

Kant on freedom and practical irrationality

Ralf M. Bader

Merton College, University of Oxford

“[T]he will stands halfway between its a priori principle, which is formal, and its a posteriori incentive, which is material as it were at a crossroads” (4:400).

I Introduction

It is clear that, according to Kant, we have transcendental freedom. It is not so clear, however, how far this freedom extends. In particular, it is unclear whether there is freedom within the prudential realm: whether we can freely choose which ends of self-love to pursue as well as how to pursue them. Relatedly, it is clear that we can be practically irrational by ignoring the commands of pure practical reason and siding with self-love instead. However, it is not clear whether Kant recognises any other forms of practical irrationality, in particular whether there is room for weakness of will in terms of implementing one’s commitment to give priority to duty, as well as room for prudential irrationality. In short: What kinds of choices does transcendental freedom encompass? Can we fail to act morally despite having a good will? Is there freedom within the prudential realm? And can we be practically irrational in prudential matters?¹

This paper identifies the basis of our transcendental freedom as well as the precise location of freedom in Kant’s theory in order to determine the extent to

¹These questions importantly bear on how one is to understand the normativity of pragmatic imperatives. First, there is the question whether pragmatic imperatives are genuine imperatives that necessitate the will. For Kant necessitation presupposes contingency. Non-compliance has to be possible in order for a requirement to classify as an imperative. This suggests that there has to be freedom within the prudential sphere if there are to be imperatives within this sphere. Yet, it is far from clear whether the relevant kind of freedom can be found in the prudential realm. Second, there is the question of how one can make sense of the analytic practical principle of willing, especially the proviso that someone who wills the end wills the means ‘insofar as reason has decisive influence on his actions’. This proviso suggests that the condition of having decisive influence can fail to be satisfied, in which case there can be room for contingency and hence for an imperative that necessitates the agent. Making sense of this proviso requires an account of the ways in which reason can fail to have decisive influence, i.e. of the ways in which one can be practically irrational in matters of prudence.

which we are free. By examining the metaphysical basis of freedom as well as by examining how practical deliberation proceeds and how incentives are incorporated to result in actions, it will be argued that freedom is located at the crossroads between reason and sensibility and that we have freedom due to being both rational and sensible creatures that face two heterogeneous types of incentives. As a result, freedom is restricted to the choice between rational self-love and self-conceit. What we can choose between are the orderings of the two sources of incentives, namely the pure incentive of reason and the empirical incentive deriving from sensibility. This ordering of the sources of incentives then determines which incentive is incorporated in a given context.

2 The basis of transcendental freedom

To determine how far our freedom extends we need to examine the basis of transcendental freedom: both the epistemic basis for considering ourselves to be free and the metaphysical basis on which our freedom rests.²

2.1 The epistemic basis

How do we come to know (from a practical point of view) that we are transcendently free? The epistemic basis for attributing transcendental freedom to ourselves is the fact of reason. Since being able to act independently of inclinations is a presupposition of being able to act morally, one can infer from one's consciousness of being bound by the moral law that one is transcendently free (cf. 5:29-31, 6:49 fn & R5434). Morality categorically requires us to act in certain ways, independently of what inclinations we happen to have. In order to be able to follow these categorical requirements, one has to be able to set aside external stimuli and instead act on the basis of the self-wrought feeling of respect in accordance with a law that one has imposed on oneself. Given that we categorically ought to comply with the moral law and that this is only possible if we can set aside our inclinations and act independently of them, it follows that this presupposition has to be satisfied and that we are, accordingly, able to do so. It is

²We will set aside a further basis for attributing freedom to ourselves, namely that freedom is a presupposition of agency. "For I can be quite indifferent as to the origin of my state in which I am now to act; I ask only what I now have to do, and then freedom is a necessary practical presupposition and an idea under which alone I can regard commands of reason as valid" (8:13). The idea that freedom is a necessary practical presupposition is often interpreted as a general claim to the effect that all actions have to proceed under "the idea of freedom" (4:448). However, it is best understood as the claim that one has to consider one's actions to not be necessitated by sensibility if the commands of reason are to be valid. Freedom is presupposed, not in order to make sense of choice in general, but in order to consider oneself to be free from necessitation by sensibility and hence free to do what reason requires.

in this way that the moral law provides us with epistemic access to our freedom and hence classifies as the ratio cognoscendi of freedom (cf. 5:4 fn).

