I

MORAL AUTONOMY

1. Autonomy and Freedom

The contemporary philosophical debate on autonomy presents several conceptual distinctions; one of these is the fundamental distinction between "moral autonomy" and "personal autonomy"[1]. According to Kant, an agent is autonomous in a moral sense when her own rules for acting are included in her will as universal law (*Groundwork of the Metaphysics of Morals* section 2 IV). Modern proponents of personal autonomy introduce rather the image of a person in charge of her life, not just following his desires but choosing which of her desires to follow[2] .

In this chapter I shall deal with the concept of moral autonomy. The first matter is the distinction between autonomy and freedom. Isaiah Berlin describes autonomy as "positive freedom". Positive freedom: «(…) derives from the wish to be self-directed and not acted upon by external nature or by other men as if I were a thing, or an animal, or a slave incapable of playing a human role, that is, of conceiving goals and policies of my own and realizing them»[3]. Indeed, Berlin also presents a kind of "negative freedom" because he thinks that a liberal government must protect citizens from external constraints. Thomas Green underlines rather the importance of positive freedom as autonomy in a Kantian sense: human will is a form of a principle that realizes itself consciously[4].

[1] See D. Johnston, *The Idea of Liberal Theory*, New York, Princeton University Press, 1994; J. Waldron, *Moral and Personal Autonomy* in J. Christman and J. Anderson (ed.), *Autonomy and the Challenges to Liberalism*, New York, Cambridge University Press, 2005.

[2] The hierarchical model is introduced by R. Dworkin. See his *The Theory and Practice of Autonomy*, New York, Cambridge University Press, 1988. See also H. Frankfurt, *The Importance of What We Care About*, Cambridge, Cambridge University Press, 1987.

[3] I. Berlin, *Two Concepts of Liberty*, in his *Four Essays on Liberty*, Oxford, Oxford University Press, 1969, p. 131.

[4] T. H. Green, *On the Different Senses of Freedom*, in T. H. Green, *Lectures on the Principles of Political Obligation*, London, 1941.

Raffaela Giovagnoli, *Autonomy: a matter of content*, ISBN 978-88-8453- 636-5 (online), ISBN 978-88-8453- 635-8 (print), © 2007 Firenze University Press

According to Kant, freedom and autonomy are two different concepts. In the practical field freedom corresponds to "practical spontaneity", which is different from "epistemic spontaneity" in its revealing of the laws of nature. According to Kant: «Reason must look upon itself as the author of its own principles independently of alien influences. Therefore, as practical reason or as the will of a rational being, it must be regarded by itself as free; that is, the will of a rational being can be a will of his own only under the idea of freedom, and such a will must therefore – from a practical point of view – be attributed to all rational beings» (*Groundwork* III, section 4). Starting from this perspective, the conditional result of spontaneity is evident: if I take myself to be a rational agent i.e. I take myself to be acting on the basis of reasons and of a reflective evaluation of my situation rather than merely responding to stimuli, I must necessarily regard myself as free[5].

It is possible for human beings to be guided by heteronomous principles (i.e. dependent on the relationships between an agent's will and the properties of objects). Autonomy requires therefore a further condition: the capacity of the will to determine itself independently of every property belonging to the objects of volition. Moral autonomy implies not merely that our actions conform to duty but that they derive "from duty": the duty itself provides a sufficient reason to act. Nevertheless, freedom and moral law imply each other. The "reciprocity thesis" can be interpreted by starting from a conception of negative freedom as including motivational as well as causal independence. Freedom corresponds to spontaneity, i.e. to the rationality of the agent: «This entails that its choices must be subject to a justification requirement. In other words, it must be possible for such an agent to offer reasons for its actions; since reasons are by their very nature universal, this means that such an agent must be willing to acknowledge that it would be reasonable (justifiable) for any rational being to act in a similar manner in relevantly similar circumstances»[6].

