Borges, Maria. "Action, reason, and causes in Kant." *Emotion, Reason, and Action in Kant*. London: Bloomsbury Academic, 2019. 5–34. *Bloomsbury Collections*. Web. 26 Apr. 2023. <<http://dx.doi.org/10.5040/9781350078390.0008>>.

Downloaded from Bloomsbury Collections, www.bloomsburycollections.com, 26 April 2023, 15:10 UTC.

Copyright © Maria Borges 2019. Released under a CC BY-NC-ND licence (<https://creativecommons.org/licenses/by-nc-nd/3.0/>). You may share this work for non-commercial purposes only, provided you give attribution to the copyright holder and the publisher, and provide a link to the Creative Commons licence.

1

Action, reason, and causes in Kant

In this chapter, I analyze Kant's theory of action and what can count as a reason for or cause of an action. I begin with the distinction between *motiva* and *stimulus* in the pre-critical lessons on ethics and metaphysics. In the *Groundwork*, Kant explains this distinction in terms of objective and subjective grounds for actions: The motive (*Bewegungsgrund*) is the objective ground of an action and the incentive (*Triebfeder*) is the subjective one. Next, I will analyze the possibility of the overdetermination of actions, understood as the possibility of two possible causes for the same action. I argue that according to the incorporation thesis, there is only one motive for an action.

Second, I try to reconcile the weakness of the will and the incorporation thesis. I will show that, according to the incorporation thesis, Kant is committed to a strong thesis concerning the causation of actions. Reason and only reason can be the cause of an action. If we accept weakness, however, we must have a humbler solution to the Kantian theory of action: the domain of rational agency does not have the same extension of voluntary action. In the model of rational agency there is no room for weakness, although it is a fact in real actions. The domain of the voluntary is, then, wider than the domain of rational agency. The weakness of the will applies to the first and the incorporation thesis applies only to the second. In this chapter I also

investigate if an action can be predicted, based on information about the inner states of the agent.

Stimulus and motiva in the precritical Kantian ethics

Is it a reason to act the cause of an action? Or is it only a way to explain it *post factum*? What is the cause of an action? There are two ways of answering this question: one lies in why we perform that action, another in what really pulls us to perform it. Two different concepts apply: a motive is the intellectual reason for doing something; an incentive is what drives one to do it. In the *Groundwork*, Kant explains this distinction in terms of objective and subjective grounds for actions: “The subjective ground of desire is an incentive; the objective ground of volition is a motive” (G, 4: 428). The motive (*Bewegungsgrund*) is the objective ground of an action and the incentive (*Triebfeder*) is the subjective one.

The distinction *Triebfeder/Bewegungsgrund* can also be found in the early Kant lessons on ethics and metaphysics, as a distinction between the Latin words *stimulus* and *motiva*. In the Collin’s *Lectures on Ethics*, one reads

Necessitation (*Nötigung*) is of two kinds, objective and subjective. Subjective necessitation is the idea of the necessity of actions *per stimulus*; or through the *causae impulsivae* of the subject. Objective compulsion (*Zwang*) is the constraining of a person through what has the greatest constraining and moving power in his subject. (LE, 27: 267)

Compulsion can be either pathological or practical, the first is the necessitation of an action *per stimulus*, and the second is the necessitation *per motiva*. Human choice cannot be necessitated *per stimulus*, since it is an *arbitrium liberum*. Animals are necessitated *per stimulus*: “So that a dog must eat if he

is hungry and has something in front of him; but man, in the same situation, can restrain himself” (LE, 27: 267). To claim that human choice is an *arbitrium liberum* is to accept that human beings can only be compelled *per motiva*, not *per stimulus*.

A similar picture is given in the *Lectures of Metaphysics*: “Every act of choice has an impelling cause <*causam impulsivam*>. The impelling causes <*causae impulsivae*> are either sensitive or intellectual. The sensitive are stimuli <*stimuli*> or motive causes, impulses. The intellectual are motives or motive grounds” (*Metaphysik* L1, 28: 254).

The sensitive impelling causes may have a necessitating power (*vim necessitante*) or impelling power (*vim impellentem*). For nonrational animals, the stimuli have a necessitating power, but not in human beings, for whom they have an impelling power.

Both in the *Lectures of Metaphysics* and *Lectures of Ethics* we can see a double level of causation, in the distinction between sensitive and intellectual impelling causes. In human beings, the sensitive impelling causes (*stimuli*) can only have an impelling power, not a necessitating one:

Stimuli (*stimuli*) thus have either necessitating power (*vim necessitantes*) or impelling power (*vim impellentem*). With all non-rational animals, the stimuli (*stimuli*) have necessitating power (*vim necessitantem*) but with human beings the stimuli do not have necessitating power (*vim necessitantem*) but rather only impelling (*impellentem*). (*Metaphysik* L1, 28: 255)

Human choice is called free choice, because it is independent from the necessitation of all stimuli. This freedom of choice is connected to practical freedom:

This practical freedom rests on independence of choice from necessitation by stimuli (*independencia arbitrii a necessitatione per stimulus*). That freedom, however, which is wholly independent of all stimuli (*stimulis*), is

transcendental freedom, which will be spoken of in the Rational Psychology (psychologia rationalis). (*Metaphysik* L1, 28: 257)

Practical freedom is the independence of choice from the necessitation of *stimuli*, while transcendental freedom is independence from all *stimuli*. In the *Lectures on Metaphysics* Kant anticipates what he will claim in the *Critique of Pure Reason*. In the latter, transcendental freedom is defined as a faculty of beginning a state in itself and as “spontaneity, which could start to act from itself, without needing to be preceded by any other cause” (KrV, A 533/B 561) and freedom in the practical sense as “the independence of the power of choice from *necessitation* by impulses of sensibility” (KrV, A 534/B 562). While in the *Dialectic*, the concept of practical freedom is a case of transcendental freedom, in the *Canon* of the *Critique of Pure Reason* Kant argues that one may have practical freedom without transcendental freedom.

