – have an instrumental role. In this respect, they contribute to the stability of constitutional democracy, but not to its justification. The weight of legal and political legitimacy rests in the procedures of communicative reason which are independent of historical practices, institutions, substantive political cultures and ethical conceptions of the good. From the point of view of a political version of Kantianism there are good reasons to support this. On my reading, it is not possible to subordinate political legitimacy to ethical commitments and particular conceptions of the good. On the contrary, legitimacy is based on principles that have a universalistic tinge, namely, the idea of public reason and the Kantian principles of freedom and equality (2).

1. Immanent components of Habermas's political theory

In *Freedom's Right*, Honneth proposes a theory of justice as an immanent analysis of institutions which consists in a normative reconstruction of the structural features already present in society. (Honneth, 2014: 3) Thus, he claims that the normative contents or values that frame a theory of justice are the result of the analysis of historical institutions and practices. This is what defines a fully immanent theory that is opposed to what he characterizes as 'conventional Kantian theories of justice'. It is important to notice that Honneth does not equate Habermas's and Rawls's theories with standard versions of Kantianism, in which political legitimacy would be completely uncoupled from the institutional framework of modern societies.[5] Rather:

---

[5] It has been asserted that this is not a proper reading of Kant's moral philosophy. The Kantian defence is that morality does not create moral norms but rather that they derive from social reality. Therefore, the categorical imperative and the universal principle of right are not purely abstract and empty categories which are uncoupled from social institutions. In Kant's moral philosophy and in Habermas's moral theory as well, the principles of morality regulate conflicts of action that 'grow out of everyday life' (MCCA, 204). Nevertheless, it is necessary to draw a more subtle distinction here because Honneth is not contesting that the contents of morality in Kant derive from social life. Rather, his argument is that the principles of morality in Kant are grounded in a fully transcendental (as opposed to immanent) fashion. As he argues, a Kantian or a Lockean theory of justice 'stipulates that the normative principles according to which we judge the moral legitimacy of social orders may not stem from within existing institutional structures, but must stand alone outside of this institutional framework'. (Honneth, 2014: 1-2) Therefore, the problem for Kant remains. I cannot examine this particular issue here because the focus of the present work is not Kant himself but Habermas's Kantianism. Freyenhagen offers a detailed account of Hegel's criticisms of Kant's moral philosophy and Kantian replies. (2011)

Both Rawls's theory of justice and Habermas's theory of law provide good examples of an approach that has its point of departure in the historical congruence between independently derived principles of justice and the normative ideals of modern societies. (Honneth, 2014: 5)

Thus, Honneth acknowledges that Habermas incorporates historical arguments in his reconstruction of political legitimacy. Therefore, normative principles of law and politics coincide with the institutions and practices of modern societies. In this respect, Honneth's criticism begins from a fair reading of Habermas and Rawls. However, the core of his position is that a theory of justice should be the outcome of the fully immanent analysis of society and its substantive institutions and practices. In Habermas it is possible to find immanent components at the centre of his theory – for example, the constitution as a self-correcting learning process. (BFN, 421; 2001: 768) The presence of these elements rules out from the outset the misleading reading that he might be developing a purely formal, abstract and empty theory of legitimacy.

In what follows I develop two central features of Habermas's political theory, which show that he incorporates an immanent methodology into his philosophy (1.1). Nevertheless, I argue, he does not develop a fully immanent reconstruction – as Honneth advocates – because even the historical components of a theory of legitimacy have a Kantian core, namely, the recognition that citizens grant to each other as free and equal persons which is contained in the post-conventional notion of normative justification – the principle of discourse (D) – (1.2).

1.1. Systematically, Habermas has developed historical arguments in the different programs he has been working on. This methodological feature of his *oeuvre* arises from his early book *The Structural Transformation of the Public Sphere* and is also developed in the *Theory of Communicative Action*. In his theory of legal and political legitimacy that is also the case. In *Between Facts and Norms* his justification of the system of rights is immanent to law. As a matter of fact, it can be argued that it is too immanent, which from the perspective of a *stronger* Kantianism is problematic. (Forst, 2011: 173; Flynn, 2011: 253-254) Moreover, Habermas asserts that the internal connection between private and public autonomy 'can develop only in the dimension of time – as a self-correcting historical process' and that consequently the constitution is a 'living project' (BFN, 129) 'that makes the founding act into an ongoing process of constitution-making across generations'. (Habermas, 2001: 768)

Thus, this second argument emphasizes the historical dimension of constitutional democracy, in which founders and present generations are part of a common project of democratic self-determination. Consequently, Habermas's theory of legitimacy acknowledges in its construction the historical dimension of modern societies. On the one hand, the form of law is an empirical and historical component; on the other hand, the constitutional practice is a learning process, which relates founders and present generations.

Concerning the first element, the legal form has functional and empirical features and is introduced from the observers' perspective, operating in the stabilization of behavioural expectations to guarantee society´s integration. Therefore, it is a component of social reality. It is part of the network of practices and institutions that produce the substantial shape of constitutional democracy. The relevant point is that due to the features of the principle of discourse and the legal form, Habermas is able to connect his formal discourse theory with a functional feature of modern societies. (BFN, 83; Finlayson, 2011: 7; Hedrick, 2010: 106) Hence, he integrates a normative principle that can operate as a way to assess the rationality of modern law and an empirical input. (Hedrick, 2010: 105) This functional character of law is expressed in the fact that in modern conditions it has to be understood as positive law.

According to BFN, citizens can and often do obey laws due to an insight into their intrinsic normative rationality. Nevertheless, it is also necessary to consider the functional requirements of positive law, because when this insight fails the individuals know that there remains the threat of prosecution and punishment.[6] As Pedersen argues, 'positive law leaves the question of motivation to the participants'. (Pedersen, 2011: 417, in Hedrick, 2010) However, even in the case that law is obeyed due to the threat of sanctions, there remains a normative core that might not be recognized and accepted by the agents but which represents the rationality internal to law.[7] Now, the point is that it does not matter in which specific grounds an agent obey the law, provide she does. This is different in morality because there the worth of the action, and its character as a moral action depends on the reasons or the insights that ground an action. Habermas's account of moral action is Kantian and deontological in this respect (and not consequentialist).

---

[6] The norms can, 'of course, be obeyed on the basis of insight […] but, they also reckon with the legal subjects who act from self-interest and free choice and whose conduct must be bindingly regulated without reference to moral motives'. (Forst, 2011: 166) This is what Habermas has defined as the 'Janus-faced character of law'. (BFN, 448)

[7] In this respect, I recall the discussion broached in Chapter Three, section 1 of part I. There, I discuss the 'internal' tension between facticity and validity.

For Habermas, this enforceability – its facticity – of law is always combined with its validity. In a context of pluralism, and from the vantage point of radical democracy, law is legitimate when it respects the self-legislation of the political community which refers to the Kantian idea of citizens as *addressees* and *authors* of their constitution. (BFN, 104)[8] Concerning the form of law, Forst claims that this concept is not justified in Habermas's theory and it is described as the result of a historical development (Forst, 2011: 167) reflected in two hundred years of constitutional law. (BFN, 129; Hedrick, 2010: 111)

The second historical component refers to the relationship between private and public autonomy as a 'learning process'. (BFN, 421; Habermas, 2001: 768) Habermas develops this thesis along two lines. The first focuses on the past where later generations continue the project of interpreting the normative content of the constitution embarked upon by earlier generations. Thus, the later generations begin with the same standards, as did the founders. In this respect, for Habermas all citizens must recognize a core in constitutionalism which is the same throughout history. This idea is at the centre of the communitarian and republican traditions. Nevertheless, they do not include the normative principle of discourse, as Habermas does, which relies on a Kantian account of practical reason: namely, the public use of reason (RPR, 38) in conditions of post-metaphysical thinking (PT: 3-9), which is shaped by the weak *'concept of normative justification'*. (IO, 45)

Concerning, the second line, it implies a change in focus from the past to the future, from the constitutional tradition to the performativity of constitution-making as a promise of future reconciliation. Here, citizens must still see themselves as the heirs of a founding generation, 'carrying on with the common project'. (Thomassen, 2010: 54) However, they see themselves as heirs and at the same time the constitution making process ultimately refers to the promise of an unrestricted exchange of reasons that presupposes the possibility of agreement but also provides space for disagreement. In this respect, again the normative idea of the public use of reason frames – albeit in an abstract and unfinished fashion – the practice of constitution making.

1.2. Thus, it would not be fair to argue that Habermas does not develop an immanent analysis in his theory of legal and political legitimacy. In this respect, at least two elements are the result of that methodological strategy. On the one hand, the form of law, on the other, the constitutional practice as a learning process. Nevertheless, this reconstruction is not performed in the fully immanent fashion that Honneth expects and there are good

---

[8] See other essential features of the validity of a legal norm in. (Finlayson, 2011: 10)

arguments which explain this theoretical decision. In Habermas, there is still a Kantian element that is independent and beyond historical practices and institutions, namely, the post-conventional notion of normative justification.[9] (D) is a universal – it is transculturally valid. However, Habermas thinks it comes to be historically as a result of the process of modernisation and rationalisation. As a weak transcendental principle it is not logically or metaphysically necessary, but necessary for our (communicative and discursive) form of life. That said, principle (D) is not fully immanent in that its idealizing presuppositions transcend actual conditions. Moreover, this component pertains to the public use of reason which is internally connected with citizens who grant each other equal rights to freedom.

In the argument concerning the practice of constitution making as a learning process developed in history, the Kantian principle of Discourse is shared by the founders and the heirs of the constitutional practice whenever they want to legitimately regulate their political life by means of law. Additionally, this normative core frames the legitimacy of constitutional democracy when it is embodied in the medium of law.

This claim is sustained in the fact that the principle of democracy includes the post-conventional notion of normative justification at its core. (BFN, 107) Thus, as far as this element is included at the centre of legal and political legitimacy, it is not difficult to understand why moral reasons have priority in the justification of legal norms over ethical-political and pragmatical considerations. (BFN, 103, 108, 113; IO, 42-43; Flynn, 2003, 434) Moreover, legitimate laws must 'harmonize with the moral principles of justice and solidarity' (BFN, 99, 155) which Finlayson has called the *Moral Permissibility Constraint* MPC. (Finlayson, 2011: 12, 2016a: 12, 14)

Consequently, even the historical features of the theory are framed in Kantian terms. In other words, in *Between Facts and Norms* and elsewhere this component prevails in the grounding of political and legal legitimacy. Thus, Honneth is correct to argue that Habermas does not build his theory as a fully immanent reconstruction of principles of justice. Notwithstanding, the former is not correct when he advocates that the theory of legitimacy has to be developed in this fashion. Habermas's political theory includes the idea of normative justification which responds to what might be called Kantian Republicanism. (MW, 110-113; BNR, 102)

---

[9] It is possible to argue that the principle of discourse is an immanent reconstruction of the practice of justification in a post-conventional context. Nevertheless, in Chapter One and Three I showed that there is still a transcendental component in (D).