Moving from this perspective, to be autonomous is to be intentionally bound by conceptual rules that are not individual desires and preferences. I maintain that an analysis of the concept of freedom requires investigations into individual motivations such as desires and preferences. In this sense, we are free to act according to our means-end reasoning without the necessity of assuming a detached, responsible perspective over personal desires and preferences. Naturally, we are free to act when external or internal constraints do not exist. Autonomy can be considered rather as that capacity human beings normally have of universalizing their own point of view and thus distinguishing subjective and objective reasons.

[5] See H. E. Allison, *Idealism and Freedom*, Cambridge University Press, Cambridge, 1996, p. 133.

[6] Ivi, p. 137.

Autonomy comes into play also when an agent has conflicting subjective reasons for acting. For this reason, motivation alone cannot be the source of autonomy and we need to consider the rationality of choice. This rationality is expressed by the fact that we can justify our choices. For example, we can explain why we prefer to drink wine rather than beer or why we recognize that an insult will offend our interlocutors.

The kind of rationality we are looking for entails dimensions that supersede merely means-end reasoning. For the sake of my analysis it is useful to consider the reasons Habermas provides for superseding contemporary interpretations of classical empiricism[7] in order to present a "communicative rationality" (i.e. that rationality oriented toward a consensus among speakers)[8]. Empiricism understands practical reason as instrumental reason. According to the agent, it is reasonable to act in a certain way when the outcome of her action corresponds to her interest, or is satisfying. In this case, action is motivated by preferences and personal ends. They are "pragmatic " or " preferential" reasons, as they deserve to motivate actions and, contrary to "epistemic" reasons, not to justify judgments and opinions. Pragmatic reasons "modify" free will only when the agent decides to follow a certain rule. In this sense, intentional motivation distinguishes itself from spontaneous motivation. According to Hume, autonomy corresponds to the undertaking of attitudes of approval or disapproval. These feelings belong to the third person perspective of "benevolent detachment" from which actors are morally judged. Every agreement on the moral valuing of a character will imply a coincidence of feelings. Approval and disapproval express likes and dislikes; thus they have an emotional nature. It is true that we all react with disapproval when someone performs a bad action. We consider a person as virtuous when he reveals himself to be useful and agreeable to us and to our friends. There exist pragmatic reasons for an agent who wants to adopt "altruistic" attitudes. Indeed, the benevolence of the interlocutors provides satisfaction to a useful and agreeable person. By starting from this perspective based on emotional attitudes the social force of reciprocal trust can develop.

Nevertheless, pragmatic reasons supporting moral attitudes and actions are convincing only with regard to interpersonal relationships of small sympathetic communities (for example family or neighbourhood). In complex societies interpersonal relationships require a moral point of view, i.e. one of autonomy, which aims at universal justice, since they become more abstract. In this sense, the members of a primary group could not immediately refer to the benevolence of persons living in a different cultural context. When feelings of obligation pertain to relationships be-

[7] J. Habermas, *The Inclusion of Other, Studies in Political Theory*, The Mit Press, 1998, chap. 1.

[8] I shall describe Habermas' perspective in chapter 6, where I discuss an intersubjective concept of autonomy in the political field.

tween strangers, the agent cannot consider them "rational" as he considers those of his group, on whose cooperation he can always rely. A theory of autonomy aiming at social justice must explain the normative primacy of the rational "ought" in a wider context.

Contractualistic tradition does not consider solidarity, rather it directly relates the normative ground of a system of justice to individual interests (this move implies the shift from "duties" to "rights"). This strategy aims at understanding whether it is rational to subordinate the individual will to a system of rules (i.e. subjective rights). In virtue of the figure of the contract, it seems possible to ground subjective rights in a symmetric way and so to construct a legal system based on a free agreement. This system is right in a moral sense if it satisfies the interests of its members equally. The social contract derives from the idea that every candidate must necessarily have rational motives for becoming a member of the collectivity and to subordinate him/herself freely to its norms and procedures. The cognitive content that turns the contract into a "moral" or "right" system is bound therefore to the collective acknowledgment of all members. Moreover, it refers to the instrumental rationality of each member who values his advantage from the perspective of his own interests.