Henry Allison tries to solve this incompatibility by showing that “practical freedom is related to transcendental freedom the same way divine freedom is related to human freedom.”¹

If human beings have practical freedom without having transcendental freedom, then their actions are not completely independent of sensory impulses, although they should be independent of determination by sensory impulses. This point will come later, when we discuss the possibility of acting without feelings.

Incentives, motives, and the overdetermination of maxims

The distinction between *Triebfeder* and *Bewegungsgrund* is crucial in contrasting actions according to duty with actions done from duty, because it underscores the distinction between what constitutes moral motive and incentive. The difference between moral and nonmoral incentives is explored

in the philanthropist example, where we find two agents with different incentives for being benevolent. Neither of them has a motive of vanity or self-interest; nevertheless, the first has a natural inclination to do good for other persons, an inner happiness in making other people happier. Although his action has a moral motive, Kant maintains that “in such a case an action of this kind, however right and however amiable it may be, has still no moral worth” (G, 4: 398). Hence, a moral motive is a necessary, but not sufficient, condition for a moral action.

When does an action have moral worth? Kant answers with the case of the second philanthropist (G, 4: 398). Unlike the sympathetic philanthropist, the insensible one performs an action with moral worth; consequently, the absence of sympathy seems to make an action morally worthy. In order to avoid the criticism of insensibility, some authors have discussed what makes an action morally worthy. They were obviously trying to avoid the uncomfortable conclusion that moral insensibility is something good.

One of these authors is Henson. Henson² tries to answer two questions related to this example of the philanthropist: (1) What does it mean to ascribe moral worth to an act? (2) Under what circumstances are we to say that one acts from duty?

There are two possible answers in Kant's work: (A) an action has no moral worth if, at the time of performance, the agent has an inclination to perform it; (B) the moral action does not require an absence of inclination, provided that respect for duty is present and would suffice to produce the dutiful action. The A answer gives us what Henson calls the *battle citation model*; the B answer provides us with the *fitness report model*. According to the *battle citation model*, an action has moral worth only if respect for duty were the only motive tending to the direction of the dutiful act. In the *fitness report model*, other inclinations could be present, provided that respect for duty was present and would have sufficed by itself, even though other motives were also present and might themselves have sufficed. According to the fitness model, there is no

need to banish all other inclinations toward an action, provided that respect would have been a sufficient reason to cause the action.

Barbara Herman in her book *The Practice of Moral Judgment*³ discusses whether the absence of inclinations is a necessary condition of a moral action: “The apparent consequence of this view . . . is at the least, troubling in that it judges a grudging or resentfully performed dutiful act morally preferable to a similar act done from affection or with pleasure.”⁴ She argues that the presence of a nonmoral motive does not mean a lack of moral worth in an action. However, she points out some problems with the idea of sufficiency employed by Henson in the fitness model. She goes further and proposes to strengthen the interpretation of the fitness model.

Suppose that a shopkeeper had two sufficient motives to be honest: the moral one and the profit one. A shopkeeper with a sufficient moral motive would perform honest actions even if the profit motive were absent. Therefore, according to the fitness model, this would be a moral action. Herman argues that the fact that the moral motive is sufficient in this situation does not imply it would be sufficient in another. Herman claims that a moral action takes place not only if the moral motive is a sufficient one in a specific situation but also if it is strong enough to prevail over other possible inclinations against the moral law that could arise in different situations. “On a greater-strength interpretation of the fitness model, an action can have moral worth”—she says—“only if the moral motive is strong enough to prevail over the other inclinations.” In the strength interpretation of the fitness model, we will praise the individual whose moral motive prevails over nonmoral motives, bringing us back to the battle citation model. Herman explains: “A greater—strength interpretation of sufficiency would then undermine the claim that there are two notions of moral worth in Kant, and leave just with the battle-citation model’s powerful moral motive.”⁵ The idea of sufficiency, however, is not enough to determine if an action has moral worth. Herman points out that even if tomorrow the circumstances change and the dutiful action as a result is

not done, this does not imply that the action that is done today does not have moral worth.

Herman agrees with Henson that it is natural to accept that in a morally worthy action nonmoral motives may be present, as long as they are not the reason for the agent to act. However, she points out that it is not obvious how a motive could be present and yet not operative.

In order to understand Kant's idea of moral worth, Herman makes a distinction between motives, incentives, desires, and causes. Kantian motives are not desires or causes in the sense of vector-like forces. Desires are incentives (*Triebfedern*), not motives for acting. Following this line of reasoning, she concludes that the doctrine of moral worth can accept overdetermination with respect to incentives, not motives.

Paul Guyer disagrees with Herman, in that he sees no reason to admit that we can have overdetermination of incentives: "An incentive cannot coexist or cooperate with any other motive as one sufficient or independent cause might coexist or cooperate with another, but can become a cause of action only by a maxim that makes it into a reason for that action."⁶

Since an incentive is not a cause, it means that there cannot be overdetermination at all, since overdetermination means that you can have two possible causes for the same event. If an incentive is not a cause per se, then the existence of many incentives cannot be seen as overdetermination.

It is very difficult, however, to see how an incentive can be present and not be taken as a motive. Perhaps what Kant really wanted to show in the philanthropist example in the *Groundwork* is that only the complete absence of feelings can assure that a certain act was done from duty. If one is immersed in a sea of feelings, it is a hard task to acknowledge that these feelings haven't played any part in the actual accomplishment of the action. The only way to be sure that a feeling did not have any role as a motive is not to have had it at all. Once love is there, or hatred, or jealousy, or even sympathy, we can never say if the action has a pure motive.