The epistemic ground for attributing freedom to ourselves is restricted to the conflict between self-love and duty. It regards our ability to set aside our inclinations and act independently of the influences of sensibility. We are justified in considering ourselves to be free insofar as being bound by morality presupposes that we are able to set aside self-love and comply with the requirements of duty. The fact of reason thus only supports a rather limited conception of freedom.

The question now is whether our freedom extends any further than this. Are we able not only to set aside inclinations and self-love altogether in favour of duty, but also to make choices, say, within the prudential realm? For instance, can we freely decide to act on some inclinations rather than others, i.e. which ends to pursue, or take certain means rather than others, i.e. which actions to perform? The fact that we only have epistemic access to our freedom on the grounds that we have to be able to set aside inclinations does not imply that our freedom is restricted in this way, but is compatible with a more expansive freedom that might also apply to prudential choices. To address this issue we need to examine the metaphysical basis of transcendental freedom.

2.2 The metaphysical basis

The metaphysical basis of our freedom resides in the fact that we have a hybrid will, what Kant calls an ‘arbitrium impurum’ at 23:379. Our hybrid nature consists in us occupying an intermediate position between, on the one hand, non-rational animals that are merely sensible and, on the other, non-sensible beings, such as the holy will, that are purely rational. As a result, we have a ‘mixed human Willkür (*libertas hybrida*)’ (R5618).

The fact that we are both sensible and rational creatures ensures that our will is confronted with two radically heterogeneous types of incentives, namely incentives of self-love (= the anticipation of pleasure) and incentives of duty (= respect for the moral law). As a result, our will is an arbitrium liberum that is affected but not necessitated. If the different kinds of incentives were both to necessitate, then contradictions would arise whenever the different incentives would pertain to incompatible actions. The will would then be necessitated to perform incompatible actions.³ Accordingly, there has to be contingency. Reason and sensibility do not determine our actions but only provide incentives that can be incorporated by Willkür. The agent’s elective faculty of choice (Willkür) can choose which of

³If the two types of incentives were homogeneous rather than heterogeneous, then they could be weighed up. The strongest set of incentives could then outweigh the others. Yet, they differ in kind and derive from different sources, corresponding to the two sides of our dualistic nature. Freedom is thus not a matter of indifference. It is not because the incentives are of equal strength that one can decide amongst them, but because they are heterogeneous and cannot be weighed up against each other.

these different sources of incentives to prioritise. Arbitrium is the “choice between inclination and reason” (R4222). As a result, the agent has liberty of indifference and can choose amongst the different options. Our freedom can in this way be traced back to our dualistic nature that combines both rationality and sensibility.

By contrast, there is no liberty of indifference in the case of beings that do not have a hybrid will. On the one hand, a non-rational animal has an arbitrium brutum that is mechanically determined by sensibility. On the other, reason determines action in the case of purely rational creatures. In both cases, there is a unique source of incentives. This ensures that the will is necessitated. There is hence no liberty of indifference and no contingency in willing. Even the holy will does not choose amongst different options.⁴ “Seine willkühr ist also eine zwar freye, aber unbestimmte Willkühr (die göttliche ist bestimmt). Das *arbitrium brutum* ist determinirt *secundum rationes sensitivas*, das göttliche *secundum intellectuales*, das menschliche durch keines. Seine Handlungen *hätten alle können* nach der *Vernunft* geschehen. Daher ist er frey” (R4226). Similarly in R4227 we are told: “Wären die Menschen vollig intellectual, so wären alle ihre Handlungen thatig determinirt, aber doch frey, und würden nur in Ansehung der veränderlichen Gelegenheiten zufällig seyn. ... Wären sie vollig sinnlich, so wären ihre Handlungen allein passiv determinirt; ihnen könnte nichts imputirt werden, und sie würden keiner Belohnungen und Bestrafungen fähig seyn. Nun sind sie zum Theil sinnlich, zum Theil intellectual, doch so, daß die Sinnlichkeit freylich das *intellectuale* nicht passiv machen kan”.