Contractualism is subject to two fundamental criticisms. First, from this perspective a universal morality becomes impossible, because it neglects the problem of the normativity of interpersonal relationships. A person can undertake a commitment only if she can expect the right response from her interlocutors in a cooperative situation. Second, it raises the "free rider" problem that shows how a person could commit herself to a cooperative praxis while being free to break the norms when a good opportunity arises.

2. The Role of Reason

For these reasons it is useful to reconsider the Kantian normative role of reason. In the *Critique of Pure Reason* the concept of reason has a "regulative function": it imposes on one's mind the infinite search of the unconditioned. Epistemological doubt forces us not to stop the search even when we think we have reached the truth. In its practical function, reason shows an opposite requirement: it must determine by itself as pure reason its own will, because only in this way is it "moral" i.e. unconstrained by nature or inclinations. In the theoretical field, reason regulates mind, which must apply to empirical dates. In the practical field it possesses a "constitutive" function: its tendency to be subordinated to subjective inclinations is criticized[9].

[9] See W. Mathieu, Introduction to Kant *Fondazione della metafisica dei costumi*, Milan, Rusconi, 1994, pp. 11-2.

The first step for demonstrating the efficacy of reason is the argumentation about the shift from morality of common sense to philosophical knowledge (*Groundwork* part I). A representation of the "ought" is rational if it is independent of sensitivity, as the moral value is peculiar to a being determining his/her own will by him/herself i.e. regardless of immediate impulses or physical necessity. How could reason be effective without empirical motivation? Undoubtedly, we cannot observe as a fact how the mere representation of a law determines the agent's behavior. On the contrary, the psychologist analyzes the relations between an agent's behavior and empirical circumstances of it, coming to find certain "regularities". He could call them "laws", but they are effective because of the empirical circumstances determining human agency. From this perspective, the agent is passive, like a body falling as a consequence of gravity. In the case of morality, the agent presupposes rather that she is the one who (rightly or wrongly) decides. This decision takes place once she is free from all empirical influences and determines her will only because of his obligation (it does not matter if he likes to or not). If the empirical circumstances are deprived of motivational force, then she can be compelled only by the pure form of the law. The law is therefore "rational" as experience is ruled out.

Rationality, in this context, is bound only to pure obligation and must be distinguished from instrumental rationality. One can act out of self-interest; for example, behaving honestly, because doing otherwise implies bad consequences, is rational and contrasts immediate egoistic impulses. But this is a case of instrumental rationality aiming for the goal of its own happiness: reason is not the source of will; it is indeed a practical reason but not a "pure" practical reason. Thus the work of pure practical reason is to rule out reasons for justifying action that limit reason to the service of sensibility.

The categorical imperative, being constitutive of volition, is therefore a peculiar principle of practical reason. It tells us to act only on maxims we can will as universal laws. What is the status of this principle? According to the Kantian discussion on the nature of practical reason, the most immediate conclusion is that it must be "formal", as it must show the possibility of acting out of self-interest. How does the principle of morality work to unify human will? Let us consider Korsgaard's arguments against the empiricist and rationalistic traditions[10]. Starting from Hume's point of view, three correlated arguments are important: (1) the role of reasoning is to ascertain the relations between things; (2) the only relation directly relevant to action is the causal relation and (3) that relation can have motivational force only if we have a desire to obtain or avoid one of the two objects thus related. In this case, the success of the action depends on our belief about reaching a

[10] See C. Korsgaard, *The Sources of Normativity*, Cambridge, Cambridge University Press, 1996, lectures 1 and 2.

certain end (in case of error it is a matter of theoretical considerations). This option hinges on the fact that we are always reliable in reference to our desires and actions. This "optimistic" view of human rationality could imply a tautology: «The problem arises when we ask what makes something someone's end. Suppose someone claims to desire a certain object. We inform him that taking a certain action is the adequate and sufficient means to the achievement of that object, yet he fails to form the desire to do that action. Then we are entitled to conclude that he does not desire the object, or does not desire it enough to inspire him to take those means. That being so, the object is not his end, and that being so, he has not failed to act on any instrumental reason that he has. If we mean by your "end" is that which you in fact pursue, it is conceptually impossible for you to fail to take the means to your end. If you fail to pursue something, then it isn't your end, and then you don't act irrationally in failing to pursue it. But then the force of saying you acted on an instrumental "reason" becomes unclear. Your desire for the end plays a role in explaining why you took the action, but there is no requirement of taking the means to your end that has any normative force for you, and so no reason on which you acted»[11].