How could we know that we are not acting from friendship, for instance, when we vote in a provost election for a friend of ours? Although we can be persuaded that we are doing so because we really think that this friend has the best platform for the university in mind, it is still the case that perhaps it is not our reason, but our feelings that are inclining us to vote for him/her. Although it is good to have feelings for our friends, the best reason to vote for a provost is not the feeling we have for him, but the fact that we consider him to be the best choice for the university.

I think, however, that Kant is not saying that only the absence of feelings will make the action morally worthy. After a long discussion with commentators, Henson's fitness report model has proved itself to be the correct interpretation of the *Groundwork* example. In a dutiful action, other inclinations may be present, provided that respect for duty is present and would have sufficed by itself for the accomplishment of the dutiful action.

Allen Wood, in the book *Kantian Ethics*, expresses the Kantian position in a correct way, saying that

we have a duty to *strive* for a pure disposition, so that the motive of duty alone is sufficient. We have this duty because it is "hazardous" to rely on motives besides duty, because the performance of duty on such motives besides is always only "contingent and precarious" (G 4: 390; KpV, 5: 73). We have no duty at all, however, to exclude other motives we might have for doing our duty. (MS, 6: 393)⁷

Kantian virtue, in Wood's conception, is not only to have the strength to follow the moral law, but also to strive for the purity of motives that lead to that end. That is not the same as having a duty to be a cold moral person. Other incentives could coexist with the moral one as long as the moral law would have been enough to accomplish the moral action. However, I disagree with Wood when he states that "we are not in the least to blame if we require incentives others than duty, so long as we in fact do it."⁸ I agree that it is better to act according

to duty moved by sensible incentives than to act in opposition to the moral law. However, the philanthropist example undoubtedly shows that the genuine moral action is the one which is done only because of duty and not because of any pleasure the agent may have. Then we could be blamed if we need other incentives besides moral law, since this would not be a pure moral action.

Incorporation thesis and weakness of the will

One problem in any theory that claims that agents act from reasons is weakness of the will. According to the traditional doctrine of the weakness of the will, the agent has a reason to act in a certain way and decide to perform the action A1, but she does A2, because she is driven by a powerful inclination to do it. This apparently contradicts what Kant states in *Religion*:

Freedom of power of choice has the characteristic, entirely peculiar to it, that it cannot be determined through any incentive *except so far as the human being has incorporated it into his maxim* (has made it into a universal rule for himself, according to which he wills to conduct himself); only in this way can an incentive, whatever it may be, coexist with the absolute spontaneity of the power of choice (of freedom). (Rel, 6: 23–24)

And some pages later, while analyzing the three degrees of evil, he mentions the weakness of the will (*fragilitas*):

I incorporate the good (the law) into the maxim of my power of choice; but this good, which is an irresistible incentive objectively or ideally (*in thesi*), is subjectively (*in hypothesi*) the weaker (in comparison with inclination) whenever the maxim is to be followed. (Rel, 6: 29)

The first quotation was baptized by Henry Allison as the incorporation thesis. Allison wrote in *Kant's Theory of Freedom* that there are two theses regarding

freedom, both of which play an important role in Kant's theory of rational agency, the incorporation thesis, and the reciprocity thesis:

The former expresses what I take to be Kant's central insight regarding rational agency. Inclinations and desires do not of themselves constitute a sufficient reason to act. But do so insofar as they are taken up or incorporated into a maxim by the agent. This means that an act of spontaneity or self-determination is involved even in inclinations or desire-based (heteronomous) actions. I view this thesis as the key to both Kant's moral psychology and his conception of agency.⁹

If we accept the incorporation thesis, an incentive can never constitute, by itself, a sufficient reason to act, but only if it is taken into the maxim, or, in another words, if it is taken as a motive. If this is the case, weakness of the will can never occur, since it is exactly the opposite case: we act from an incentive, against the maxim. In weakness of the will, an incentive seems to constitute a sufficient reason to act. If we accept that weakness of the will can be the case sometimes, then, at least in these situations, the incorporation thesis does not hold.

How can one solve this problem in Kant's philosophy? Should we accept the incorporation thesis and deny the possibility of the weakness of the will? Or should we accept the undeniable fact of weakness of the will and accept at least some exceptions to the incorporation thesis.

In a discussion with Marcia Baron, Allison suggested that weakness of the will is not really weakness, but self-deception. When an agent explains her action based on weakness of the will, she is misleading others and herself about what her motive, and her maxim, really was.¹⁰

The contradiction between the incorporation thesis and the weakness of the will can be explained in the following four propositions:

- (1) S knows the principle P1
- (2) S can act according to P1

- (3) S freely chooses P1 as a Maxim M1
- (4) S acts contrary to M1

Allison's choice is to deny 3; that is, the agent S really chose P1 as a Maxim.

I do not think that Kant really presents a solution to this puzzle in his work, although I will try to see what kind of answer could be taken as a Kantian solution.

Rational agency and irrational actions

In his famous article "Action, Reasons and Causes," Donald Davidson argues that actions have causes and these are composed of desires and beliefs. Following Davidson, we are now used to talking about causes of actions.