There are two types of incentives on which we can act. Our freedom consists in being able to freely decide between these two sources of incentives. In particular, we can order the two sources of incentives, either giving priority to duty or to self-love.⁵ We are not (causally) necessitated or determined by our incentives, but are instead merely affected by them. It is up to us which side of our nature to prioritise and thereby determine which incentive to incorporate. The conception of freedom explained in terms of the metaphysical basis, accordingly, coincides with that which we attribute to ourselves on the basis of the fact of reason. It is likewise restricted to the choice between rational self-love and self-conceit. Freedom is located at the crossroads between reason and sensibility.

3 Moral irrationality

Practical irrationality can (potentially) arise along two dimensions:

I. END-IRRATIONALITY

⁴A pure will has positive freedom in the sense of spontaneity. It is not determined mechanically but by intelligible/rational grounds.

⁵Freedom regards not the incorporation of particular incentives, but the ordering of the sources of incentives. This means that we do not directly choose which incentives to incorporate, but choose the fundamental *Gesinnung* that determines the incorporation of incentives.

one can be irrational in terms of the commitments one makes, i.e. in terms of the ends that one pursues.

2. ACTION-IRRATIONALITY

one can be irrational in terms of the way in which one implements one's commitments, i.e. in terms of the actions that one takes.

One can either make the wrong commitments or fail to adequately implement one's commitments. Both of these dimensions can potentially be found in each of the two practical domains, namely the moral domain and the prudential domain.

3.1 Moral end-irrationality

Transcendental freedom makes room for moral end-irrationality, which consists in having a bad will. By giving priority to self-love when one should instead prioritise duty, i.e. choosing self-conceit over rational self-love, one makes the wrong kind of commitment. Freely choosing a bad *Gesinnung* amounts to ignoring the requirements of morality. One is not responsive to the requirements of pure practical reason and hence practically irrational.

It is possible for an agent to become self-conceited by a free choice on the part of *Willkür* that gives priority to self-love. A self-conceited agent will attempt to satisfy his inclinations even when doing so is impermissible. Even though he has no reason to perform this kind of action, given that the resulting happiness lacks value and is merely agreeable, he is nevertheless motivated to perform these actions since pleasure results from satisfying his inclinations.⁶ This kind of immoral behaviour is possible because the agreeable can motivate even when it is not good. Pleasure results from performing actions that satisfy the agent's inclinations, ensuring that there is motivation, even when happiness lacks value and does not give rise to reasons. This makes it possible for the agent to side with self-love rather than with duty, taking an option for which he has no reasons.⁷

3.2 Moral action-irrationality

Whilst one can be practically irrational by prioritising self-love over duty, there is no room for practical irrationality in the form of moral action-irrationality. This is because one's *Gesinnung*, which is the agent's ordering of the sources of incentives

⁶Though it lacks value, the agent is likely to mistakenly consider it to be good (and hence to be reason-giving), i.e. the agent is likely to be subject to an illusion of value. Moreover, the agent is also likely to engage in self-deception, trying to convince himself into thinking that his action does not contravene against morality after all.

⁷This fact merely makes this form of irrationality possible but does not explain it or render it intelligible, something that cannot be done given that it involves a transcendently free choice of a bad *Gesinnung* for which no explanation can be found. All that we can do is to identify the conditions that make it possible. We cannot, however, explain what makes it actual.