The rationalist view (Samuel Clarke, John Balguy, and Richard Price in the 18th century, William Whewell in the 19th century, and W.D. Ross, H. A. Prichard, and Derek Parfit in the 20th and 21st centuries) supersedes the problem of Hume, who considers action as nothing more than a movement caused by a judgment or idea. In a rationalistic sense, action is not merely caused by a judgment, but rather guided by it. According to this "realistic" or "externalist" view, action refers to true substantive moral principles for its justification, i.e. certain act-types or action are inherently right or wrong. Another externalist view maintains that irreducibly normative reasons for actions simply exist. We act according to normative verities because we apply our knowledge that an action is right by choosing it. Korsgaard rightly observes that simple choice guided by knowledge of the right reason does not explain "moral" obligation. In my opinion, we need to make clear the normativity of the principles of rationality by reference to two different cases: (1) the case in which we act for subjective reasons and (2) the case in which we act for objective reasons.

3. Action and Normativity

The conclusion of the argumentation sketched above leads to the thesis that if the principles of practical reason are to be normative, there

[11] C. Korsgaard, *Practical Reason and the Unity of the Will (lecture II),* forthcoming, pp. 3-4.

must be principles of the logic of practical deliberation. In the *Groundwork* a fundamental distinction between hypothetical and categorical imperatives is introduced. These represent the normative source of personal identity, as constitutive principles of actions. According to Kant, action is determining oneself to be the cause of some end. In the case of the hypothetical imperative, the normativity of action implies that the commitment to realize an end binds the agent, obligating her to take the necessary means. In this sense: «(…) this is a commitment that you may fail to meet. Finding the means daunting, frightening, tedious, or painful, you cannot face them and do not go forward. Finding yourself nevertheless unprepared to decide that the end is not worth it, you cannot not change your mind and you cannot not go back. A paralyzed will is not the same thing as one that has simply failed to operate; an abortive effort at self-determination has taken place. The standard represented by hypothetical imperative, though constitutive, is normative as well»[12].

The categorical imperative is normative in an unconditional sense. For example, if I say "don't lie" I presuppose that lying is unconditionally wrong and this is not the same as saying "don't lie if you want to preserve a friendship". How can we demonstrate the absence of subjective conditions of the categorical imperative? In this context, the action is put to the test of the "universalization of the maxim". A maxim is a subjective principle of action, and the problem is to know whether the chosen maxim corresponds to the law, i.e., the universal principle. A corresponding version of this thesis is the following: act in accordance with the maxim you can will as a universal law. Naturally, it does not imply any "particularistic" will, which means that an agent acts according to momentary impulses. The unity of will is due to the categorical imperative. This is a constitutive standard of action because conformity to it is constitutive of an exercise of the will of the self-determination of a person. A significant interpretation of the universalization test is the one presented by John Rawls[13]. By using the metaphor of the "veil of ignorance" an autonomous moral legislator must, in Kantian terms, abstract from personal differences. The central point for both Kant and Rawls is that: «(…) for purposes of trying to adjudicate fairly and reasonably among *competing* principles and values, certain considerations must be ruled out of the court. For example, the fact that a principle would benefit *me, my* family, and *my* country instead of someone else, someone else's family and country, is not *in itself* a reason for anyone, as a moral legislator, to favor that principle. In other words, at the level

[12] C. Korsgaard, *Autonomy, Efficacy, and Agency (Lecture III)*, on-line, p. 1.

[13] J. Rawls, *A Theory of Justice*, Cambridge, Harvard University Press.

of deliberation about basic principles, morality requires impartial regard for all persons»[14] .