However, many contemporary authors defend the thesis that actions are not caused by desires and beliefs. In a recent book, *Rationality in Action*, John Searle claims, against the so-called classical model of action, that rational actions are not caused by beliefs and desires and that, in general, only irrational and nonrational actions are caused by beliefs and desires. He writes:

In the normal case of rational action, we have to presuppose that the antecedent set of beliefs and desires is not causally sufficient to determine the action. . . . We presuppose that there is a gap between the causes of the action in the form of beliefs and desires and the "effect" in the form of the action. This gap has a traditional name; it is called "the freedom of the will."¹¹

In Kant's works, the difference between irrational and rational actions is expressed in the well-known distinction between *arbitrium brutum*, *arbitrium sensitivum*, and *arbitrium liberum* in the *Critique of Pure Reason*. Human beings have *arbitrium sensitivum*, meaning they can be affected by sensibility,

but the sufficient reason can only be given by reason. This is related to the presupposition of practical freedom, as the capacity for independence of being determined by the inclinations.

If a set of beliefs and desires is not enough to determine the action, should one consider that reason is the cause of an action? If so, how can one admit weakness of the will, when an incentive moves the agent to act contrary to what has determined the will? Although some commentators argue that reasons can be the cause of an action, I will try a different solution.

Kant does not use the expression “cause of an action,” with the term “cause” being reserved for the phenomenal world. When discussing motives and incentives, he talks about the ground of determination of the will (*Bestimmungsgrund*), which is a motive (*Bewegungsgrund*). Incentives (*Triebfedern*) are responsible for the subjective determination of the faculty of desire. In the passage of *Religion* on which the so-called incorporation theory is based, Kant does not state that only motives can be a cause of an action, but that only motives can determine the will.

If I have a strong desire D, but I decide to act for reason R, I will perform the action Ar. In another case, if I decide to perform the action Ar, and nonetheless I act according to my desire and perform the action Ad, I act contrarily to the previous determination of my will. It could be that I changed my mind between the first determination of the will and the action. This will not be a case for weakness, but I just changed my mind about which action I wanted to perform, meaning that the determination of will has changed. But there are cases in which I really decided to perform the action Ar, based on any previous maxim, rationally decided and—in the cases of weakness—I do Ad. If the actual action is not the same as the decided action, it falls off the strict model of rational agency, understood as the domain of actions that are maxim based. If we accept weakness, we must have a more humble solution: the domain of rational agency does not have the same extension as all voluntary action. In the model of rational agency there is no room for weakness, although it is a fact

in real actions. The domain of the voluntary is, then, wider than the domain of rational action and can also contain irrational actions. What are irrational actions? Actions that are accomplished against the determination of the will.

In order to understand rational agency and the possibility of irrational actions, we have to ask whether all actions are accomplished according to maxims.

Do we always act from maxims?

The incorporation thesis states that incentives can only determine the will if incorporated into a maxim. Do we always act from maxims? In the *Reflexionen* on moral philosophy, Kant writes:

Character is the general ruling *principium* in the human being for the use of his talents and qualities. Thus it is the constitution of his will in good or evil. A human being who has no constant *principium* of his actions, hence no uniformity, has no character. . . . The human being can have a good heart, but still have no character, since he is dependent on changing circumstances and does not act in accordance with maxims. (*Reflexionen* 1113, 15: 496)

Character is the capacity to act according to maxims, which means acting with a certain regularity. People who let emotions determine their actions, in the sense that they act contrary to what they have decided, have no character. These people act in a very unpredictable way. In this sense, one can have a human being with a good heart and no character, and another one, with a good character, although he lacks tender emotions. Character, for Kant, is not the same as virtue; however, whoever has character is less likely to act in an irrational way, in the sense of acting without maxims, or against his own maxims.

The idea that an agent cannot have his action determined by incentives unless these are incorporated in his maxim leads to the need to investigate Kant's idea of maxims. I will attempt to investigate what the notion of maxims

consists of and whether the subject always acts according to maxims. This is relevant if we consider that Kant discusses weakness of the will in *Religion*, indicating the possibility that, at certain moments, the agent opts for a maxim and acts contrary to it.

What is a maxim? In the *Critique of Practical Reason*, Kant gives a precise definition of maxims: they are practical principles, which are defined as “propositions that contain a general determination of the will, having under it several practical rules” (KpV, 5: 19). Maxims, as well as the practical law, are principles of will. However, they differ in the extension of their validity: “They are subjective, or *maxims*, when the condition is regarded by the subject as valid only for his own will, but are objective, or practical *laws*, when the condition is recognized as objective, that is, valid for the will of every rational being” (KpV, 5: 19).

According to this definition, maxims are principles which apply to the will of a particular subject, and therefore are not a practical law valid for every rational subject. In the *Groundwork*, Kant offers the definition of maxims at two points: “A maxim is the subjective principle of volition; the objective principle . . . is the practical law” (G, 4: 400n); “a *maxim* is a subjective principle of acting, and must be distinguished from the *objective* principle, namely, practical law” (G, 4: 421n); it is “the principle in accordance with which the subject *acts*; but the law is the objective principle valid for every rational being, and the principle in accordance with which *he ought to act*, i.e., an imperative” (G, 4: 421n).

In the first section of *Groundwork*, we are told that actions are morally judged according to their maxim:

An action done from duty has its moral worth *not in the purpose* which is to be attained by it but in the maxim in accordance with which it is decided upon, and therefore does not depend upon the realization of the object of the action but merely upon the *principle of volition* in accordance with which the action is done without regard to any object of the faculty of desire. (G, 4: 400)

In other words, maxims are principles of action which are judged by the categorical imperative for the ascertainment of their capacity to be considered a law. Thus, the action of the agent who makes a false promise is not judged specifically as an action, but instead as a principle that underlies this action. In this case the principle can be formulated as a maxim *to make a promise with the intention of not keeping it*. It is not the action, but the maxim that is morally judged.