(cf. 6:36), determines what maxims one adopts. One either has a good will, which consists in a fundamental commitment to duty such that one makes the pursuit of self-love conditional upon moral permissibility. Or one has a bad will, which consists in a fundamental commitment to self-love such that one pursues one's self-interest independently of the requirements of morality and complies with duty only on condition that doing so furthers (or at least does not detract from) the interests of self-love, i.e. conditional on 'prudential permissibility'. Kant's rigorism consists in recognising only these two possibilities (cf. 6:22 fn) – either one gives priority to duty (= rational self-love) or to self-love (= self-conceit). The goodness of a *Gesinnung* is thus not a matter of degree but is binary.⁸

Willkür freely chooses the agent's *Gesinnung* and thereby indirectly selects specific maxims. The *Gesinnung* is the agent's fundamental maxim. It functions as the ground of the adoption of all other maxims. It "applies to the entire use of freedom universally" (6:25). The choice of specific maxims in any given context is determined by the *Gesinnung*. Specific practical rules are put forward by instrumental reasoning. These are evaluated by pure practical reason as to whether or not they are universalisable. They are then either adopted or rejected/modified on the basis of the agent's fundamental *Gesinnung*. Whether an incentive is incorporated into the rule, in which case it becomes causally efficacious and thereby becomes the agent's principle of action (= his maxim), depends on whether it is permissible. If it is universalisable and hence permissible, it will be adopted and an incentive will be incorporated, independently of whether or not the agent has a good will.⁹ Yet, if it is impermissible, then it will only be adopted by a self-conceited agent, since such an agent does not pay any heed to the impermissibility verdict but gives priority to self-love. By contrast, it will be rejected/modified (whether it is rejected or modified depends on whether its universalisation yields a contradiction in conception or merely a contradiction in will) by an agent who has a good will, since he subordinates considerations of self-love to those of duty.

⁸Whilst there is a commitment to rigorism at the transcendental level, this is compatible with a gradable evaluation at the empirical level, since the empirical manifestation of the underlying intelligible character can vary continuously.

⁹Even though incentives do not automatically/mechanically give rise to actions but have to be incorporated, incentives will be incorporated independently of whether or not one has a good will as long as the practical rule put forward by instrumental reasoning is permissible. In those case there is no alternative and hence no possibility to not act on that incentive. We should thus reject the claim that "[p]art of our freedom is the ability to say 'No' to any desire that we have" (Baron: 1993, p. 431). We can only say 'No' when morality tells us not to act on a particular incentive, not however when morality does not object. Even though there is no choice in terms of whether to incorporate the incentives of self-love into permissible practical rules, agents always do have choice with respect to the principle on the basis of which one incorporates incentives, namely whether do so conditionally or unconditionally. Moreover, it should be noted that the agent does not accept or reject particular incentives, but instead accepts or rejects practical rules that are put forward by instrumental reasoning. Only practical rules can be assessed for universalisability – incentives do not have the right logical form.

In other words, specific maxims are a function of:

1. instrumental reasoning applied to the agent's desires (which includes both present desires and ones that the agent anticipates to have and from the satisfaction of which he can accordingly anticipate to receive pleasure) operating with the agent's beliefs about means-ends connections and the empirical circumstances in which he finds himself, resulting in a practical rule that specifies the actions to be taken in order to best satisfy his desires,
2. the permissibility of the practical rules put forward by instrumental reasoning, which is determined by pure practical reason on the basis of an assessment whether these practical rules are universalisable or lead to contradictions (in conception or willing) when universalised, and
3. the agent's *Gesinnung*, which either leads to the adoption or to the rejection/modification of the proposed maxim on the basis of its permissibility, whereby the maxims suggested by instrumental reasoning will always be adopted by a self-conceited agent who gives priority to self-love, whereas such maxims will only be adopted on condition that they are morally permissible in the case of an agent who has a fundamental commitment to rational self-love.

The fundamental maxim, together with facts about the universalisability of proposed practical rules, in this way determines the incorporation of incentives and hence the adoption of the maxims on which an agent acts. As a result, no room opens up for weakness of will when it comes to implementing one's commitments. It is not possible for an agent to have a good will but then fail to follow through by adopting specific maxims that are incompatible with this fundamental commitment. The adoption of specific maxims is determined by the agent's *Gesinnung* and does not involve any further choices. We are not dealing with a sequence of choices, whereby one first chooses one's *Gesinnung* and then makes further (transcendentally free) choices that require sufficient strength of will to implement this commitment and overcome the temptation to deviate therefrom.¹⁰

The *Gesinnung* determines the adoption of specific maxims. Specific maxims are a function of the fundamental maxim together with facts about the context in which the agent finds himself. The choice of specific maxims is thus not separable from the choice of the *Gesinnung*.¹¹ As a result, the specific maxims on which one

¹⁰Accordingly, we should reject attempts to explain weakness of will in terms of being swayed by contrary inclinations that lead to us having specific maxims that deviate from our general maxims (cf. Hill: 2012 sections 4 & 5, as well as Johnson: 1998 who distinguishes between specific 'motivating maxims' and general 'justifying maxims').