This interpretation of the categorical imperative focuses on the dimension of choice that is fundamental for the contemporary Kantian philosophers. It entails the difference between "act" and "action". An act is performed for a certain end and for this reason Kant called it "heteronomous action". On the contrary, "acts" are subject to individual choice, i.e. are "objects" of individual choice and become therefore actions as they are guided by the categorical imperative. Another way to demonstrate the role of the categorical imperative for human action is to examine the difference between human and animal choices[15]. Indeed all animal choices are subject to external causes. The animal acts from natural laws, i.e. nature provides laws of causality of animal action. What is peculiar about human choice? Korsgaard refers to Aristotle's classification of actions in order to distinguish between merely voluntary actions and choices. The former are peculiar to animals that act according to natural laws. Animals represent the objects of their desires because they have a form of intelligent adaptation in perceiving their environment. Human beings have the possibility of moral choice, i. e. of classifying actions as morally good and bad.

But in order for us to be considered as autonomous we must distinguish technical knowledge from moral knowledge. Aristotle discusses this topic in his *Nicomachean Ethics* where he maintains that in the ethical field we cannot have prior knowledge of the means to reach certain ends[16]. Moreover there are not particular ends but moral knowledge determines all the goodness of moral life. For this reason, moral knowledge is bound to individual deliberation and reflection. Because of the fact that we do not previously know the means for reaching determinate ends, moral knowledge is internally related to ethical consciousness, which must refer to the concrete situation. The most important consequence of this thought is that moral consciousness directly reflects on the means for determining the moral validity of ends. An individual's reflection on his ends is *eo ipso* a moral commitment. It is not bound to sensible perception but rather orients itself to the actual situation, i.e. it is consciousness of the ends and the means necessary for it. What is right is the result of a reflection that is not constrained by passions causing the loss of control. Moral knowledge is that which includes our consciousness of means and ends, and is therefore opposed to a pure instrumental knowledge.

[14] T. E. Hill, *Autonomy and Self-respect*, New York, Cambridge University Press, 1991, p. 45.

[15] See Korsgaard, *Autonomy, Efficacy, and Agency*, op. cit.

[16] H. G. Gadamer clarifies this distinction in *The Problem of Historical Consciousness* in P. Rabinow & W.M. Sullivan (eds.), *Interpretive Social Science. A Second Look*, California University Press, Berkeley, 1987.

Moral knowledge pertains not only to individual moral action but also to interpersonal relationships. In this case we need not only the concept of *Phronesis* but also that of *Synesis*. The latter refers to the phenomenon of the comprehension among persons. *Synesis* is an intentional modification of moral knowledge whenever we need to value the agency of our interlocutors in an ethical sense. In this case, moral evaluation means taking the place of the other, placing oneself in the situation in which the other must act. This mutual relationship is based on a kind of "affinity" which allows the discernment of the other's situation and the tolerance resulting from that knowledge. The importance of these thoughts for the Kantian tradition is the possibility of undertaking a reflective attitude over personal impulses and desires in order to consider others' peculiar situation. This fact allows us to supersede merely instrumental attitudes so that we can consider persons as ends in themselves.

4. The Deep Deliberator

As we have already seen in the last paragraph, animals are not guided by hypothetical and categorical imperatives for a fundamental reason: animals do not choose the principles of their own causality, they do not choose the "content" of their instincts. Moving from this thought, we can recognize two senses of autonomy or self-determination: «In one sense, to be autonomous or self-determined is to be governed by the principles of your own causality, principles that are definitive of your will. In another, deeper, sense to be autonomous or self-determined is to choose the principles that are definitive of your will. This is the kind of determination that Kant called "spontaneity". Every agent, even an animal agent, is autonomous and self-determined in the first sense, or it would make no sense to attribute its movement to it. Only responsible agents, human agents, are autonomous in the second and deeper sense»[17].