In the second section of the *Groundwork*, Kant submits a few maxims to the categorical imperative, expressed in the formula of nature's law: "Act as if the maxim of your action were to become by your will a universal law of nature" (G, AA 4: 421). Again, it is not the specific action of a particular suicide which is judged, but the principle "from self-love I make it my principle to shorten my life when its longer duration threatens more troubles than it promises agreeableness" (G, AA 4: 422).

The difference between the rule of a particular action and a maxim as a subjective practical principle is more easily understood when we examine imperfect duties. An example of an imperfect duty toward others is the duty of benevolence. That being benevolent is an imperfect duty shown in the *Groundwork*, in the example of the philanthropist (G, AA 4: 398). In the *Doctrine of Virtue*, the maxim of benevolence is mentioned as part of a duty to promote the happiness of others (MS, AA 6: 456). However morally right it may be to foster the maxim of benevolence, it does not tell us what specific action should be recommended in a particular case. When and with how much should we help those in need? This question pertains to casuistry: how far should one spend one's resources in practicing beneficence? Surely not to the extent that one would finally come to need the beneficence of others. The maxims of benevolence relate to an imperfect duty and allow a latitude in order to fulfill this duty; consequently, each action should be examined on a case-by-case basis.

If Kantian ethics accesses and judges maxims of action, one assumes that the rational subject acts according to maxims, or subjective practical principles

of action. Maxims express general policies of action, or principles according to which a subject really acts.¹² Maxims can be expressed in the form: in S situations, I should perform A actions. The maxim of fake promises can be expressed in this form: every time I find myself in trouble (S situation) I should make a false promise (A action).

Maxims are principles which a rational subject fosters freely, self-imposed rules valid for his will. Rüdiger Bittner¹³ highlights the difference between empirically determined maxims and regularities of a subject's behavior. He provides the following example: suppose that after each insult aimed at me follows a behavior on my part that is meant to be understood as vengeance toward the subject who insulted me. We can describe this empirical and psychological regularity as "S does not let any insult go unpunished." Even if this regularity really verifies itself and we can in fact predict S's behavior, it is simply a regularity of action, but not a regularity which is desired as such. This is different from the example given by Kant himself when he says that somebody may have as a maxim not letting any insult go unpunished. In this case it is not a question of a mere regularity verified by an external agent, but the subject himself wants this regularity. The maxim, therefore, is a desired regularity, a principle which the subject provides for himself.

The Kantian theory of action presupposes that the rational subject acts through self-imposed principles which dictate the kind of action to be realized in certain situations. Such regularity in his conduct is not an empirical or psychological regularity, but one desired as such. However, a rule desired by my own will is not enough to constitute a maxim. We can think of a rule desired by my own will which establishes only the means to a particular end, for example "I should exercise three times a week." This kind of regularity is desired as a prescription which serves as a means to attain a certain particular end, be it health or good physical shape. One could give another example of rules of this kind: "I want to dine every Monday at friends' houses."¹⁴ Or "I want to look for shelter from the rain so as to not get wet."

Maxims are more than mere regularities of conduct, even if desired by the subject himself. They are connected to life goals, to the comprehension of who I am and what I want for myself. Thus, “Not making false promises when I’m in trouble” is articulated as a rule of life for me: honesty and truth are more important than benefiting from every situation.

Maxims and affects

Could actions caused by intense affects,¹⁵ such as anger or dread, be classified as actions which follow a maxim? Defining maxims as practical subjective principles and establishing that the criteria for judging the morality of actions falls upon maxims, Kant allows us two interpretations: either every human action can be expressed through maxims or *only* actions which can be expressed through maxims are appropriate for a test of morality. Allison sustains the first position: with the exception of reflex behavior, every human action can be expressed according to maxims. Onora O’Neill¹⁶ holds the latter:

Because a universality test for autonomous beings does not look at what is wanted, nor at the results of action, but merely demands that certain standards of consistency be observed in action, it has to work with a conception of action that has the sort of formal structure that can meet (or fail to meet) standards of consistency.

According to O’Neil, only actions which express synthetically structured principles or descriptions can be consistent or inconsistent. By demanding that an action be judged only if described according to its maxim, Kant would maintain that an action, in order to be morally accessible, must have a specific formal structure.¹⁷

We should understand the difference between these two possible interpretations: according to the first, non-reflex actions performed by a

rational subject are always expressed through maxims. According to the latter, only acts expressed through maxims can undergo moral investigation, and thus, those which do not present this formal structure fall outside the range of the Kantian theory of action.

In order to resolve this matter, we must first question whether we always act according to maxims. What would be the Kantian position?

According to *Religion within the Limits of Reason Alone*, as we have seen, an inclination can only be the cause of an action so far as the individual has incorporated it into his maxim (Rel, 6: 24), which Allison calls the “incorporation thesis.”¹⁸ According to this thesis, if desires and sensitive inclinations incline an *arbirtrium sensitivum* toward an action, they do not, in themselves, constitute a sufficient reason for action. They can only constitute such if the subject freely decides to take them as motive, only if they are taken or incorporated in the individual’s maxim; in other words, taken as motive by an act of freedom. This means that an act of spontaneity or self-determination is involved even in actions based on dispositions and desires. This theory implies that one may, therefore, act through inclination, but that, even so, there was an act of spontaneity by the individual which took this inclination as motive.

Even if the theory of incorporation seems extremely abstract to those who sustain the theory that desires and beliefs are sufficient reasons for action, it reflects the common conflict between reason and sensibility. In order to illustrate this conflict, suppose, for example, that a married man feels desire for another woman. His desire will not be a sufficient reason for action. Between the desired and the concupiscent action there will be a free decision on the individual’s part, in which other reasons, morals, and discretions will be taken into account. Only then will this process of deliberation, or self-determination, result in a decision regarding the action. The theory of incorporation is a little stronger still: it does not simply address a decision regarding a particular action, but the choice of a principle of action. When the individual becomes

determined to cheat on his wife due to desire, he is electing a principle, aside from a particular course of action.