¹¹Even if maxims are not separable from the *Gesinnung*, there can be a gap between maxims and actions. Cf. "denn zwischen der Maxime und der That ist noch ein großer Zwischenraum" (6:46).

acts cannot come into conflict with one's *Gesinnung*. There is no further choice to be made and hence no room for practical irrationality (in the form of weakness of will) to arise. This means that if one has a good will, then one will only adopt universalisable maxims. If someone adopts an impermissible maxim, then it is not possible for this agent to have a good will and merely to lack sufficient strength of will to implement his commitment.¹² Instead, acting on impermissible maxims implies having a bad will.

There are only further choices when dealing with extended courses of actions that consist of multiple choices. In such a case, an agent can initially adopt the correct priority ordering but then later on change his fundamental maxim and thereby overturn his initial commitment. This, however, is a matter of changing one's commitments, rather than retaining the commitment and failing to follow through. In short, such an agent is to be thought of as having lapsed in his commitment, rather than as merely having failed to implement his commitment.¹³

Neither the choice between the two fundamental dispositions, nor the adoption of specific maxims, is a matter of the comparative strength of the different incentives facing the agent. We should not think of choice as being the outcome of competing forces.^{14,15} Instead, *Willkür* spontaneously chooses its *Gesinnung*. The adoption of specific maxims then turns out not to be a matter of choice at all but is determined by one's fundamental disposition (together with the context). That the choice of one's *Gesinnung* is not determined or affected by the strength of various incentives should be clear, given that the fundamental disposition concerns the ordering not of particular incentives that have particular strengths but the ordering of sources of incentives. The strengths of incentives cannot be rel-

¹²Frailty/weakness of the heart (cf. 6:29) can be understood in terms of self-deception (cf. Allison: 1990, p. 159). One deceives oneself into thinking that one does in fact have a good will and only fails to live up to one's commitment due to weakness when what is really going on is that one has made the wrong commitment. In this way, one uses the notion of weakness of will to attempt a cognitive reconciliation of having acted impermissibly with one's self-conception as someone who has a good will. This means that one engages in self-deception (*vernünfteln*) not only to attempt to convince oneself that one's actions are not in conflict with morality, but also that one has the right disposition even when one has acted impermissibly. Self-deception can, in this way, be found at the level of both the specific maxim and the underlying disposition.

¹³This means that we are to think of the good will roughly along the lines of what Ameriks has called the 'particular intention' view, rather than the 'general capacity' view or the 'whole character' view (cf. Ameriks: 2003, chapter 7). However, one should note that 1. the good will is to be construed as a fundamental *Gesinnung* that determines the particular intention/maxim on which the agent acts in a particular choice situation but is not to be identified with this intention/maxim, and that 2. in addition to the notion of a good will which is concerned with the agent's disposition in a particular choice, we also have the notion of virtue (construed as an extensive magnitude) which is concerned with the choices that an agent makes throughout his life (cf. Bader: 2015).

¹⁴At least, this holds as regards the noumenal level – the phenomenal manifestations of noumenal choices can take the form of competing forces.