This kernel is grounded in a notion of practical reason, and not in the historical features of modern societies. Moreover, this Kantian substance is apprehended in the principle of discourse which is an intersubjective conception of practical reason. Naturally, Habermas's conception of a weak *'concept of normative justification'* (IO, 45) – that alludes to the capacity to redeem validity claims in discourse – has an intersubjective character. In Chapter One and Three I examined the relationships between this concept and Kant' notion of pure practical reason and I concluded that at bottom the post-conventional notion of normative justification – which is at the centre of Habermas's *oeuvre* – belongs to the Kantian family.

Nevertheless, Habermas's dialogical idea of autonomy presupposes a break with the monological conception of autonomy that structures Kant's account of pure practical reason. In other words, Kant's reconstruction of the moral point of view, the categorical imperative, defines a procedure that every agent can perform on their own to check the universalizability of moral norms. By contrast, Habermas has an intersubjectivistic understanding of practical reason and autonomy. (See Chapter One) However, there are many elements which both philosophers share, and this is one of the reasons that allow Habermas to consider his position as a form of Kantian pragmatism (T&J, 83-130) and also of Kantian Republicanism. (MW, 110-113; BNR, 102) In the case of the latter, the Kantianism in Habermas arises in the medium of law itself. In the interpretation that I have been developing, this empirical input – the legal form – has to be framed to protect and enact the basic rights to exercise private and public autonomy.[10] Therefore, the normative component frames the form of law and then the historical argument is dependent on the former.

Habermas argues correctly that in a *weak* sense his theory of political legitimacy includes Hegelian elements. (1998: 384) He would be a Hegelian in the *strong* sense if he were developing a fully immanent reconstruction. Now, in my interpretation this *weak* Hegelianism runs alongside a *weak* Kantianism, and rightly so. The weight of the justification of political legitimacy is normative, in a *weakly* Kantian sense. As I have shown, even the immanent elements of the theory incorporate the Kantian notion of moral reason. The form of law itself, which in *Between Facts and Norms* is reconstructed as an historical feature includes normative Kantian components. The American legal philosopher Lon Fuller proposes that the legal form includes several normative elements which he has called the 'morality of law': being publicly promulgated, clearly formulated and not applicable

---

[10] See Chapter Three, part I, section 2.

retroactively. (1969: 33-38) One might include others: the autonomy of the judiciary, the system of rights and its relationship with subjective rights that guarantee equality and freedom. Of course, these features have a normative character, but still, it is possible to understand them as procedures and contents which guarantee legal and political legitimacy.[11]

The 'system of rights' certainly has a normative flavour. These rights have as their principal function the protection of equality and freedom among legal subjects. Furthermore, these rights should be coherent with fundamental rights, which, Habermas claims, have both a legal and fully moral nature. No one can deny that Human Rights are moral norms because they square with everyone's interest. Therefore, they are universalizable norms, and have a deontological character. In this respect, Human Rights are internally connected with Human dignity, which acquired its current canonical expression in Kant. (Habermas, 2010: 465) Nevertheless, they are also legal entities which are embodied in and supported by institutions and practices. This is what pertains to its *Janus-Faced* nature. (BNF, 454; Flynn, 2003: 432) This nature of basic rights shows how Habermas's notion of legal and political legitimacy squares with the analytical interpretation of Hegel's critique of Kant: moral content is embodied and supported by a modern *Sittlichkeit*.

Finally, the very independence of the judiciary can be interpreted as another argument for the primacy of a normative core in the form of law. From the point of view of the tradition of juridical positivism, this independence is related to the necessary expertise of the jurists. However, this can be understood also as the autonomy that the legal system needs in order to protect the freedom of legal subjects. For example, it expresses the protection of the integrity of the person in front of others and in front of institutions that have more political power; and also, the liberal principle of the defence of minorities against the collective will of majorities.[12]

Concerning the constitution practice as an historical process, the founders and heirs share Kantian notions of freedom and autonomy. This is a universalistic and still untapped

---

[11] Nevertheless, this morality of law does not refer to the morality of Discourse Ethics. It is a thin and narrow notion of morality, confined to certain aspects of the legal system, which does not reaches the post-conventional level of morality of Discourse Ethics. It does not regulate every action, not bind every moral agent.

[12] However, to be democratic the collective will cannot contradict the system of rights which protects public and private autonomy; otherwise it is simply not democratic. It can be called something else, but not democracy. Historically, whenever a political system stops respecting private autonomy in the name of the general will, what is taking place is something different to what could be conceived as a proper democratic system.

normative core of normative insights. It is still untapped, because the constitution making is a learning process that is not closed from the outset. At the beginning of the constitutional history of many countries, subordinated groups were not granted rights. For example, the political rights of women did not exist until the second half of the twentieth century. The normative core of democracy clashed constantly with this situation. Through historical struggles women finally could also be part of the democratic process. My question here is if politics were only an immanent process, then how is it possible to get from an order in which one group does not have political rights to another where it does? Constitutional democracy contains a universalistic core which when developed and interpreted through public discourses, gives a concrete shape to a legitimate political system, in which the *addressees* of the norms are also their *authors*. (BFN, 104, 120)

To sum up, in this section I have shown that Habermas includes immanent components in his political theory, but does not develop all the contents of a political conception of justice, and the notion of political legitimacy in a fully immanent fashion, because both these elements are subordinated to the Kantian core, namely, the principle of discourse. This principle is embodied in the medium of law by means of the interpenetration that I have examined in detail in Chapter Three. This means, that whatever Habermas may claim, a Kantian concept of autonomy frames Habermas's notion of constitutional democracy. Furthermore, the constitution as a self-correcting learning process also retains a Kantian notion of normative justification. Namely, founders and heirs understand themselves as bounded by the idea of freedom and equality. Thus, as far as this Kantian component is included at the centre of legitimacy, it is understandable that moral reasons have priority in the justification of the 'system of rights' over ethical-political and pragmatical considerations. (BFN, 103, Flynn, 2003, 434) In this respect, in issues concerning legitimacy Habermas insist in the priority of the right, as he does in the program of Discourse Ethics. In what follows, I examine a further Hegelian component of Habermas's notion of legal and political legitimacy, namely, the idea of a democratic political culture.

2. Democratic *Sittlichkeit* and political culture

A further Hegelian criticism of Habermas's political theory refers to the priority for democracy of the existence of citizens who embrace democratic political values. In this respect, Richard J. Bernstein (1998: 287-305) and Albrecht Wellmer (1998: 39-62) argue that a democratic regime needs an active citizenry with the will to deliberate. This active citizenry 'accept reasonable and rational arguments', and are 'oriented toward the common

good rather than their own particular interests'. (Hedrick, 2010: 129) This element is constitutive for democracy and according to these critics it has only an instrumental role in Habermas. As Hedrick reports, for them in Habermas's theory, 'the dependence of the legal order on democratic *Sittlichkeit* is asymmetrical: a community's ethos is called on to support the ongoing legitimation of the legal order, but only after its essential structure has been determined'. (2010: 130)

In other words, in Habermas this democratic political culture is subordinated to the Kantian principles which frame the practice of self-legislation, namely, the principle of discourse (D) and the principle of democracy. The weight of the argument in his theory rests on the normative core that structures the system of rights and democracy. In Bernstein's terms the problem is that the more Habermas insists upon the 'purity' of his theory 'the more formal and empty it becomes'. (Bernstein, 1998: 289) On Bernstein's account a discursive theory of democracy has to be built as a hermeneutical circle that presupposes a democratic ethos. Such a democratic substantive ethos does not determine specific norms, values, and decisions. While it is still formal it is substantive at the same time. According to Bernstein, Habermas should be more faithful and recognize that 'all social and political theory involves ethical presuppositions and commitments'. (Bernstein, 1998: 290) However, this raises the problem of those ethical commitments (ethical as opposed to moral).

This criticism is based on an insight that is present in Habermas's methodological approach. In *Theory of Communicative Action* he combines both a reconstructive component and the point of view of the participants. In the case of political legitimacy, in the former, the fundamental features of a legal order are based on its capacity to guarantee the rational and autonomous governance of society. From the perspective of the participants, political issues are substantive problems. In relation to the charge that I am discussing in this section, Habermas's political theory focuses on the side of the reconstruction of principles and it does not seem to 'be capturing the reasons that participants in practices of self-government would give for viewing those practices as justified'. (Hedrick, 2010: 125)

The earlier theory of universal pragmatics does not suffer from this problem since the reconstructive perspective seeks to make explicit the knowledge that participants presuppose in their actual communicative practice. In other words, the reconstructive and the participants' approach coincide. That is not the case in Habermas's political theory. According to Bernhardt Peters (1994), modern institutions such as law and politics are

based not on a substantive insight, but in rational procedures. The problem is that from the participants' perspective, there 'is still the convincing force of substantive arguments that (ideally or rationally) leads them to accept certain propositions as true, certain norms or institutional orders as legitimate'. (Peters, 1994) Thus, in Habermas there is a difference between the knowledge that participants have and the reconstructive knowledge generated by the theory.

Rawls's concept of *overlapping consensus* (PL, 133) gives a central role to the perspective of the participants in a political theory. Therefore, his theory seems to involve two elements which are difficult to reconcile. On the one hand, he is developing a theory of justice in the Kantian tradition. On the other hand, he is responding convincingly to a Hegelian charge: namely, that the Rawlsian citizens participate in public justifications oriented towards reaching a shared conception of justice embedded in their diverse comprehensive doctrines. Thus, *Political Liberalism* formalizes certain components which are already present and accepted in the life-world of the citizens. Thereby, in this theory at least some of the substantive views that are reconstructed from the participant's perspective have a place in the legitimacy of politics. In this respect, Hedrick asserts that Habermas's argument should make 'some concessions to the Rawlsian position, by conceiving of reasonableness as a substantive commitment to a rational and communicative form of life shaped by a process of legitimate law-making'. (Hedrick, 2010: 127)

However, this commitment to a particular form of life is not feasible in light of the fact of pluralism. (BNR, 102; Finlayson, 2011: 19) In this context democratic practice in modern societies cannot be grounded in a pre-political ethical core. Political legitimacy can be constrained only 'by the preconditions for rational discourse through law' and it does not depend on a particular form of ethical life. (Hedrick, 2010: 127)

As I have shown, Habermas recognizes that political legitimacy needs to *be met by* a coherent political culture *halfway* (BFN, 302, 461), but this does not mean that it depends entirely on this form of life. It is important to notice that the form of life is not a thick conception of the good, but to a form of life where moral agents and citizens have the disposition to solidarize with other citizens and to act on insight into moral and political norms. As agents at stage 6 they are disposed to be bound by the results of moral and democratic discourses. Moreover, the allegiance to this form of life does not mean the end of the public autonomy of the citizens. This is because the concrete shape that democracy, the 'system of rights' and the constitution take is something that has to be constantly

worked out by the citizens. It is important to emphasize that Habermas's claim that a form of life must meet morality (and deliberative democracy) halfway, does not amount to the claim that a substantial ethos is required as the basis of political association. This is because morality and politics are embedded in a form of life, but their validity and legitimacy depend on the principle of Discourse and its specifications in the principle of universalization (U) and the principle of democracy.