Kant, however, mentions weakness of the will, in which we are aware of the maxim but lack sufficient strength to follow it. Weakness of the will, the first degree of evil, is an example of an action which cannot be submitted to a maxim precisely because it was realized contrary to the maxim consciously assumed. Here we have two possibilities: either we admit that the actions which follow weakness, realized contrary to the maxims previously decided by the agent, such as in the case of strong emotions, are not to be subjected to moral accessibility as they do not have an adequate formal structure, or we admit that there is no such thing as weakness, which is self-deception. The first position brings two consequences: a positive one, as it would solve the alleged contradiction between the incorporation thesis and weakness of will; and a negative one, since it would mean that Kantian morality only has adequate tools for judging actions which have been expressed in a certain formal structure.

The second position answers that weakness itself should be seen as something for which we are responsible; in other words, it is self-imposed. Undoubtedly, this answer is a way of maintaining the internal coherence of the Kantian system. This position was sustained by Allison, who held an insightful discussion about this point with Márcia Baron.¹⁹

In *Idealism and Freedom*, Allison states that weakness should be seen as something for which we are responsible; in other words, it is self-imposed. If the tendency to subordinate the cause of morality to the causes of self-love is evil, weakness must be understood as the opening to temptation. It is only our tendency for self-deception which leads us to see it as a hard fact. Weakness of the will, as a natural error, presents a problem to Allison. One of Baron's objections to this lies precisely in the incompatibility between the theory of incorporation and this first moment of evil. Allison answers by saying that weakness should not be taken as an error or as hard fact, but identified as

an opening to temptation; this opening or susceptibility, on the other hand, is the condition which facilitates something like weakness, which is self-deceptively seen by the individual as hard fact, part of his nature, which he laments but is not responsible for. Baron insists that, in weakness, there is a real commitment to moral law, which would not exist in the other two degrees: impurity and perversity.

If the interpretation of weakness as self-deception elucidates the relation between the theory of incorporation and weakness through the denial of the latter, it faces, however, two problems: the Kantian texts, which really do address the classical issue of moral weakness, and the experience of the common human being, for whom weakness of the will is a fact. I will now attempt to investigate which would be the Kantian answer to these problems. My hypothesis, to be investigated, is that weakness exists, but must be fought with virtue. Then virtue is to be defined as *fortitudo*, as the strength to fight strong inclinations.

Another query relevant to the relation between maxims and affects has to do with the fact that, at times, the agent adopts a maxim and *chooses* to act against it. Now it is no longer a question of weakness of the will, as weakness is an exception not reflected in the maxim. Let us consider, for example, that we have as maxim “Not Lying.” Suppose that in a certain situation we conclude that lying would be profitable to us, and we effectively lie. How would Kant analyze the situation? He would probably claim that our maxim is not “Not lying,” but instead “lying when it is profitable to us.” Thinking that our maxim was “Not lying” was simply the result of an act of self-deception. Kant goes beyond that, expressing in *Religion* that we have a prior commitment toward a disposition to adopt maxims of self-love or moral maxims: “The disposition, i.e., the ultimate subjective ground of the adoption of maxims,”—in other words, the many maxims particularly adopted by an individual—“can be one only and applies universally to the whole use of freedom” (Rel, 6: 25).

Not only does Kant state that actions must be coherent with the chosen maxim, but that the maxims chosen by an individual are not independent from each other in their moral character. To Kant, it would be contradictory

to suppose that an individual might be committed to a universal principle of morality which prescribes a certain kind of action and, at the same time, to another maxim which prescribes an action which is incompatible to the maxim of morality, “nor can a human being be morally good in some parts, and at the same time evil in others” (Rel, 6: 24).

With the exception of the actions which can fall under the description of weakness of the will, the relation between actions and their maxims is one of a previous moral commitment, a commitment which is either one of not making any exceptions of any kind to the moral law, or of doing so, in case one is inclined toward personal gain in some way.

Although the majority of Kantian moral work presupposes the model of an agent who acts according to her will, it leaves open the possibility of an agent who does not have the sufficient strength to do what she has decided to do. Virtue is the only way to make all actions conform to the rationality of the model of rational agency. In analogy with Aristotle, only virtue could heal the akratic. But virtue in Kant, unlike Aristotle, will not be based on habituation and cultivation of good character, but on the effort to build a strong will that could fight against the force of inclinations.

But are all akratic actions irrational? In a line of argument that goes back to Davidson, akratic action has the following characteristics:

- 1 It is intentional. Since the agent acts according to a reason and not by compulsion, one can say that the agent has acted intentionally.
- 2 It is conflicting. Since the agent has a better reason not to have done what she has done, one has a conflict of reasons.
- 3 It is irrational. Since the agent has acted for a reason while having a better reason to have done otherwise, then it is incoherent, and, consequently, irrational.

There are, however, many authors who maintain that akratic actions are not irrational. One of the better arguments for this position is the argument of

internal reasons. For Bernard Williams, the only authentic reasons for action are founded in a motivational system that includes beliefs, preferences, personal projects, and our emotional states. Contrastingly, external reasons are not necessarily based on our personal motivational system and can sometimes be excessively impersonal and abstract. For supporters of this view, since our best judgment is not founded on our motivational personal system, to act against it, is not irrational.

This solution is possible for the non-Kantian philosopher, who admits that our commitment to our rational maxims will be the same as our external reasons. Since sometimes they are not based on our own desires and inclinations, we may end up acting in a different way from how we had decided to act.