The figure of the "deep deliberator" presented by Korsgaard aims at demonstrating that autonomy needs a "normative" concept of reason relating to the "reflective self". I think that Korsgaard's criticism of Thomas Nagel's realism could elucidate this point[18]. According to Nagel: «Why isn't the reflective individual just someone with more information, who can therefore make choices which may be different but which need be no less purely personal – or even temporally fragmented? How do reasons, law, and universality get a foothold here – one that cannot be dis-

[17] Korsgaard, *Autonomy, Efficacy and Agency*, op. cit. p. 9.

[18] See C. Korsgaard, *The Sources of Normativity*, Cambridge University Press, Cambridge, pp.217-233.

lodged? Presumably it has something to do with the difference between reflective and unreflective consciousness, but why should awareness of self bring with it this further regularizing effect?».[19] Kant would respond that this happens because the will is a causality and a causality must operate in accordance with laws. But two further points show Nagel's refusal of this thesis. (1) If will is self-determining it could be the case of disconnected choices as well as those guided by consistent laws. In this sense causality does not allow the distinction between merely following a rule and applying the categorical imperative. (2) According to (1), the reflective self must be more universal that the unreflective self (i.e. the self guided by desires and impulses), because it achieves its self-conscious awareness by detaching itself from the individual perspective. Korsgaard underlines the necessity of assuming a detached perspective but he recognizes also the importance of determining "how" the achievement of reflective distance leads the agent to identify himself as a "person", i.e. to have a normative conception of himself as a person.

This option represents a problem in Nagel's account because it is important to establish why reasons for acting could be general. A relevant fact is that the person possesses a kind of normativity for judging whether some reasons are more objective than others. The agent's causality is normative as he directly perceives power; consequently individual exercises of power become perceivable. In this sense: «(…) willing is *self-conscious* causality, causality that operates in the light of reflection. To will is not just to be a cause, or even to allow an impulse in me to operate as a cause, but so to speak, to consciously pick up the reins, and make *myself* the cause of what I do. And if I am to constitute *myself* as the cause of an action, then I must be able to distinguish between *my* causing the action and some desire or impulse that is "in me" causing my body to act. I must be able to see *myself* as something that is distinct from any of my particular, first order impulses and motives, as the reflective standpoint in any case requires. Minimally, then, I am not the mere location of a causally effective desire but rather am the *agent* who acts *on* the desire. It is because of this that I endorse acting in a certain way now; I must at the same time endorse acting the same way on every relevantly similar occasion».[20]

Normative principles of the will possess the function of bringing integrity and unity to human actions. The reflective self is the self which is capable of achieving a reflective distance from immediate impulses. The reflective attitude unifies the self "not" because it has some reason to want or anticipate that it will persist into the *future*.

[19] T. Nagel, *Universality and the Reflective Self*, in Korsgaard, *The Sources of Normativity*, p. 201.

[20] Korsgaard, op. cit. p. 227-228.

The interpretation of Kantian principles of rational will is discussed by Hill who criticized Nagel's theory of reasons[21]. The point of contention is the role of the rational will for the choice to do something implying projects and pains. The problem of metaphysical determinism is irrelevant here because it is not a case of predicting a future action. Also the deliberation from the instrumental point of view, i.e. the consideration of means, risks, costs, sacrifice, etc., has nothing to do with deliberative reflection on previous ends and commitments. Deep deliberation presents two fundamental features: (1) the searching for "justifying reasons" and not "motivating reasons" and (2) the agent's performing of actions from choices that reveals himself as agent.