In the case of politics, this might sound problematic given that the normative core of a democratic constitutional state looks underdetermined, ethically underdetermined, so to speak. However, for Habermas this is an advantage and is more congruent with the fact of pluralism and the conditions that this fact brings about. Furthermore, it responds to the demands of a critical perspective, because the outcomes of the discourses are left up to the participants to determine. That is the reason why Habermas defines the first three categories of rights that form the system of rights as *unsaturated placeholders* that are 'more like legal principles that guide the framers of constitutions'. (BFN, 126) In principle, public autonomy is not constrained by procedures neither from the beginning nor from the end. Nevertheless, self-legislation cannot contradict – what I want to call following Baynes – a thin core (Baynes, 2016: 107-108) of universal norms. This condition is related to what Finlayson has called the *Moral Permissibility Constraint.* (Finlayson, 2011: 12; 2016a: 12, 14)

Habermas's answer to these criticisms relies on his Kantian Republicanism within the context of a society characterized by the fact of social and ideological pluralism. In modern conditions, ethical appeals to ground a legal order are not possible. Therefore, its justification has to rely on a procedural basis. This defines Habermas´ commitment to a Kantian Republicanism which seeks to:

> Transform the moral level of justification into procedures of political self-legislation and attach justice less to the general principles than to the democratic legitimation of norms and laws. (Forst, 2011: 164-165)

Nevertheless, the question remains whether Habermas is able to illuminate the actual practice of everyday human reasoning, which from a Hegelian perspective is oriented *from* and *toward* substantive norms as grounds with which to evaluate the legitimacy of modern legal and political orders. There are strong reasons for arguing that Habermas's theory can illuminate this: the modern condition is characterized by a plurality of conceptions of the good. Therefore, a procedural theory in the Kantian tradition has more chances of

illuminating the dimension of validity in its universality.

So far, I have discussed a criticism of Habermas's theory of constitutional democracy from a Hegelian point of view. This charge is based on the distinction between the theorist's perspective and the participants' perspective. It has been argued that Rawls in his concept of *overlapping consensus* seems to be more aware than Habermas of the participants' perspective. (Hedrick, 2010: 127) Namely, the substantive principles that are based on the comprehensive doctrines that the participants embrace have a central place in *Political Liberalism.*

To answer this criticism I shall appeal to Habermas's Kantian Republicanism and his notion of post-metaphysical thinking. (PT: 3-9) In this paradigm, as a result of the fact of pluralism, according to Habermas, it is not possible to ground legal and political legitimacy in any particular doctrine of the good life or in any set of specific ethical commitments. If that were the case, then we would not know which one would be the right conception of the good. A proper account of legitimacy needs to acknowledge and deal with the fact of pluralism. Therefore, only a procedural theory – understood as a Kantian Republicanism – answers the question of how of legitimate laws can be produced that can to be obeyed due to their intrinsic rationality.

II. The co-originality and the return of Hegel

In the 1995 debate between Habermas and Rawls, the sketch that emerges is that the theories of both philosophers are 'two-stage' constructions. (RH, 63, 76; Habermas, 2001: 766-782; Forst, 2011: 171) In the first stage an abstract core of basic rights is constructed from the vantage point of the nonparticipant. (BFN, 118) These rights are defined as 'unsaturated placeholders that guide the framers of constitutions'. (BFN, 126) In a second stage, they gain their actual shape through the exercise of public autonomy. (Habermas, 2001: 778) This reconstruction supports the Kantian reading of Habermas's theory. At the first stage the normative component is incorporated beyond democratic institutions and practices. In this way, a post-conventional notion of normative justification bears the weight of legal and political legitimacy.

With this version of the relationship between public and private autonomy in hand the question that I want to examine is whether Habermas's Kantianism is immune to Hegelian criticisms. To develop these issues, I discuss the co-originality between private and public autonomy in the Habermas-Rawls debate. Here, it is possible to conclude that both are part

of the Kantian *family,* because at bottom they share a conception of moral reason and autonomy at the centre of their political theories. (Forst, 2011: 153; Laden, 2011: 135) This component results in the return of the Hegelian criticisms that target the priority of normative procedures and an abstract set of basic of rights prior to the democratic process (1). We will then see whether these Hegelian objections can be answered by Habermas's Kantian Republicanism (2).

1. The Habermas-Rawls debate, the co-originality and the Hegelian objection

According to Habermas, his debate with Rawls is a *familial dispute.* (RPR, 25) Perhaps the fundamental reason why he names it in those terms is that both share a Kantian notion of practical reason and autonomy. (Forst, 2011: 153; Laden, 2011: 135) If they endorse this Kantianism, then their dispute is a good place to examine the particular shape of this component of Habermas's theory. In this context, I focus first on the issue of the co-originality of private and public autonomy. I argue that Habermas and Rawls share a common Kantian conception of legitimacy and we can show this by means of an analysis of their 'two-stage' reconstruction of a 'system of rights' (1.1). Then, I discuss this aspect of Habermas from the point of view of two criticisms that are inspired by Hegel's critique of Kant. The first concerns the thesis of the co-originality. The second concerns the inclusion of substantive components in Habermas's political theory (1.2).

1.1. Habermas's criticism of Rawls, in light of the co-originality thesis, is that the latter gives primacy to liberal rights over democracy. (RPR, 42) In other words, in the well-known dispute between the *moderns* and the *ancients* (following Constant), Rawls shifts the balance on the side of the modern conception of liberties, by privileging individual autonomy. Public autonomy, so Habermas argues, is modelled at the level of the original position. This means that the possible outcomes of the democratic practice are decided in advance. As a result, this autonomy is present just on the first level of the theory and 'does not fully unfold in the heart of the justly constituted society'. (RPR, 42) Public autonomy is part of the original position but later when the veil of ignorance is lifted, Rawls´ citizens confront principles and norms that already have been anticipated within the theory and are beyond their control. Thus:

> They cannot reignite the radical democratic embers of the original position in the civic life of their society, for from their perspective all of the *essential* discourses of

legitimation have already taken place within the theory; and they find the results of the theory already sedimented in the constitution. (RPR, 42)

Consequently, the citizens are confronted with norms and also with political questions that have been already anticipated in the theory. Therefore, Rawls fails to do justice to the co-originality of private and public autonomy. In this respect, Habermas argues that this leads to a democratic deficit in the theory – or the privileging of political philosophy over democratic politics.

In *Political Liberalism: Reply to Habermas,* Rawls rejects this criticism. He asserts that his construction of a four-stage sequence [original position, constitutional convention, legislation and adjudication] is misunderstood by Habermas, because at every step there is space to check the contents. In his own words, 'The four-stage sequence fits, then, with the idea that the liberties of the moderns are subject to the constituent will of the people'. (RH, 71) Thus, according to Rawls in his theory there is a proper place for public autonomy. He argues that even 'the political conception of justice, like any other conception, is always subject to being checked by our reflective considered judgements'. (RH, 66)

Nevertheless, the answer given by Rawls is not convincing because he is referring to the four-stage as a model for the reflexive application of the principles of justice: it is not up to the citizens to decide them; they are given before discourses of political legitimation take place. In *Political Liberalism* this is also the case. There Rawls 'advances the thesis that the collective use of reason in questions of 'constitutional essentials' and 'basic justice' is subject to the restriction that these questions should be answered only on the basis of 'political values' and not with reference to contested doctrines'. (Forst, 2011: 170) Thus, as Forst argues, Rawls's account of political legitimacy recognizes principles of justice that have priority over the self-legislation of a political community. In other words, the citizens are first addressees of rights and only after that their authors.

Furthermore, additional evidence that supports this reading is that Rawls's theory does not seem to go beyond a monological account of reason with political-public intent towards a theory of deliberative democracy. (Forst, 2011: 171) In consequence, Rawls' answer does not successfully deal with Habermas's criticism and the former does not adequately capture the co-originality of private and public autonomy. In *Justice as Fairness* a system of rights oriented to protect individual freedom is prior to the public practice of self-legislation.

That said, although Rawls's reply is ultimately not satisfactory, it nonetheless makes two important contributions. On the one hand, it helps to clarify his position concerning the balance between private and public autonomy.[13] On the other, it advances an interesting interpretation of Habermas's co-originality thesis. According to Rawls the theory of *Justice as Fairness* and Habermas's theory of constitutional democracy are necessarily 'two-stage' reconstructions. (RH, 76) In this respect, Habermas cannot avoid a methodology that assigns to private autonomy and human rights a moral meaning and consequently a normative priority. The first statement that refers to the moral content is something that Habermas does not deny. (BFN, 455-456) Nevertheless, the claim that Habermas is giving priority to individual freedom over collective self-determination is at odds with his official version of the co-originality thesis which implies among other things that private and public autonomy have 'equal-weight'. The idea of a 'two-stage' reconstruction necessitates the specification of particular rights prior to their institutionalization. According to Forst, Rawls argues 'that Habermas also cannot get around a 'two-stage' theory construction that assigns to human rights a moral content and normative priority'. (2011: 171)

Now, as I showed in Chapter Three, this opens an ambiguity in Rawls's critique of Habermas. On the one hand, the 'system of rights' become part of positive law only when it is understood as condition for the institutionalization of political autonomy. This is Larmore's interpretation of Habermas's thesis of the co-originality. (1999: 64-65) In this respect, although Habermas allots an intrinsic value to the 'system of rights', he fails because in his theory basic rights and political rights are supposed to have only a functional place in the institutionalization of collective self-determination.

Nevertheless, the other reading is that by means of the 'two-stage' construction Habermas needs to assign normative priority to the 'system of rights' and this is detrimental to public autonomy. Ingeborg Maus (1995: 825-882) is critical of this reconstruction of the co-originality, whereas Jeffrey Flynn (2003, 2011) and Rainer Forst (2007) agree that the priority of a moral content is the only alternative remaining which would allow Habermas to give a proper justification to the 'system of rights' and to ensure co-originality. (See also Baynes, 2016) This reading is more coherent with Habermas's theory in general and with the development of his notion of legitimacy in writings after *Between Facts and Norms*. (MW, 110-113; Habermas, 2001: 766-782; 2010: 462-480)

---

[13] In this respect, Rawls recognizes that by means of the debate 'I have been forced to think through and re-examine many aspects of my view and now believe I understand it better that I once did. (RH, 91)

In summary, Habermas argues that the principle of discourse is embodied in the medium of law. Then, the 'system of rights' that arises has normative priority over the democratic process. By means of the 'two-stage' reconstruction (2001: 766-782; 2011: 296) Habermas is acknowledging the priority of the 'system of rights', before the concrete practice of self-legislation gets in motion. Hence, he errs more on the liberal side – which gives priority to the 'system of rights' – than on the republican side – that put the emphasis on the actual exercise of public autonomy. Interestingly enough, when Habermas opens the debate and pushes Rawls to recognize himself in the Kantian family, the unexpected outcome is that the debate pushes Habermas to acknowledge that he also belongs more to that family that he admits in BFN.[14]

Despite what Habermas says to the contrary in BFN, in his theory of legitimacy the weight moves towards the 'system of rights' before it gains actual shape through public autonomy. Then, as in Kant, a normative concept has temporal and conceptual priority over the historical practice of self-legislation.[15] As I have shown, this interpretation of Habermas's theory is at odds with the official view set out in BFN - and also in BNR. In that version, the justification of the 'system of rights' is fully immanent to law. (Forst, 2011: 173; Flynn, 2011: 253-254) Nevertheless, following Forst and Flynn I claim that the 'system of rights' needs the support of a normative component. (Forst, 2011; Flynn, 2003; 2011; Baynes, 2016) As I have argued in Chapter One, this content refers to a modified version of Kant's notion of autonomy. Naturally, this re-emergence of a certain Kantianism lays Habermas open once again to the Hegelian critique we have been attempting to rebut. The question is whether and to what extent the Kantian version of Habermas can adequately answer these criticisms.