For the Kantian point of view, however, the problem is already there.

Trying to solve the puzzle

After the discussion between Allison and Baron, some other Kantian philosophers tried to solve the puzzle from a Kantian point of view: if we cannot act without taking incentives in the maxim, or taking incentives as motives, how is the weakness of the will possible in Kantian terms?

Iain Morrison tries to solve this problem. He addresses the problem by explaining the apparent contradiction in weakness “typically, weakness of the will is understood as the phenomenon whereby an agent is somehow overcome by a desire, upon which she knows she should not act. But such an occurrence does not seem possible when all desires must be incorporated into maxims before they can be acted upon.”²⁰ The contradiction lies in the fact that weakness implies the presence of unmaximized or unprincipled action. He explains that to be committed to a maxim is to regard the actions as justified or good, in some sense, and that maxims seem to make actions good. Morrisson proposes the possibility:

Kant can solve this problem by claiming that having a maxim (i.e., committing to a course of action) is not exactly identical to regarding an action as justified or good (in some sense). That is to say, he can reconcile the ubiquity of principled action and the existence of weakness of the will by modifying the notion of principled (i.e., maximed), such there is principled (i.e., maximed), and yet weak, action.²¹

In order to show that a weak action can also be a principled action, Morrison makes a difference between happiness-related maxims and pleasure-related maxims. An agent may fail to follow a happiness-related maxim and follow a pleasure-related maxim instead. Morrison takes the example of an agent who desires to eat a lot of chocolates but also desires to lose weight. She will have to select which desire is more important and will form an interest or a maxim based on that. Suppose that losing weight is more important to this agent, she will form maxims directing her behavior toward her *idea* of happiness. But suppose that the agent eats chocolate, against her maxim of losing weight, will this action lack a maxim? According to Morrison, this action is done according to a pleasure-related maxim. He claims that

on Kant's account, then, weak action is chosen and somewhat rational. . . . We act weakly when we focus on the immediate end of our choice and ignore the sense in which this choice is "not good" for happiness. In short, Kant can reconcile principled action with weakness of the will insofar as there can be, on his account, a difference between being motivated by happiness and by pleasure.²²

Although based on a refined explanation, the solution Morrison offers is not far from the Allison solution. According to Allison, weakness is self-deception, because the agent really chooses the weak action, therefore taking another maxim as her rule. For Morrison, also, the agent chooses the action that will give her immediate pleasure. For both authors, the weak action is not without a maxim, it is a principled action motivated by pleasure. In this sense

Morrison, as well as Allison denies the existence of the weakness in the radical sense of an action without any maxim.

Knowing people and predicting actions

Now I would like to address an important question regarding the philosophy of action in Kant: can we predict people's actions? Sometimes we wish this were possible. We try every day to explain actions and to predict them, based on what we know about the person, her character, her psychological moods, her preferences. But is this a possible task? Of course if we do not know the psychological inner states of a person, then this prediction is at least difficult. But suppose that we know this person deep inside, suppose that we knew, as Leibniz would put it, the "concept" of this person,²³ could accurate prediction then be possible? Kant considers the problem of freedom in human action in the *Critique of Practical Reason*. The aim of this *Critique* is to show

whether pure reason of itself alone suffices to determine the will, or whether it can be a determining ground of the will only as empirically conditioned. Now, there enters here a concept of causality justified by the *Critique of the Pure Reason*, although not capable of being presented empirically, namely that of freedom; and if we can now discover grounds for proving that this property does in fact belong to the human will (and so to the will of all rational beings), then it will not only be shown that pure reason can be practical, but that it alone, and not reason empirically limited, is unconditionally practical. (KpV, 5: 15)

Kant asserts here a notion of causality that is not a natural causality. However, Kant needs to solve the contradiction between freedom and nature. This Kantian contradiction has an analogy with the Leibnizian contradiction between eternal truths and freedom. Kant tries to solve the contradiction

between freedom and the mechanism of nature by stating that the necessity of nature appertains only to the attributes of the thing that is subject to time-conditions, consequently only to those of the acting subject as a phenomenon. Regarding the subject as a phenomenon, the determining principles of action reside in what belongs to past time, and is no longer in his power. Among what is not in his power, his own past actions can be included and the character that these may determine for him in his own eyes as a phenomenon. But the same subject can also view himself as not determined by any law of nature, to whom nothing is antecedent to the determination of his will:

But the very same subject being on the other side conscious of himself as a thing in himself, also views his existence also *in so far as it is not subject to time-conditions*, and regards himself as only determinable by laws which he gives himself through reason; and in this existence of his nothing is, for him, antecedent to the determination of his will, but every action, and in general every determination of his existence, changing conformably with inner sense, even the whole sequence of his existence as a sensible being—is to be regarded in the consciousness of his intelligible existence as nothing but the consequence, and never as the determining ground of his causality as a *noumenon*. (KpV, 5: 97)

The modifications of the subject's existence, varying according to his internal states, are nothing to him as a *noumenon*. They do not give any determining principle to him as a *noumenon*. The only causality here is self-determination. In his conscience of himself as a *noumenon*, his psychological states and external states should be viewed as a result of self-determination.