Let us now ask whether pleasure and pain provide the necessary common denominator of rational choice. Extreme hedonism identifies rational choice with choice that maximizes the agent's balance of pleasure over pain, and so counts nothing but pleasure and pain as ultimate reasons. This perspective comes to the conclusion that the foregoing of innocent pleasures for the sake of highly desired and valued states of affairs after one's death is always irrational. A more modest claim maintains that pleasure and pain are not the only ultimate reasons but they are "always" ultimate reasons. The problem here is that if in a certain moment of his life an agent takes a perverse pleasure in performing a bad action it is not a question of balancing pro and contra: the very nature of pleasure makes it inadequate to count as justifying reason. Also Nagel's consideration of severe physical pain as an ultimate reason for acting needs to refer to the procedural condition of rational deliberation: claiming to discount severe pain is a good sign of a disorder in one's thinking about practical matters. The agent's inclinations are therefore not viewed as forces which fix one's ends without one's cooperation: «Naturally, one assumes that one *will* not choose to discount inclinations without a reason, and acknowledging something as a reason implies caring about it. But searching for reasons is not simply trying to discover one's inclinations, just as weighing reasons is not simply trying to introspect the relative strengths of one's inclinations. One may find that one "cannot" typically express a refusal, not a disability. Like Martin Luther's remark, "here I stand, I can do no other", it does not complain of powerlessness but rather expresses sustained commitment»[22].

Theories based on prudence represent a modification of the Kantian view as they give regard for one's future. In this context, immediate pleasures, desires, etc. necessarily give the agent reasons to act; the agent is the same person over time so if an experience is of a kind to make his reason

[21] See Hill op. cit., chap. 12.

[22] Ivi, p. 183.

favor it for himself now, the same features must make his reason, other things being equal, favor it for himself later. Also this perspective moves from the presupposition that certain ends in themselves necessarily give the agent reasons to act. The fact that an agent is the same rational agent over time does not need to be construed as referring to a set of attitudes typically presupposed in deliberation: «For example, I'm now responsible for and to myself later, and I later will be responsible for my choices now as well as then, and responsible to myself still later»[23]. The notion of "responsibility" could however be interpreted by considering both the causal and the normative dimensions. In this sense, when the agent takes a deliberative standpoint she has not only the ability to influence her future choices indirectly but also a capacity to make plans and resolutions for his future relevant for later deliberations. An agent responsible for present choices will therefore acknowledge that she is the author of the character and consequences resulting from her current choices[24].

In the consideration of individual projects, the Kantian idea of respect for oneself takes the place of the necessity of certain experiences as reason-giving. This move is due to the thesis that deep deliberators are concerned to choose so that their choices stand up, at least at that time, to critical scrutiny of and by themselves. Values and inclinations may vary with time, but these are what the agent subjects to scrutiny and not a fixed presupposition of it.

5. The Reason View

In this last section I shall present arguments that moderate the Kantian internalist conceptions presented above. Individual self-reflection needs to confront itself with objective reasons, i.e. reasons that are external to the pure exercise of practical reasoning. This is a very important point because the consideration of the content of our beliefs and actions seems fundamental in order for a performance to be autonomous.

The starting point of my discussion is Susan Wolf's observation that the relationship between autonomy and responsibility implies the possibility of acting in "discordance" with reason. It is indeed difficult to see how an agent is autonomous only if she acts always in accordance with reason. In this case, the work is done by the notion of responsibility with the consequence that we do not need the notion of autonomy. It can also be pointed out that: «(…) if one lacks the ability to act in accordance

[23] Ivi, p. 185.

[24] On this subject see also F. Duque, *Liberdad y sacrificio: Deber ser para dejar ser*, «Revista portuguesa de Filosofia», 61, 2005, 667-686.

with Reason, one cannot be responsible even if one is autonomous. For dogs and psychopaths might conceivably be autonomous in the sense that they might be ultimate sources of their own actions, able to act on no basis. But because they lack the ability to act on a basis – in particular, the basis of Reason – they are not responsible in the sense that would allow them to be deserving of deep praise and blame»[25].