1.2. Hitherto, I have shown that Habermas's version of the co-originality grants priority to 'the system of rights' over the exercise of public autonomy. Therefore, a Kantian motive returns because a previous 'system of rights' with a universalistic meaning arises before the democratic practice takes place. Now, I discuss two criticisms of Habermas's thesis of co-originality that are prompted by Hegel's critique of Kant. The first concerns the subordination of democratic institutions and practices to a Kantian core of universal rightness. The second is strongly connected with the first, although it does not target the

---

[14] I am indebted to Rainer Forst for this interpretation of the Habermas-Rawls debate.
[15] It is important to note that this is a matter of controversy. According to Baynes 'The difficulty is reflected in the question whether Kant is best understood as a natural rights theorist or a social contract theorist'. (2016: 135) For example, according to Habermas and O'Neill (1975) the universal principle of right is a subsidiary formula of the categorical imperative.

priority of private autonomy *per se*. Rather, it focuses on the issue that the system of rights is in the first stage abstractly reconstructed and it does not incorporate substantive components of ethical life from the beginning.

The first is the charge that Habermas and Rawls formulate against each other. It claims that the priority of the system of rights is detrimental to democracy and the actual practice of self-legislation. (RPR, 40-45; RH, 63-81; Maus, 1992) In the vocabulary of Hegel's critique of Kant, institutions and practices are ancillary to a moral core that constrains them.[16] Thus, the notion of practical reason and autonomy – that in Habermas frames the system of rights – has priority over social space and historical time. Of course, from a Hegelian point of view this is not the right way to proceed. As we have seen, Honneth argues that a theory of justice should abstain from offering principles of justice prior to immanent analysis. (2014: 5)

The second criticism is developed by Frank Michelman in his contribution to the collection entitled *Habermas on Law and Democracy.* The claim here is that the 'system of rights' cannot be constructed in an abstract mode. Michelman asserts that law is always an institutional domain; it is embedded in social space and historical time from the beginning. Thus, Habermas's 'two-stage' reconstruction that begins from abstract rights, or what he calls 'unsaturated placeholders' (BFN, 125), is challenged. In Michelman's view, there is 'nowhere to go' if the starting point is only an abstract definition of rights, beyond any particular context. Instead, it is always necessary to frame these rights taking into account substantive institutions and practices. In this respect:

> For suppose we start, as the discourse-theoretic paradigm would have it, with a liminally abstract principle of right […] Then there is nowhere to go from there – nowhere to go from where we start- except by steps that cannot be taken without reference to institutional-practical questions of concrete definition in context. (Michelman, 1998: 321)

Thus, Habermas's 'two-stage' reconstruction (Habermas, 2001) that begins from abstract rights, or what he calls 'unsaturated placeholders' (BFN, 125), is challenged from a Hegelian point of view.

---

[16] Undoubtedly, this is a modern democratic reading of Hegel. (See Cortella, 2015)

Furthermore, according to Michelman Habermas is wrong when he incorporates Kant's distinction between questions of justification and of application.[17] More precisely, 'the dissociation is analytical, not empirical-practical' (Michelman, 1998: 321) something that Habermas does not seem to acknowledge. In questions of legal and political legitimacy the disjunction is misleading because 'Constitutionalism requires enactment *at the originary* level of what are called *laws,* because constitutionalism *means* a rule of law, a government of laws'. (Michelman, 1998: 321) Hence, the 'two-stages' reconstruction is undermined from a Hegelian point of view. The system of rights should contain from the beginning substantive rights that among other things are shaped not only in terms of their rational justifiability but also of their practical applicability. (1998: 322)

The balance proposed by Habermas between the 'system of rights' and collective self-determination can be targeted by this criticism as well. In light of this charge, it is plausible to argue that the substance should be incorporated from the beginning by means of the exercise of self-legislation or public autonomy.[18] If that were the case, then the 'system of rights' and public autonomy would co-originally produce the legal order, and both principles would have equal weight. In Habermas's theory, the co-originality fails and the constitution cannot rule: it cannot be enacted if it does not incorporate substance and the norms of its application at the first stage. In this Hegelian spirit, Michelman asserts that the 'system of rights' needs to have already a substantive basis in a concrete form of ethical life.

In the next section, I address these two Hegelian criticisms of Habermas's thesis of co-originality. I assert that not only this position can be defended but also that in modern conditions it is the best candidate to respond to legal and political questions in a context of social and ideological pluralism. At the present time, it is not possible to ground the legitimacy of a legal order in its contingent development, in a national history or in a conception of the good life. Rather, a basic system of rights can be generally accepted if is based on a universal core of normative rightness that finds support by means of an accommodating political culture.

---

[17] In *Justification and Application* Habermas addresses this criticism which was formulated by Wellmer in *The persistence of modernity,* from the point of view of Discourse Ethics. In that book Habermas accepts that discourses of justification have to be complemented in a second stage through discourses of application. (JA, 37-39) Nevertheless, there is still a primacy of justification. In this respect 'Deontological ethical conceptions assume in the final analysis only that the moral point of view remains identical; but neither our understanding of this fundamental intuition, nor the interpretations we give morally valid rules in applying them to un-foreseeable cases, remain invariant'. (JA, 39)

[18] This charge resembles one of Hegel's criticisms of Kant. This specific criticism asserts that Kant's empty formalism obscures the consideration of institutions and practices. (See Hegel, 1977: 258; Benhabib, 1986: 309) I examine in detail this issue in Chapter Two, section 1.

2. Kantian Republicanism as a solution to the Hegelian challenge

So far, I have described two Hegelian criticisms of Habermas's thesis of co-originality between private and public autonomy. In what follows, I develop replies to them. The first objection targeted the primacy of private autonomy over the exercise of public autonomy which results from the 'two-stage' reconstruction of Habermas's theory. To this charge I argue that this strategy is cogent as far as private and public autonomy require a normative core of universal rightness (2.1). The second challenged the 'two-stage' methodology because it did not include ethical substance in the definition of legal norms from the beginning. I will assert that Habermas is right when he claims that there is a common core to different legal orders that is fleshed out considering the context in a second stage (2.2).

2.1. Bearing in mind the first criticism, I argue that it is necessary to establish a 'two-stage' reconstruction of a system of rights which gives priority to private autonomy over public autonomy. In the first stage an abstract frame is reconstructed by means of the external perspective of the nonparticipant. (BFN, 118) This is the perspective of the philosopher who '*tells* citizens which rights they *should* acknowledge mutually if they are legitimately to regulate their living together by means of positive law'. (BFN, 126)

Hence, the conditions for legal and political legitimacy are at hand from the beginning. Moreover, the system contains the basic rights that citizens must grant each other, not only to be *addressees* of the legal order that regulates their political life but also its *authors*. In other words, the system of rights protects private autonomy and at the same time, establishes the necessary conditions for the exercise of public autonomy. In the second stage, the system is shaped and gains concrete form by means of public autonomy. This is where pragmatic, ethical-political and moral reasons (BNF, 108) are incorporated and the system of *unsaturated placeholders* is fleshed out with content. Strictly speaking, this stage is not performed within the theory. Rather, it is shaped by means of self-legislation.

There are good arguments to maintain that this is the correct method to employ. Habermas asserts that, 'moral principles must be embodied in the medium of coercive and positive law'. (MW, 113) In this respect, Habermas argues that basic rights are both moral and legal entities (BFN, 454) that have to be justified twice – morally and legally –. Hence, this means that the moral content is not just the thin and narrow morality of law that Fuller identifies. Rather, it pertains to the post-conventional notion of normative justification which is reconstructed in the program of Discourse Ethics.

Thus, a 'two-stage' reconstruction is necessary to include these principles in the medium of law, and this is what is assured at the first stage. Otherwise, whether self-legislation would or would not incorporate this kernel is an open question. Historical cases show that by means of the exercise of popular sovereignty or by the will of the political legislator, undemocratic or even tyrannical outcomes might take place.

The 'two-stage' reconstruction is the best methodology to respond to the question of legitimacy. It ensures that in the first stage the Kantian notion of practical reason will be included in the 'system of rights' by means of the embodiment of procedures in the legal medium. (Forst, 2011; Flynn, 2003, 2011) The 'system of rights' as a basic structure of *unsaturated placeholders* is justified by the normative notion of post-conventional justification. (2011; Habermas, 2011: 296) This 'system of rights' is oriented towards guaranteeing private and public autonomy and sets the basic conditions that the practice of self-legislation meets, if it is going to be legitimate not only in terms of its procedures but also in terms of its outcomes. Furthermore, when at the second stage substance is included, moral reasons have priority over other considerations, namely, ethical-political or pragmatical. (BFN, 103) As Finlayson argues, 'no legitimate law may violate any valid moral norm. He calls this the moral permissibility constraint. (2011: 12; 2016a: 12, 14) In summary, the two-stage model enables levels of normativity to be ordered in the way that Habermas favours — firstly, moral reasons, secondly, ethical reasons, and thirdly, pragmatic reasons. This is because in Habermas's view, moral reasons have been already filtered by the normative net of the principle of discourse – via the principle of universalization. In this way, they have priority over other considerations. Because the MPC guarantees that legal systems enshrine human rights as basic rights, Habermas can claim that legitimate laws always operate within the bounds of moral permissibility. This is the priority of the moral.

2.2. Concerning the second criticism, in the same volume where Frank Michelman develops his criticism, Habermas writes a brief reply. The criticism states that the 'system of rights' cannot be abstractly reconstructed because a legal order consists of substantive laws from the beginning. In this same vain, Michelman charges Habermas with endorsing the distinction between questions of justification and questions of application. From Michelman's Hegelian point of view, the norms that compose the rule of law need to include in their definition the conditions of their application, otherwise they cannot rule and be enacted. Thus, the system of rights needs to be modulated by the historical and

social conditions that it aims to regulate. In other words, the system does not depend only on its rational justifiability but also on the context dependent conditions of its applicability.