¹⁵Cf. McCarty: 2009 and Timmermann: 2003, pp. 203-207 for accounts of the will in terms of competing forces.

evant, since we do not order incentives but their sources. Since the notion of strength is not applicable, the understanding of weakness of will as involving a lack of sufficient strength is likewise inapplicable.¹⁶

Sensibility can only lead us astray when we are dealing with affects and passions.¹⁷ Affects bypass the higher faculty of desire altogether. This means that practical reason does not come in and that one does not act freely. Actions brought about by affects are not free deeds (cf. 6:223). Instead, this kind of action is mere behaviour that is not based on maxims and that merely involves the lower faculty of desire. Unlike free deeds, which flow from (and hence cannot contravene against) one's *Gesinnung*, this kind of behaviour is mechanically brought about. Accordingly, it is not attributable and imputable to the agent.¹⁸

Actions that are based on passions, by contrast, do involve the higher faculty of desire.¹⁹ Even though they are based on maxims, they are nevertheless not free actions.²⁰ This is because passions interfere with pure practical reason. They prevent the agent from acting on the basis of pure practical reason and hence cannot be controlled, thereby not leaving the agent any alternative options (cf. 6:29 fn & R1514-1515). Although this kind of action goes via the higher faculty of desire, it is not attributable to the will.²¹ The problem is not that we have an agent who is presented with various options and then chooses the wrong option due to weakness of will (contra Morrisson: 2005), but that the agent's sensibility

¹⁶The incommensurability of moral and prudential incentives, moreover, causes further difficulties for thinking of choice in terms of comparative strength.

¹⁷For a detailed and helpful account of affects and passions and their role in empirical psychology cf. Frierson: 2014, chapters 7.1 & 7.2

¹⁸One can only be responsible and blameworthy for letting oneself become susceptible to affects (likewise for passions).

¹⁹Cf. "Leidenschaft setzt immer eine Maxime des Subjects voraus, nach einem von der Neigung ihm vorgeschriebenen Zwecke zu handeln. Sie ist also jederzeit mit der Vernunft desselben verbunden, und bloßen Thieren kann man keine Leidenschaften beilegen, so wenig wie reinen Vernunftwesen" (7:266).

²⁰Cf. "Der Affect thut einen augenblicklichen Abbruch an der Freiheit und der Herrschaft über sich selbst. Die Leidenschaft giebt sie auf und findet ihre Lust und Befriedigung am Sklavensinn" (7:267; also cf. 7:264). The claim that there are maxim-based actions that are not free actions and that do not flow from the agent's *Gesinnung* does not conflict with the claim at 6:25 that the *Gesinnung* is the ground of the adoption of specific maxims. This claim is restricted to maxims that have been freely adopted and does not apply to maxims that result from passions. Moreover, this should not be surprising, given that Kant distinguishes between humanity and personality (cf. 6:26), which suggests that there can be agents that can engage in instrumental reasoning and act on maxims, yet that lack transcendental freedom. We can, accordingly, think of affects as operating at the level of animality and passions at the level of humanity, whilst freedom only comes in at the level of personality.

²¹Kant notes at 7:267 that passions are "(der Form nach) ... nicht bloß *pragmatisch* verderblich, sondern auch *moralisch* verwerflich" since they mistakenly treat a proper part of the agent's end as being the whole end, which runs contrary to the formal principle of reason (cf. 7:266). This, however, does not mean that actions resulting from passions are objectionable, nor that they are in any way imputable to the agent.

is interfering with reason and preventing him from acting in any way other than what is dictated by the passions.²²

Inclinations, accordingly, cannot interfere with a good *Gesinnung* and cannot overpower our commitment to prioritise duty. Although affects and passions can make us act in objectionable ways, they cannot make us choose something that is contrary to our fundamental maxim. We can only fail to act on the basis of our *Gesinnung* if we do not act freely at all. There is thus no room for practical irrationality in the moral domain when it comes to implementing one's commitment. Weakness of will only comes in the form of not making the right commitment, that is, failing to give priority to duty and instead favouring self-love, but not in the form of not carrying out the commitment one has made. In short, there is only room for moral end-irrationality but not for moral action-irrationality.

3.3 Imperfect duties and latitude

Having a good will implies that one only adopts universalisable maxims and hence does not contravene against duty. This applies straightforwardly in the case of perfect duties. Yet, things become more complicated when it comes to imperfect duties. This is because imperfect duties involve latitude.