What Wolf calls the "Autonomy View" is the view of metaphysical free will. The problem according to this perspective is that the ability necessary for responsibility is "bidirectional" – it is an ability to do one thing *or* another, an ability to do X or something other that X. On the contrary, according to Wolf's "Reason View" the ability necessary for responsibility is unidirectional – it is an ability to do one sort of thing, which is compatible with the *in*ability to do anything else. This fundamental difference entails an agreeable clarification of the flexibility of the agent's identity. The Autonomy View maintains that a responsible agent is flexible insofar as she is able to choose and act in a way that is not forced upon him by uncontrollable features or events of her past. It is indeed the Reason View that can explain flexibility: an agent partly acts in accordance with Reason if she is sensitive and responsive to relevant changes in her situation and environment. Acting according to Reason means having the peculiar ability to choose and act for the "right" reasons. The source of this "normativity" requires a distinction between human beings and non-human beings. According to the Reason View we have the intellectual power to recognize the True and the Good, and it has nothing to do with the metaphysical power to choose and act out one path of action or another independently of any forces that could represent potential constraints. Autonomy, i.e. acting according to the True and the Good, implies two kinds of explanations that are related. In the case of action, we can first point out that in the process of socialization the agent was taught to act justly, and was subsequently positively reinforced for doing so. Second, we can point out that it is right to act justly, and go on to say why she knows this is so. These explanations are likely to be related: if it were not right to act justly, he might well not have been taught that it was. And if the person had not been taught that she ought to act justly, the person might not have discovered this on her own. These explanations are therefore compatible: one can be determined by the Good *and* determined by the Past.

Moreover, acting according to the True and the Good entails a wider notion of "responsibility" which implies not only the dimension of the moral point of view. In this sense, subjective reasons for doing something, such as drinking coffee, exhibiting a fondness for purple or spending so

[25] S. Wolf, *Freedom within Reason,* New York, Oxford University Press, 1990, p. 68.

much time philosophizing, come into play. :«According to the Reason View, it is up to me if my decisions to do these things are made in the light of my knowledge or of my access to knowledge of the (true and good) reasons for doing and not doing them (assuming as well that my doing these things is dependent on my decisions to them). For example, I am responsible for drinking coffee if in deciding whether to drink it, I am in a position to know, appreciate, and act on the reasons for and against drinking it. If, on the other hand, I am not in such position – if perhaps I am hypnotized to drink coffee, or deceived about what is in my cup – then I am not responsible from drinking it»[26].

The important result of Wolf's thesis is that the freedom necessary for responsibility is not just the freedom that allows one's action to be governed by one's own reason, but also a freedom that allows one's reasons to be governed by what reasons there are. This theoretical option could entail the presupposition that there are objective values in a Platonic sense that we have the ability to recognize. But the assumption of what Wolf calls "normative pluralism" shows a different way of interpreting the objectivity of "normative facts": seeing the world rightly involves seeing reasons for and against many different options. In this sense, maximum freedom and responsibility would presumably imply being able to see all of them. It is indeed difficult to have capacities, knowledge and time to engage in such an enterprise. The explanation of this capacity is easy to understand if we think that the case of the agent is moved by praise or blame. But, according to Wolf, from this perspective it is hard to isolate the special ability to appreciate reasons and values, because it seems necessarily bound to intellectual moral reflection and introspection. Rather we must consider the plurality of values belonging to different cultures: these values as normative facts are the content of our practical reasoning: «If inevitable features of myself – my gender, my race, my nationality, for example – and rationally arbitrary choices and twists of fate shaped my values and decisions, this does not seem to me to place objectionable limits on my status as a free and responsible agent. As long as these non-rational determinants do not prevent me from a sufficiently open-minded and clear-headed assessment of my values to allow me to see whether they fall into the range of the reasonable, and as long as my blindness to some other reasonable alternatives does not lead me to acts of intolerance or prejudice, then it seems that, for most intents and purposes, I am free and responsible enough. These non-rational determinants are, after all, what give us our individuality and distinctiveness. If, at the limits, they can be in tension with our freedom and responsibility, in more central cases they provide the basis for substantive identity and

[26] Ivi, p. 91.

an attachment to the world without which no interest in freedom and responsibility could arise»[27].

Despite this interesting conclusion, the task of a theory of autonomy is, in my opinion, to give a plausible account of the way in which a person can criticize the non-rational determinants of his identity if it is the case. This requirement is fulfilled if we can describe the normative (rational) structure of that point of view that allows the agent to be autonomous, i.e. able to discern and justify a wide range of (subjective and objective) commitments.

[27] Ivi, p. 146.