Faced with this criticism, Habermas claims that there is a common core to different constitutions, which does not depend on their particular histories. This kernel concerns the implementation of basic rights, which have a universalistic meaning (Habermas, 1998: 389) that is introduced at the first stage. At this point, it is important to emphasize that Habermas mentions the term 'universalistic' here when he is replying to criticisms to his political theory. In his *oeuvre*, this concept has a moral connotation which pertains to a validity claim to rightness, to the property of being morally justified. After all that he has written on Discourse Ethics, the term *universalistic* is a concept with a certain history and a heavy moral weight. Among other things, this universalism means a rejection of *ethical relativism* 'which holds that the validity of moral judgments is measured solely by the standards of rationality or value proper to a specific culture of form of life'. (MCCA, 121) Consequently, if the system of rights at the first stage already incorporates a moral kernel, as far as it has a universalistic meaning, the legal order cannot be grounded merely on the contingent contents of a particular culture, a national history or a form of ethical life.

So it is clear that the Kantian notion of moral reason has priority in Habermas's theory. It is a moral priority in the full sense, not just an instrumental and functional priority. Nevertheless, this does not mean that Habermas's rules out the historical and social dimensions. At the second stage the *framers of constitutions* flesh out the system of rights taking into account pragmatical, ethical-political and moral reasons. (BFN, 108) In this respect, Habermas asserts that the universalistic meaning that is at the centre of legitimacy is introduced alongside competing horizons of interpretation. (Habermas, 1998: 389-390) The substance that Michelman wants to include from the beginning is incorporated in Habermas at the second stage in which public autonomy shape the system of rights.

III. Morality, institutions, practices and the fact of pluralism

Section I and II specified that there is a normative Kantian ground for the production of legal and political legitimacy. In this final section the issue is whether and to what extent it is reasonable to expect that particular citizens will support and endorse this Kantian foundation. In other words, and following Hegel's critique of Kant, the question is how is it possible to develop a post-conventional morality and an accommodating *Sittlichkeit* at the

level of particular wills?[19] Now days this problem should be examined in light of a social context of value pluralism that did not apply in Hegel's time.

In order to answer this, some more or less sympathetic critics of Habermas have proposed to blur the distinction between morality and ethics. (Michelman, 1998; McCarthy; 1991; Bernstein; 1998 and Putnam; 2002) There are good reasons to reject this alternative, among other things the fact of pluralism.[20] Nevertheless, Habermas is correct when he claims that 'Rawls salvages a valuable insight of Hegel's critique of Kant', because 'Moral commands must be *internally* related to the life-plans and lifestyles of affected persons in a way they can grasp for themselves'. (MW, 112) The question is, is this also the case in Habermas's version of democratic and political legitimacy?

The 1995 debate between Rawls and Habermas is an excellent place to discuss this problem. Here, Habermas argues that due to the fact of pluralism, a procedural normative core is needed to establish principles of political justice. Nevertheless, this fact also makes this hardly achievable, if effectively pluralism really cuts deeper (Finlayson, 2011: 19) and it not only concerns different religions, ideologies but even conceptions of practical reason.

In the following sections, I will discuss the debate regarding the normative principles of justice and the return of the Hegelian challenge. This objection is based on Hegel's criticism of the will as a tester of maxims. Here, I focus on the dissociation between the normative principles of legitimacy and the particular wills of the citizens (1); then I examine a Habermasian answer to this challenge which relies on Habermas's Kantian concept of political autonomy. I suggest that this answer is theoretically plausible, but empirically difficult to achieve bearing in mind the fact of social and ideological pluralism. However, at the end I examine a Kantian solution that relies in a thin and weak '*concept of normative justification*'. (IO, 45) (2).

---

[19] It is important to remember that this is a pressing issue for Habermas because he admits that legitimacy demands the support of an active citizenry and an accommodating political culture. This thesis was examined in the first section of this Chapter. There I showed that Habermas admits that constitutional democracy demands an accommodating political culture. Now, the question is how to expect that the citizens will endorse this culture, if legitimacy has a moral core and this kernel is based on agent-neutral reasons and not in agent-relative reasons. This issue becomes even more difficult in a context of deep pluralism where even conceptions of practical reason are contested.

[20] The concept of *Constitutional Patriotism* (BFN, 500; Müller, 2007) is additional evidence which shows that Habermas supports the distinction in the political domain. The allegiance to a legal order does not depend on a particular national or ethical history. It is ground on the loyalty towards the universalistic *embers* of the constitution.

1. The normative principles of justice and the return of Hegel

In what follows, I discuss the 1995 debate between Habermas and Rawls concerning the normative principles of justice. (RPR, 34-40; RH, 55-63; MW, 92-113) The picture that arises from the debate in general, and in this criticism in particular, is that the *weak* Kantian component has priority over the historical and functional arguments in Habermas's version of legal and political legitimacy. This has been read in the sense that normative moral components are crucial to political legitimacy (1.1). Then, from a Hegelian point of view a criticism returns because it seems problematic that constitutional democracy can expect the support of the citizens when the fundamental reasons that constitute the moral core of political legitimacy have to be agent-neutral (1.2).

1.1. Habermas opens his second objection to Rawls discussing the *political* meaning of *justice as fairness*. This feature is motivated by the fact of social and ideological pluralism that is the main problem that *Political Liberalism* addresses. The question that this theory aims to answer is:

> How is it possible that there can be a stable and just society whose free and equal citizens are deeply divided by conflicting and even incommensurable religious, philosophical and moral doctrines. (PL, 133)

Rawls argues that the political conception needs to 'be neutral toward conflicting worldviews' (RPR, 34) and this means that it does not go beyond the political realm – for example it does not reach the controversial arenas of morality, metaphysics and religion. Otherwise, it would not be reasonable to expect that the citizens would endorse this political conception, due to their different interpretations of the good. Thus, Rawls aims to rule out any comprehensive doctrine as the justificatory ground of the political conception.[21]

According to Habermas, this 'strategy of avoidance' (MW, 92) indicates certain 'unclarity' and 'indecisiveness as to how the validity claim of the theory itself should be understood'. (RPR, 34) In this respect, this strategy has to do with not presenting the theory as valid because true, but as valid because justified by shared political values. The heart of Habermas's criticism is revealed when he asks 'whether the overlapping consensus, on which the theory of justice depends, plays a cognitive or merely instrumental role'. (RPR,

---

[21] Rawls's concept of 'comprehensive doctrine' has three different senses: a) world views, religious and secular; b) actually existing moralities, or conceptions of the good; c) philosophical theories of one kind of another, including moral theories such as utilitarianism or Kantianism. (PL, xviii, Finlayson, 2016a: 8)

34) Rawls seems to be more concerned with the question of 'whether a society constituted in accordance with the principles of justice could stabilize itself' (RPR, 35) and therefore, Habermas argues, he fails to do justice to the 'epistemic meaning' of his own theory, because he does not recognize that it necessarily makes a claim to truth. This problem is obscured by simply introducing the predicate 'reasonable' instead of 'true' and this distinction for Habermas has an ambiguous meaning. (RPR, 38-40) Due to this strategy of avoidance, Habermas contends, Rawls collapses the distinction between justified acceptability and actual acceptance. (RPR, 36) In this respect, if Rawls is still part of the Kantian family, then he needs to clarify and acknowledge that the overlapping consensus possess a cognitive meaning and not a purely functional role to guarantee social stability.

Rawls in his rejoinder asserts that the answer to these criticisms 'lies in the way in which political liberalism specifies three different kinds of justification and two kinds of consensus'. (RH, 56) At this point, I do not need to develop Rawls's defence. Rather, I focus on his criticism of Habermas that goes alongside his reply. For the former, the latter takes 'too many theoretical hostages to fortune' (Finlayson, 2011: 19) and this charge is explained by Rawls's claim that Habermas's position is ''comprehensive' while mine is an account of the political and it is limited to that'. (RH, 47) According to Rawls, the specific kind of comprehensive doctrine that Habermas develops is a philosophical theory, while *justice as fairness* concerns the *political* 'leaving philosophy as it is'. (RH, 48)

In his reply, Habermas does not contest the charge that his theory is 'comprehensive' and philosophical. (MW, 93) Rather, his claim is that his version of legal and political legitimacy and Rawls's 'cannot ultimately avoid giving full weight to requirements of practical reason that *constrain* rational comprehensive doctrines rather than merely *reflect* their felicitous overlapping'. (MW, 94) According to Habermas, a political theory that claims to be neutral with respect to worldviews, cannot move itself entirely within the narrow domain of the 'political', as Rawls construes it (RH, 47) 'and steer clear of stubborn philosophical controversies'. (MW, 93) This condition of neutrality undermines the *epistemic* status of the political conception, and sets at the centre of legal and political legitimacy a normative kernel.

Habermas has good arguments to justify the claim that his theory of legitimacy is not comprehensive and philosophical in the negative way that Rawls asserts it is, even if it incorporates at its core a Kantian conception of moral reason. Habermas's account of normativity is post-metaphysical, and therefore it rules out any sort of metaphysical

doctrine of reason and the world (PT: 3-9); and it is procedural and not substantive. (MCCA, 197) Moreover, his principle of justification has a *weakly* transcendental status, because validity claims are embodied in the medium of social practices and institutions, although they appeal to standards of rightness which are context transcending.

In this respect, the Habermas-Rawls debate leaves room for an interpretation in which morality has a central place in Habermas's theory of legal and political legitimacy. Now, at this point let us recall the main features of legitimate laws in Habermas's account of them.[22] The first is that a particular law is legitimate if has been properly passed by a recognized legal process (principle of democracy); the second is that it does not violate any valid moral norm MPC (*moral permissibility constraint*); and the third is that it achieves some measure of congruence from each citizens's ethical values.