So considered, a rational being can now rightly say of every unlawful action he performed that he could have omitted it even though as appearance it is sufficiently determined in the past and, so far, is inevitably necessary; for this action, with all the past which determines it, belongs to the one single phenomenon of his character, which he gives to himself and in accordance

with which he imputes to himself, as a cause independent of all sensibility, the causality of those appearances. (KpV, 5: 98)

Although this view could be seen as odd, it agrees perfectly with judicial assumptions, in which nobody can deny responsibility for his own acts and put the responsibility on some physical necessity. If the agent, however, attributes this action to some bad habits, he is still responsible for this. That is why when an immoral action is performed, the subject experiences “a painful feeling produced by the moral sentiment, and which is practically void in so far as it cannot serve to undo what has been done”(KpV). Even if the agent cannot undo what is done, this painful feeling produced by the moral sentiment is right, because it shows that the agent could have done otherwise. Kant accepts that he can regard his error as a natural consequence, as a result of bad habits; however, the subject knows that he cannot deny he was responsible for what he had done. That is why Kant states that the pain is quite legitimate, since

the sensible life has, with respect to the intelligible consciousness of its existence (consciousness of freedom), the absolute unity of a phenomenon, which, so far as it contains merely appearances of the disposition that the moral law is concerned with (appearances of the character) must be appraised not in accordance with the natural necessity that belong to it as appearance but in accordance with the absolute spontaneity of freedom. (KpV, 5: 99)

What is Kant's position about the predictability of actions? If one has the intellectual intuition of a man, could his action be predicted? Kant states in the *Critique of Practical Reason* one of the most interesting ways to understand psychology and freedom of the will. He claims that

it were possible to have so profound an insight into a man's mental character as shown by internal as well as external actions, as to know all its motives, even the smallest, we could calculate a human being's conduct for the future

with as much certainty as a lunar or solar eclipse; and could nevertheless maintain that the human being's conduct is free. If, that is to say, we were capable of another view, namely, an intellectual intuition of the same subject (which is certainly not given to us, and instead of it we have only the rational concept), then we would become aware that this whole chain of appearances, with respect to all that the moral laws is concerned with, depends on the spontaneity of the subject as a thing in itself. (KpV, 5: 99)

This view seems contradictory. How could we foresee the actions of a subject in the same sense that we predict the eclipse of the moon and at the same time attribute his actions to the spontaneity of the subject as a thing in itself, of the determination of which no physical explanation can be given? The answer Kant gives is that character is itself the product of a choice. Even those who show since childhood such early wickedness, that they are taken to be born villains, are judged for what they do or leave undone (KpV, 5: 99–100).

Although there is the possibility of uniformity of conduct, the evil principles were, however, at one point voluntarily adopted. And this was the choice of the noumenal self that will have the character of the subject as a result. To have character, in Kantian terms, is to be predictable. The man of principles has character: of him, we know exactly what to expect, he does not act on the basis of his instinct, but on the basis of his will (Ant, 7: 285).

Character differs from temperament, which is a natural disposition. Temperaments may be influenced by the physical condition of a person; consequently this division follows that of the humoral division:

Thus it follows that the temperaments, which attribute only to the soul, may perhaps also be influenced mysteriously by the physical condition of a person. Furthermore, since, first, they permit principal classification into temperaments of feeling and activity, and since, second, each of them can be linked with excitability (*intensio*) or the slackening (*remissio*) of the vital power, it follows that only four simple temperaments (similar to the four

sylogistic figures through the *medius terminus*) can be established: the sanguine, the melancholy, the choleric, and the phlegmatic. (Ant, 7: 286–87)

The sanguine person is light-blooded, he is disposed to gaiety and always hopes for success, while the melancholic person is heavy blooded and disposed to melancholy, being uneasy, mistrusting, and critical. The choleric person is hot-blooded, the one who is quickly ablaze like a straw fire. He is impetuous, his dominant passion is ambition, and he likes to be involved with public affairs. The fourth temperament is the cold-blooded phlegmatic. This temperament, as we will see later is the one which is most appropriate for morality. He controls his affections very easily, “he does not get angry easily; he first considers whether he should become angry” (Ant, 7: 290).

If temperament can make people predictable, it is not yet character. While temperament means what nature makes of men, character is already what man makes of himself. If temperament gives a natural prediction, character gives a rational prediction of actions, since “a character relates to that property of the will by which the subject has tied himself to certain practical principles which he has unalterably prescribed for himself by his own reason” (Ant, 7: 292).

Kant assumes, however, that there is still a difficulty concerning the relation between the combination of freedom and the mechanism of nature. The difficulty is as follows: even if it is admitted that the supersensible subject can be free with respect to a given action, although as a subject also belonging to the world of sense, he is under mechanical conditions with respect to the same action. But if we allow that *God* as universal first cause is also *the cause of the existence of substance*, it seems as if we must admit that a man’s actions have their determining principle in something which is wholly out of his power, in the causality of a Supreme Being distinct from himself, and on whom his own existence and the whole determination of his causality are absolutely dependent (KpV, 5: 100).

In this case, man would be only a marionette and self-consciousness would not give him freedom, but only make him a thinking automaton. Kant thinks,

then, that Leibniz's answer is misleading: to rely on human freedom and on the decrees of God is not to accept human freedom. If this were the case, then consciousness of his spontaneity would be mere delusion. If space and time were then qualities of things in themselves, then finite beings would be only parts (attributes and modes) of God. If the ideality of time and space is not adopted, nothing remains but *Spinozism*, in which space and time are essential attributes of the Supreme Being himself, and the things dependent on him are not substances, but merely accidents inhering in him.

The only way to escape deterministic views is to say that time and space do not belong to things in themselves but only to appearances.

The difficulty mentioned above is resolved briefly and clearly as follows. If existence in time is only a sensible way of representing things which belongs to thinking beings in the world and consequently does not apply to them as things in themselves, then the creation of these beings is a creation of things in themselves, since the concept of a creation does not belong to the sensible way of representing existence or causality, but can only be referred to noumena. (KpV, 5: 102)

To solve the contradiction, Kant then maintains that God is not a creator of appearances, thus he is not the cause of actions in the sensible world.