Someone who considers a maxim of non-beneficence will be required by pure practical reason to modify it and render it universalisable. Morality in this way requires him to adopt a maxim of beneficence. Yet, pure practical reason does not require specific acts of beneficence, but simply a maxim to not only further one's own happiness but also that of others (cf. 6:393). Since the latitude that is involved in this imperfect duty allows for a range of possible actions, the question arises whether there is anything to determine which action the agent performs. Having a good will only ensures that some action in this range is performed. Accordingly, it would seem to be up to the discretion of the agent when and how to be, say, beneficent. The latitude that is involved in imperfect duties might thus open up room for freedom when it comes to complying with imperfect duties.²³ In that case, freedom would not be restricted to the choice between rational self-love and self-conceit. Instead, we would have freedom within the moral sphere.

This latitude might also allow for the possibility of weakness of will. Since this duty does not require specific actions, the agent has leeway in how to comply with the duty of beneficence. In particular, one can fail to act in a beneficent way in any given context yet still have adopted a maxim of beneficence. However, an agent who adopts a maxim of beneficence but then always opts for non-beneficent

²²In this way, things can go wrong in the practical sphere in ways that are analogous to how error in the theoretical sphere can be due to the unnoticed influence of sensibility on the understanding (cf. A294-295/B350-351).

²³Cf. "The widest duties, such as beneficence, leave a wide 'playroom' for free choice as to when, how, and to what extent to promote the end" (Hill: 2012, p. 115).

action would seem to fail to live up to his commitments. Although each action by itself is unproblematic, the various actions taken collectively do not seem to conform with the maxim of beneficence. Such a problematic pattern of non-beneficent actions would seem to be possible without the agent lapsing in his commitment. As long as the agent does not fail to be beneficent as a matter of principle, i.e. as long as he does not switch to a maxim of non-beneficence, but simply makes use of the latitude that is involved in a maxim of beneficence, such an agent does not have to be construed as no longer having a good will. This suggests that an agent can have a good will, yet nevertheless not adequately implement his commitments when it comes to imperfect duties.

To address these concerns, we need to supplement the account of how actions are generated by adding a further step. In particular, we need to explain how maxims give rise to actions. We have seen that adopting a maxim amounts to incorporating an incentive into a practical rule, thereby making it efficacious and providing it with motivational force. The practical rule in this way becomes the agent's principle of action (= maxim, i.e. the subjective principle on which he acts). Moreover, we have seen how the fundamental maxim, i.e. the *Gesinnung*, determines the incorporation of incentives into specific maxims, by making the adoption of maxims contingent on their universalisability in the case of a good *Gesinnung*. The specific maxims that agents adopt can then feature in practical syllogisms, whereby the maxim functions as the major premise.²⁴ Judgement subsumes a particular action-token under the type that is specified in the maxim, allowing reason to draw an inference resulting in an action.

This is crucial for understanding imperfect duties. The categorical imperative requires us to reject maxims of non-beneficence and instead adopt disjunctive maxims that include beneficent actions alongside non-beneficent actions. We have to modify the non-universalisable maxim of non-beneficence in such a way as to render it universalisable. We thereby end up with a limited maxim that is concerned not only with one's own happiness but also with that of others (cf. 5:34-35). The resulting maxim is disjunctive and allows for the subsumption of both self-regarding and other-regarding actions. This is what explains the latitude of the duty of beneficence, namely that either kind of action can be subsumed under the maxim.²⁵ Importantly, it is not up to *Willkür* to pick one of the disjuncts. *Willkür* only selects the maxim of beneficence, but not which action to subsume thereunder. Instead, it is up to judgement to subsume actions under the disjunctive principle. Since judgement is not a volitional faculty, it is not a matter of choice as to which disjunct is acted upon. This implies that the agent does not have a free choice between the different disjuncts. There is hence no freedom in this respect, but only the operation of a non-volitional faculty of judgement.

²⁴Cf. "The term 'maxim', like 'condition', is also taken from logic, deriving from *sententia maxima*, the name of the first major premise in a polysyllogism" (Beck: 1960, p. 81).

²⁵Cf. "Kantian beneficence" (Bader: manuscript) for a more detailed account.