In his understanding of Habermas's political theory, Baynes reads the centrality of morality and the fact that legitimate laws does not violate any valid moral norm in the sense that legitimacy and legitimate laws contain a moral core. In this regard, he asserts that 'Citizens must simultaneously both presuppose and strive to articulate a basic political consensus […] focused on the idea of a core morality that all citizens can endorse as valid for the same (publicly available) reasons'. (Baynes, 2016: 179) Baynes seems to think that this moral core comes about because democratic constitutions legally enshrine human rights as basic rights, and human rights have a fully moral content. In virtue of this moral core, not only valid moral norms, but also legitimate laws satisfy the validity requirement expressed by Baynes: 'basic political norms […] are legitimate only if they […] could be agreed to by all citizens as participants in a practical discourse for the same (publicly available) reasons'. (2016: 170)

If Baynes's interpretation is correct then whenever we look at the conditions that legitimate laws have to meet, then we need to examine the validity requirement of morality. Concerning Habermas's understanding of the validity of moral norms, it has been argued (see Chapter Two) that for him valid reasons are agent-neutral and not agent-relative and only the former type are apt to provide a normative moral justification.[23] The fact that at the centre of legitimacy he installs a Kantian notion of practical reason and autonomy

---

[22] In Chapter Three and in this Chapter I have discussed in detail these features.
[23] As shown by Finlayson, there are agent-relative reasons that can be universalizable. For example, everyone has a reason to avoid pain, to themselves. Therefore, the moral worth of reasons does not get totally captured by their universalizability. Rather, their moral status depends in the fact that they are universalizable but at the same time exactly the same reason for everyone. This can account as a moral reason. (Finlayson, 1999) This issue it is discussed in detail Chapter Two, section 3.

explains that he also understands his debate with Rawls in terms of the agent-relative and agent-independent distinction. Habermas charges Rawls that in the latter's theory agent-independent reasons support 'the "reasonableness" of a political conception, while […] agent-relative reasons establish the claim to moral "truth"'. (MW, 94) In contrast, for Habermas agent-independent reasons furnish a normative core of legitimacy and not the agent-relative reasons of the citizens embedded in their particular comprehensive doctrines. In Habermas's words:

> Reasonable citizens cannot be expected to develop an overlapping consensus so long as they are prevented from jointly adopting a moral point of view independent of, and prior to, the various perspectives they individually adopt from within each of their comprehensive doctrines. (MW, 94)

In the final analysis for Habermas agent-neutral reasons are at the centre of legitimacy and this implies that a version of the Hegelian question returns. This is because, from the particular point of view of the citizens, it does not seem to be plausible that they will find the motives to sustain both the political conception and legitimate laws, without bringing their agent-relative reasons into play.[24] This problem is even more burdensome bearing in mind the fact of pluralism that not only pertains to conceptions of the good but also to the concept of moral reason itself. In what follows, I expound this problem (1.2).

1.2 Considering both the place that a moral Kantian core has in Habermas's political theory – following Baynes's interpretation – and that these principles, Habermas claims, can only be established by agent-neutral reasons, the problem that I want to examine is how individuals with their own agent-relative motives, can relate to legitimate laws. This issue is connected to Hegel's critique of Kant, namely, the break between the concept of the will and the empirical will.

According to Hegel in the *Philosophy of Right,* Kant's concept of the will is distinguished by '*formal correspondence with itself,* which is no different from the specification of *abstract indeterminacy*'. (§135) Thus, the moral will is only identical with itself and it cannot incorporate particular contents and interests and consequently remains purely abstract. In

---

[24] It is important to remark that this reconstruction of Habermas's theory relies on commentators like Baynes. Surely, his interpretation is based on Habermas's letter. That said, from another interpretation of Habermas's theory of legitimacy is possible to claim that he does not say that political norms can only be justified by moral reasons and as moral reasons (agent-neutral reasons). Legitimate laws can be justified not only by moral reasons but also by ethical and pragmatic reasons. And the latter two may well be seen as agent-relative considerations.

this regard, as far as the Kantian concept of moral validity rules out agent-relative reasons, it is incomplete, it is an empty formalism. Hegel's concept of ethical life was his strategy to solve this issue. By means of *Sittlichkeit* 'the individual can be 'at home with himself'' (Pinkard, 1999: 227) and she can find her *particular* way to be moral.

Due to his Kantian concept of practical reason, Habermas's theory is not immune to this Hegelian criticism. In light of Baynes's reading – like moral norms – legitimate laws are valid when they are shaped in everyone's interest, when they are universalizable. And valid norms are grounded on agent-neutral reasons, and agent-relative motives are not susceptible of being universalized, because they refer back to particular, individual agents and their 'pathological' feelings and inclinations. Legal and moral norms are not justified by agent-relatives motives, and this condition leaves 'the epistemic core of moral validity intact'. (IO, 7) This is related Habermas's account of the cognitive content of normativity which rests on the fact that validity claims to rightness are analogous to validity claims to truth. (MCCA, 196) Inasmuch as Habermas endorses this distinction – which allows him to secure the epistemic status of his concept of practical reason and normative rightness – the question that returns is how the agents can support the normative core that legal and political legitimacy demand.

Modern legal orders have a component that guarantees the obedience among citizens: the threat of sanctions. As a matter of fact, this dimension of law targets one of the deficits that 'accompany a post-conventional rational morality', namely, 'the motivational uncertainty that results from the fact that moral insight does not guarantee compliance'. (Baynes, 2016: 137) Morality is binding but has to be self-imposed. It cannot be enforced externally, although, someone can have moral feelings of shame, blame, etc., whenever she violates a moral norm. Thus, modern law aids morality to stabilize behavioural expectations and supports the integration of modern societies. Nevertheless, a normative order cannot be based only on this functionalist feature. It also needs the quasi-voluntary and rational assent of the citizens, and in Habermas's picture, this legitimacy is based on a modified version of Kant's concept of autonomy. In this regard, the debate with Rawls can be interpreted in the sense that the citizens are not only conceptually bound by a legal and political order, but also they are 'morally' bound. In other words, they have a moral commitment, so to speak, with constitutional democracy.

Considering this reading of Habermas's political theory, then legitimacy is very hard to achieve. Habermas can afford as he does to admit that morality (valid norms that pass the

test of universalization) is a scarce resource in modern pluralist, multicultural societies. (JA 91) But legitimate law is supposed to compensate for this. As I have shown, legitimate law, according to Habermas, functions in modern societies as a compensation for the diminishing moral (and shared ethical) substance. And thus helps achieve social integration, and it can only provide that help because it is more abundant and easier to 'produce' than valid moral norms. (BFN 98-99) Hence, in introducing his revisions, Baynes weakens Habermas's sociological account of the socially-integrating function of law and democracy.

2. Kantian Republicanism and the fact of pluralism

In what follows, I discuss a Habermasian answer to the Hegelian criticism of the will as a tester of maxims this time applied to legal and political legitimacy. According to Habermas's Kantian Republicanism the freedom of the citizen, which certainly appeals to his agent-relative reasons, depends on everyone else's freedom, which is an agent-neutral reason. Thus, in this concept of autonomy – which incorporates an element of mutual recognition – it is possible to answer to the Hegelian question (2.1). This solution is theoretically plausible but socially difficult to achieve, if the fact of pluralism not only concerns different religions and ideologies but also conceptions of practical reason. However, I examine the plausibility of a Kantian answer to the dilemma opened by the fact of pluralism (2.2).

2.1 From his Kantian Republicanism Habermas claims that 'Nobody can be free at the expense of anybody else's freedom'. (MW, 113) This formulation is an echo of Kant's formula of humanity of the categorical imperative, namely, the recognition of other as ends in themselves. (4:429) Besides, it incorporates an element of mutual recognition, as I think Kant's formulation already does. Applied to the political realm, this concept becomes the notion that the freedom of the citizen, which certainly appeals to her agent-relative reasons, depends on everyone else's freedom, which is an agent-neutral reason. Thereby, by means of his Kantian Republicanism Habermas solves the Hegelian question concerning the reconciliation between the concept of the will and the particular will, such that the agent-relative freedom of the person is co-original with the agent-neutral freedom of the other persons.

This solution squares with Hegel's insight that the individuation of particular agents takes place by means of practices of socialization. (MCCA, 199) In this regard, the freedom of the 'individual cannot be tied to the freedom of everyone else in a purely negative way,

through reciprocal restrictions'. (MW, 113) Rather, the social dimension of freedom can only be actualized in the frame of an association of free and equal persons, where 'all members must be able to understand themselves as joint authors of laws to which they feel themselves bound individually as addresses'. (MW, 113) Consequently, the autonomy of the particular person depends on the mutual recognition of the autonomy of everyone else.

Now, the Habermasian response to the Hegelian challenge seems to be coherent and cogent from the theoretical point of view. However, it does not address properly the question concerning the motivation of the participants. In this respect, Finlayson claims that Habermas was wrong when he excluded agent-relative reasons from moral justification because some agent-relative reasons are universalizable. (1999) The solution that Habermas and Habermasians could find to this Hegelian charge would be to relax this condition and allow that some agent-relative reasons can justify norms.

In the Kantian and Hegelian idea of mutual recognition there is a dialectical relationship between agent-relative and agent-neutral reasons that supposes the loosening of the distinction and this results in the solution of the problem. In this way, the agent-relative reasons of the particular persons who are *addressees* but also *authors* of the legal order are co-original with the agent-neutral insight that their private autonomy emerges only when is enjoyed by everybody else. By means of this dialectic between the particular will – agent-relative reasons – and the normative will – agent-neutral reasons – that the Kantian notion of self-legislation incorporates, the Hegelian problem can be solved. Albeit: this demands that we slacken blur the distinction between agent-relative and agent-neutral reasons.

In summary, I have answered Hegel's critique of the will as a tester of maxims from the point of view of Habermas's Kantian Republicanism. Nevertheless, the picture gets more complicated if we discuss this problem in view of the fact of pluralism. In this context and considering the perspective of the agents, the question returns because how is it reasonable to expect that empirical persons would put into practice a procedural notion of autonomy supporter by their particular points of view?

2.2 The Habermasian solution to the Hegelian challenge seems problematic to achieve in the context of social and ideological pluralism. If this fact really cuts deep (Finlayson, 2011: 19) and 'reasonable disagreement about the nature of the good life and even about the philosophical foundations of morality […] is the situation we should expect' (Larmore,

1995: 63) then, how is it possible to presuppose that the citizens will comply with laws based on publicly available (agent-neutral) reasons?

In my interpretation of the Habermas-Rawls debate, in the end the fact of pluralism demands a Kantian kernel for the production of legal and political legitimacy: it is no longer possible to ground legitimacy on a particular conception of the good – religious, philosophical or of any other sort. At the same time, in light of one of the Hegelian insights that Habermas incorporates, constitutional democracy requires an active citizenry that endorses these moral norms. Otherwise, it cannot be effective. Notwithstanding, the fact of pluralism makes it problematic to achieve this Hegelian end, because today the social and cultural context is not the same as it was in Hegel's time. Hence, it is unreasonable to expect a common *Sittlichkeit* that will support a normative notion of post-conventional justification at the centre of legal and political legitimacy. Moreover, and following the results of the previous section, the existence of citizens endorsing principles of political justice seems at least difficult in the frame of a theory that strictly distinguishes between agent-relative and agent-neutral reasons.

In this respect, the gap between the concept of the will and the particular will that Hegel founded on Kant's moral philosophy arises once again at the political stage. To address this issue one option would be to revisit Rawls's alternative. According to his philosophy, the support that *Political Liberalism* demands of the citizens arises from the comprehensive doctrines that each of them endorse. However, the answer that Rawls can provide depends on the interpretation of his concept of *overlapping consensus*. Habermas's critique shows that Rawls is not sufficiently clear in this respect and this ambiguity does not rule out the *modus vivendi* interpretation of *Political Liberalism*. (RPR, 37) Nonetheless, Rawls certainly does rule out the *modus vivendi* interpretation of *Political Liberalism*. (2011, 2016a) Moreover, the opposite version that will give a centrality to a Kantian core in Rawls's theory still needs to be convincingly proved.

Another option is to blur the distinction between normativity, morality and ethics, as some critics of Habermas propose. (Michelman, 1998; McCarthy; 1991; Bernstein; 1998 and Putnam; 2002) Notwithstanding, this alternative does not seem to square with the fact of pluralism. In a context of diversity in terms of particular conceptions of the good the distinction between ethical and deontological reasons is necessary because it motivates citizens to distance themselves from their ethical commitments and view specific

disagreements in a more general light. (Müller, 2007) It is not possible to expect that political and legal issues can be solved satisfactorily appealing to ethical commitments.

There is still a Kantian alternative to envision a possible answer this dilemma. In a first step, it is important to elucidate what is the object that needs to be embraced by the citizens at the centre of legal and political legitimacy. This content is the principle of discourse (D), which is a weak '*concept of normative justification*' (IO, 45) that refers to a 'thin […] idea of justification in which individuals are viewed as mutually accountable agents'. (Baynes, 2016: 115) It is weak because it does not incorporate a particular conception of the good life, but relies purely on a procedural notion of practical reason. Furthermore, it is non-negligible, since if someone wants to participate in legal and political discourses it is undeniable. Why? Because whenever a person wants to utter validity claims in discourse, she needs to recognize the same possibility for everybody else. (T&J, 94)

Once this weak '*concept of normative justification*' is embodied in the legal medium, it emerges the 'system of rights' which afterwards gains shape through political self-determination. This content is the pre-condition of constitutional democracy. It is its condition of possibility, because it is impossible to imagine this practice without this Kantian element of mutual recognition. Now, is it possible that someone can reject this *weak* and *thin* content and at the same time be reasonable? It is true that today is not appealing to make claims for some immovable content as the ground of legal and political. However, the *weak* and *thin* character of the Kantian content that legitimacy demands at its core makes this call less demanding. Moreover, in so far as rational discourse does not refer to a substantive norm, but to a procedure, it is reasonable to expect that citizens will put it into motion in a context of pluralism. This practice demands certain commitments and rules, for example, they have to recognize others as autonomous and equal participants in discourse; the inclusion of all the affected; and sincerity. The everyday practice of giving justifications shows that in way or another we are committed with these rules. Otherwise, perhaps we would not say anything.

Conclusion

In the introduction I mentioned the 'Kant oder Hegel?' question and I claimed, following Robert Brandom, that Habermas sees himself closer to Kant. (2015: 34) Many passages of his *oeuvre* support this interpretation. However, his theory takes on board Hegel's critique of Kant's moral philosophy. Consequently, the normative foundations of Habermas's theory incorporate Kantian and Hegelian components. At the same time, I have shown that Hegel is best seen as extending Kant's rationalist morality by critiquing and supplementing it, but not rejecting it. This work demonstrates that this is also the spirit that informs Habermas's project of social criticism. Habermas asserts that he is a Kantian and according to the *Post-Kantian* interpretation Hegel was also a Kantian.

In light of this background, in what follows, I will sketch the main conclusions that arise from this thesis. But first it will help to recall the main pillars of the argument. The first was referred already in the words that I have said at the beginning of this conclusion, namely, the *Post-Kantian* interpretation of Hegel. From the latter it is possible to elucidate a second: Hegel's critique of Kant. This is because the *Post-Kantian* reading provides a particular way to understand the intention and the letter of Hegel's critique. A third component that allowed me to set the stage of this thesis is Habermas's Kantianism. I will begin discussing the latter.

It is not difficult to find the name of Kant in Habermas's project of social criticism. However, it is more problematic to examine to what degree Habermas's is a Kantian. In Habermas's *oeuvre* the presence of Kant is wide and deep. And since STPS, towards BFN and in the debate with Rawls, the philosopher of Königsberg seats at the front. In Chapter One I focused on Habermas's moral theory and I studied its Kantian main features. In Chapter Three I examined Habermas's Kantianism in his political philosophy and I concluded that this theory also has at its centre Kantian components. In these two

Chapters I proved that Kant is not only an important philosopher with whose work Habermas engages, but rather, that Kant helps Habermas to set the foundations of his critical theory.

Inasmuch as Habermas incorporates this Kantian component at the core of his theory, then it is relevant to examine whether and to what extent Hegel's critique of Kant applies to Habermas. Certainly, Habermas develops a *modified* and *attenuated* version of Kant's notion of moral reason. And these features allow him to claim that his project of social criticism can consistently rebut Hegel's critique. Indeed, Habermas addressed the issue over four decades ago (MCCA, 195-215) and his answer has convinced many critics that we are not dealing with a real problem. That said, Habermas takes a *philosophical hostage to fortune* which makes his Kantian position vulnerable to Hegel's critique of Kant, namely, the stringent distinction between agent-neutral and agent-relative reasons, and the claim that only the former are apt to justify moral norms. Habermas's moral theory trades on a view of impartiality in which valid norms have to be justified by agent-neutral reasons and never by agent-relative considerations. This reconstruction of impartiality is explicitly mentioned in Habermas's Discourse Ethics. (IO, 7, 43) And in this theory he seems to conflate agent-relative reasons with reasons that are particular, i.e. in each case only mine. Hence, he does not allow that reasons of the kind of – I want to avoid my own suffering – apply relatively to everyone. Rather, moral justification refers to norms which discard all sorts of considerations that refer back to the particular person (and her interests, inclinations and needs).

In my opinion, this strict distinction between agent-relative and agent-neutral reasons implies that Hegel's criticism of the will as a tester of maxims applies to Habermas. In Chapter Two I examined in detail this objection. Roughly speaking, this charge states that Kant's moral philosophy introduces a break between the will of morality (universal laws) and the particular will of the agent (her interests, inclinations and needs). Insofar as the latter criticism applies to Habermas, a twofold dilemma arises. The first part is that Habermas cannot explain how morality can include enough content to guide moral agents in social space and historical time. The second part is that he cannot provide an adequate account of how particular agents can be motivated to act according to moral norms. In Chapter Two I address the first part of the dilemma. I think that only with that answer we can begin to properly address the second part. Indeed, from a Kantian point of view if one

answers the first part then the second fades because, as Korsgaard claims, 'the reasons why an action is right and why you do it are the same'. (Korsgaard, 1986: 10)

My proposal to deal with this issue rests on slackening the strong distinction between agent-relative and agent-neutral reasons. By an analysis of two other components of Habermas, namely, the idea of mutual recognition and the concept of ideal role taking, I argue in Chapter Two that this modification is the best alternative to fix this problem. In the concept of ideal role taking the particular will of the agents are channelled and shaped in terms of the universal will of morality. Why? Because throughout the process of discourse the agents take as their own the point of view of impartiality. In this regard, agent-relative considerations become agent-neutral reasons, or rather the commitment to give equal weight to other people's agent-relative reasons is agent-neutral. Likewise, in the idea of mutual recognition, the agent-relative interest of the particular will to be recognized can only be actualized inasmuch as the agent-neutral reason of everyone else to be recognized is also taking into account. Both solutions imply that in Habermas the strict distinction between agent-relative reasons and agent-neutral reasons is slackened. In other words, ideal role taking and mutual recognition show that moral justification can rest not only on agent-neutral reasons but also on universal agent-relative considerations as well.

With regard to Habermas's political theory, it is slightly more difficult to build Hegelian criticisms. In Habermas's understanding of Kant's *Rechtslehre* the former claims that the latter builds a monistic view that conflates morality and politics. In BFN Habermas avoids this sort of construction because it does not give a proper account of the complexity of modern societies. More precisely, it is implausible and incorrect to claim that morality alone is capable of integrating highly differentiated social orders. (BFN, 8, 29) Rather, it is necessary to rely on the functional features of modern law and politics. Thus, this work is more sociologically informed and appeals to an immanent reconstruction of legitimacy. And the latter means that the justification of legal and political norms depends on the internal features of the institutions and practices of modern law and democracy. To a certain degree, here Habermas closes the gates to any sort of moral justification of legitimacy. Nevertheless, Habermas takes few more *philosophical hostages to fortune*, namely, the *Moral Permissibility Constraint* which implies that legitimate laws must 'harmonize with the moral principles of justice and solidarity'. (BFN, 99, 155) And among the

considerations that can underpin legitimate laws moral reasons have priority. (BFN, 103, 108; IO, 42-43)

In writings after BFN Habermas begins to open the gates even more and it comes to seem that he builds a moral argument in the justification of political legitimacy. In his debate with Rawls, Habermas asserts that legitimacy demands a moral core which secures the epistemic status of the political conception. (MW, 97) He also says that political justice stands on its own moral feet. (MW, 111) Moreover, already in BFN one reads that Human Rights are legal entities and at the same time moral entities. (454) And in IO Habermas claims that moral considerations enter into the justification of Human Rights. (191)

In his interpretation of Habermas's political theory Kenneth Baynes takes these lines of argumentation to mean that legitimate laws contain a moral core. In this regard, he says that 'Citizens must simultaneously both presuppose and strive to articulate a basic political consensus (focused on the idea of a core morality mentioned above) that all citizens can endorse as valid for the same (publicly available) reasons'. (Baynes, 2016: 179) If this is correct, then not only valid moral norms but also legitimate laws need to satisfy the validity requirement, namely, they have to be justified on the basis of "the same' i.e. agent-neutral reasons, and not on agent-relative reasons. The upshot of this reading is that political legitimacy is quite hard to achieve. Habermas can afford to make this claim in the case of morality which becomes a scarce resource in modern societies (JA, 91) because legitimate law is supposed to compensate for this deficit. (IO, 256-58) Albeit, Baynes's stringent core morality weakens Habermas's account of the socially-integrating function of modern law and politics. The solution that I have proposed to fix this problem entails making the validity requirement of legal and political norms less stringent. In this view, not only valid moral norms but also legal norms can be justified on the basis of impartial reasons but also on the basis of universal agent-relative reasons.

Perhaps, the fact of reasonable pluralism invalidates this solution because it is difficult to expect that all citizens are going to put into practice rational discourse whenever they want to deal with conflicts of action. That said, I believe that the *thin* and *weak* features of Habermas's notion normative justification combined with the relaxed view of impartiality leave enough space to presume that, if not all, then most of the citizens can engage with this practice. Among other things, they have to recognize others as autonomous and equal

participants in discourse; they cannot reach an agreement and exclude some; and they have to be sincere whenever they utter validity claims. I do not think that the allegiance to these commitments and practices is exceptionally demanding. To finish just one brief example: the social struggles for recognition show the interest of the vast majority to take part in the construction of a public sphere in which the notions of inclusion, autonomy, publicity and equality are embedded.

Bibliography

Alexy, R (1990), 'A Theory of Practical Discourse,' in Seyla Benhabib and F. Dallmayr, eds. *The Communicative Ethics Controversy*. Cambridge, MA: MIT Press.

Allison, H (1990), *Kant's Theory of Freedom*. Cambridge: Cambridge University Press.

_____ (2004), *Kant's Transcendental Idealism*, New Haven: Yale University Press.

Apel, K (1998), *Auseinandersetzungen*. Frankfurt am Main: Suhrkamp.

Aristotle (2009), *Nichomaquean Ethics*. Oxford: Oxford University Press.

Baynes, K (1992), *The Normative Grounds of Social Criticism: Kant, Rawls and Habermas*. Albany: SUNY Press.

_____ (2016), *Habermas*. New York: Routledge.

Benhabib, S (1986), *Critique, Norm and Utopia*. New York: Columbia University Press.

_____ (1990), 'Afterword: Communicative Ethics and Contemporary Issues in Practical Philosophy', in S. Benhabib and F. Dallmayr (eds,) *The Communicative Ethics Controversy*. Cambridge, MA: MIT Press.

Bernstein, R (1998), 'The Retrieval of the Democratic Ethos', in M. Rosenfeld and A. Arato (eds,) *Habermas on Law and Democracy: Critical Exchanges*. Berkeley and Los Angeles: University of California Press.

Bohman, J and W. Rehg (2014), 'Jürgen Habermas', *The Stanford Encyclopedia of Philosophy* (Fall Edition), Edward N. Zalta (ed.).

Brandom, R (1994), *Making It Explicit*. Cambridge, MA: Harvard University Press.

_____ (2015), 'Towards Reconciling Two Heroes: Habermas and Hegel', *Argumenta, The Journal of the Italian Society for Analytic Philosophy*. 1(1).
Cohen, J (1994). 'Pluralism and Proceduralism', *Chicago-Kent Law Review* 69: 589–618.

Cohen, J (1999), 'Reflections on Habermas on Democracy', *Ratio Juris* 12: 385–416.

Constant, B (1988), *Political Writings*. Cambridge: Cambridge University Press.

Cortella, L (2015), *The Ethics of Democracy: A Contemporary Reading of Hegel's Philosophy of Right*. SUNY Press.

Fichte, J (2000), *Foundations of Natural Right*. Cambridge, United Kingdom. Cambridge University Press.

Finlayson, J (1999), 'Does Hegel Critique of Kant's Moral Theory Apply to Discourse Ethics?', In Dews, Peter, ed, *Habermas: A Critical Reader*. Malden, MA: Blackwell.

_____ (2000), 'Modernity and Morality in Habermas's Discourse Ethics', *Inquiry* 43: 319–340.

_____ (2005), 'Habermas's Moral Cognitivism and the Frege-Geach Challenge', *European Journal of Philosophy* 13: 319–344.

_____ (2011), 'Introduction: The Habermas Rawls Dispute — Analysis and Reevaluation', in J. G. Finlayson and F. Freyenhagen (eds) *Habermas and Rawls. Disputing the Political*. New York: Routledge.

_____ (2016a), 'Where the right gets in. On Rawl's criticism of Habermas's conception of legitimacy', *Kantian Review,* 21 (July).

_____ (2016b) 'Structural Transformations of Discourse Theory: Habermas on 'Justice' and the Right and the Good'. Manuscript.

Flynn, J (2003), 'Habermas on Human Rights: Law, Morality and Intercultural Dialogue', *Social Theory and Practice,* Vol 29, No. 3.

_____ (2011), 'Two Models of Human Rights: Extending the Rawls–Habermas Debate', in J. G. Finlayson and F. Freyenhagen (eds,) *Habermas and Rawls. Disputing the Political*. New York: Routledge.

Forst, R (1994), *Contexts of Justice. Political Philosophy beyond Liberalism and Communitarianism*. University of California Press.

_____ (2007), *The Right to Justification*, Trans Jeff Flynn. New York, Colombia University Press.

_____ (2011), 'The Justification of Justice: Rawls and Habermas in Dialogue', in J. G. Finlayson and F. Freyenhagen (eds,) *Habermas and Rawls. Disputing the Political*. New York: Routledge.

Freyenhagen, F (2011), 'Empty, Useless, and Dangerous? Recent Kantian Replies to the Empty Formalism Objection', *Bulletin of the Hegel Society of Great Britain*. 63: 163: 186.

Fultner, B (2003), 'Translator's Introduction'. In T&J.

Fuller, L (1969), *The Morality of Law*. New Haven CT: Yale University Press.

Habermas, J (1998), 'Reply to Critics', in M. Rosenfeld and A. Arato (eds,) *Habermas on Law and Democracy: Critical Exchanges.* Berkeley and Los Angeles: University of California Press.

_____ (2001), 'Constitutional Democracy: A Paradoxical Union of Contradictory Principles?', *Political Theory,* XXIX, 6.

_____ (2010), 'The Concept of Human Dignity and the Realistic Utopia of Human Rights', *Metaphilosophy,* XVI, 4.

_____(2011), 'Reply to Critics', in J. G. Finlayson and F. Freyenhagen (eds,) *Habermas and Rawls. Disputing the Political.* New York: Routledge.

Hegel, F (1977a), *Phenomenology of Spirit.* Oxford: Oxford University Press.

_____ (1977b), *Natural Law.* Philadelphia, PA: University of Pennsylvania Press, 1977.

_____(1991), *Elements of the Philosophy of Right,* trans. H. B. Nisbet. Cambridge, United Kingdom. Cambridge University Press.

Hedrick, T (2010), *Rawls and Habermas. Reason, Pluralism, and the Claims of Political Philosophy.* Standford, California: Standford University Press.

Honneth, A (1991), *Critique of Power: Reflective Stages in a Critical Social Theory.* Cambridge: MIT Press.

_____ (2014), *Freedom's Right: The Social Foundations of Democratic Life,* trans, Joseph Ganahl. Cambridge, MA: The MIT Press.

Höffe, O (1994), *Immanuel Kant,* Albany, NY: SUNY Press.

Hume, D (2000), *A Treatise if Human Nature.* New York: Oxford University Press.

Ingram, D (2010), *Habermas: Introduction and Analysis.* Ithaca, NY: Cornell University Press.

Kohlberg, L (1981), *The Philosophy of Moral Development: Moral Stages and the Idea of Justice.* San Francisco, California: Harper & Row.

_____ (1984), *The Psychology of Moral Development: The Nature and Validity of Moral Stages.* San Francisco, California: Harper & Row.

_____ (1986), 'A current statement on some theoretical issues'. In S. Modgil & C. Modgil, (Eds.), *Lawrence Kohlberg: Consensus and controversy.* Philadelphia, PA: The Falmer Press.

Korsgaard, C (1985), 'Kant's formula of universal law'. *Pacific Philosophical Quarterly* 66, no. 1-2: 24-47.

_____ (1986), 'Skepticism About Practical Reason', *Journal of Philosophy* 83, n°. 1: 5-25.

_____ (1993), 'The reasons we can share: An attack on the distinction between agent-relative and agent-neutral values'. In *Altruism*, ed. Ellen Frankel Paul, Fred Dycus Miller, and Jeffrey Paul. Cambridge: Cambridge University Press.

Laden, A (2011), 'The Justice of Justification', in J. G. Finlayson and F. Freyenhagen (eds,) Habermas and Rawls. Disputing the Political. New York: Routledge.

Lafont, C (2004). 'Moral Objectivity and Reasonable Agreement: Can Realism be Reconciled with Kantian Constructivism?,' *Ratio Juris* 17: 27–51.

Larmore, C (1995), 'The Foundations of Modern Democracy: Reflections on Jürgen Habermas', *European Journal of Philosophy,* 3:1.

Luhmann, N (1996), *Social Systems,* Stanford University Press.

_____ (2004), *Law as a Social System,* Oxford: Oxford University Press.

Mackie, J (1977), *Ethics: Inventing Right and Wrong*, Harmondsworth: Penguin.

Maus, I (1992), *Zur Aufklärung der Demokratietheorie. Rechts- und demokratietheoretische Überlegungen im Anschluß an Kant.* Frankfurt am Main: Suhrkamp.

_____ (1995), 'Liberties and Popular Sovereignty: On Jürgen Habermas's Reconstruction of the System of Rights', 17 *Cardozo Law Review.* 825.

Mead, G (1967), *Mind, Self and Society, from the standpoint of a social behaviourist.* Chicago, Illinois: Chicago University Press.

McCarthy, T (1978), *The Critical Theory of Jürgen Habermas.* Cambridge, MA: MIT Press.

_____ (1991), *Ideals and Illusions.* Cambridge, MA: MIT Press.

McMahon, C (2000), 'Discourse and Morality', *Ethics* 110: 514–536.

Michelman, F (1996), 'Between Facts and Norms by Jürgen Habermas', *Journal of Philosophy.* 93 (6):307-315.

_____ (1998), 'Family Quarrel', In M. Rosenfeld and A. Arato (eds,) *Habermas on Law and Democracy: Critical Exchanges.* Berkeley and Los Angeles: University of California Press.

Müller, J (2007), *Constitutional Patriotism.* Princeton: Princeton University Press.

Nagel, T (1986), *The View From Nowhere,* New York: Oxford University Press.

O'Neill, O (1975), *Acting on Principle.* New York: Columbia University Press.

_____ (1989), *Constructions of Reason: Explorations of Kant's Practical Philosophy.* Cambridge: Cambridge University Press.

Peters, B (1994), 'On Reconstructive legal and Political Theory.' *Philosophy and Social Criticism* 20, no. 4 (1994): 101–134.

Pettit, P (1982), 'Habermas on Truth and Justice'. *Royal Institute of Philosophy Lectures* 14:207-228.

Pinkard, T (1994), *Hegel's Phenomenology: The sociality of Reason.* Cambridge: Cambridge University Press.

_____(1999), 'Virtues, Morality, and Sittlichkeit: From Maxims to Practices', *European Journal of Philosophy*, 7: 2.

_____ (2002), *German Philosophy 1760-1860. The legacy of Idealism.* Cambridge: Cambridge University Press.

Pippin, R (1989), *Hegel's idealism: The Satisfactions of Self-Consciousness.* Cambridge: Cambridge University Press.

_____ (1991), *Modernism as a Philosophical Problem: On the Dissatisfactions of European High Culture.* Oxford: Basil Blackwell.

Putnam, H (2002), *The Collapse of the Fact/Value Distinction and Other Essays.* Cambridge: Cambridge University Press.

Reath, A (1997), 'Introduction'. In Kant's *Critique of Practical Reason.*

Rehg, W (1991), 'Discourse and the moral point of view: Deriving a dialogical principle of universalization', Inquiry, 34:1, 27-48

_____ (1994), *Insight and Solidarity: A Study in the Discourse Ethics of Jürgen Habermas.* Berkeley, CA: University of California.

_____ (1996), 'Introduction', in BFN.

Rosenfeld, M and A. Arato, eds. (1998). *Habermas on Law and Democracy.* Berkeley, CA: University of California Press.

Snare, F (1990), *The Nature of Moral Thinking.* London, Routledge.

Savigny, F (1980), *System of the Modern Roman Law,* Hyperion.

Thomassen, L (2010), *Habermas: A Guide for the Perplexed.* London: Continuum.

Weber, M (1978), *Economy and Society: An outline of interpretative Sociology*, Berkeley.

Wellmer, A (1991), *The Persistence of Modernity.* Cambridge, MA: MIT Press.

_____ (1998), *Endgames: The Irreconcilable Nature of Modernity*, Cambridge: MIT Press.

Williams, B (1981), *Moral Luck,* Cambridge: Cambridge University Press.