Which disjunct is selected is something that is attributable to the agent (since it is the result of the (spontaneous) operation of the faculty of judgement, i.e. it has its origins in the agent and not in external causes), but is not within his control. Since it is not a matter of choice, the agent cannot be praised/blamed for having picked one disjunct rather than the other – the only thing for which he can be held responsible is the choice of the disjunctive maxim and hence for the fact that some disjunct or other is acted upon. The agent is thus responsible for ϕ -ing, since this action is performed as a result of having been subsumed under a maxim that the agent has freely chosen. Likewise, the agent would have been responsible for ψ -ing, had that action been subsumed instead. Yet, the agent is not responsible for ϕ -ing rather than ψ -ing.

4 Prudential irrationality

The crossroads at which the will stands and where it must make a transcendently free choice is that between the two possible orderings of the sources of incentives. Our freedom is restricted to the choice between rational self-love and self-conceit. There is, accordingly, no freedom in the non-moral sphere. This seems to preclude the possibility of practical irrationality in the prudential realm.

Theoretical irrationality in prudential matters, by contrast, can arise straightforwardly. On the one hand, the input with which instrumental reasoning operates can be faulty. This can be because the agent is ignorant of or mistaken about (i) his ends, (ii) means-ends connections, or (iii) the causal context in which he finds himself, i.e. about which means are available. On the other hand, the processing of the input can involve mistakes if the agent does not correctly calculate the best means for the ends.

Particularly widespread is ignorance of one's ends. Although everyone necessarily has happiness as an end, it is epistemically indeterminate what happiness consists in. Every finite ratiocreature by its very nature as a sensible being with needs will have various ends of self-love throughout his or her life. Yet the agent often does not know what these ends are. It is for this reason that prudence merely issues counsels rather than strict imperatives (cf. 4:418). The gout sufferer, for instance, does not have a sufficiently clear and determinate conception of his future desires for there to be anything within his deliberative framework that can outweigh the certain pleasures that result from satisfying his present desires (cf. 4:399). Since the future is uncertain, more immediate inclinations are epistemically privileged over more distant ones. As a result, he acts contrary to his interests objectively construed (i.e. he would act differently if he were to know the relevant facts), yet not contrary to what is best given his epistemic situation.

We will set aside these theoretical problems and work with idealised agents who know their ends as well as the relevant causal connections and facts about the causal environment in which they act.

4.1 Setting ends

“An end is an object of free Willkür, the representation of which determines it [i.e. Willkür] to action, whereby it [i.e. the object] is brought about. Every action thus has its end and since no one can have an end without having himself made the object of his Willkür into an end, so it is an act of *freedom* of the acting subject, not the effect of *nature*, to have an end of actions” (6:384-385; also cf. 6:381).

We set our own ends by means of a free choice on the part of Willkür. Many interpreters have taken this to imply that there is room for freedom in the prudential sphere, insofar as we can decide which ends of self-love to set for ourselves. On this view, (non-moral) ends need to be freely adopted or endorsed in order for them to become normative and to provide reasons for actions. The idea is that the agent starts out with an array of (possibly conflicting) desires and then decides which of them to endorse, thereby forming his conception of happiness. As a result of freely endorsing certain desires, they acquire the normative status of being the ends that he is to pursue.²⁶ Hypothetical imperatives then tell the agent which means he has to take to realise these ends.²⁷

This endorsement account should, however, be rejected.²⁸ Although we set the ends of our actions, this does not imply that we have any choice amongst

²⁶Cf. “The Incorporation Thesis of itself entails that desires do not come with pre-assigned weights. On the contrary, it is the value placed on a desire or inclination by an agent that gives it its ‘motivational force,’ its status as a reason to act” (Allison: 1996, p. 113).

²⁷This account might allow for both prudential end-irrationality and prudential action-irrationality. End-irrationality is not straightforward. Since endorsing certain desires is what confers normativity upon them, normativity is not prior to endorsement, which ensures that there are no external (prudential) standards with respect to which the agent can be criticised. In particular, one cannot criticise an agent’s end-setting in terms of resulting in less happiness than an alternative way of endorsing desires, since what the agent’s happiness consists in is determined by the desires he has in fact endorsed. His endorsement of desires fixes what his happiness consists in, such that there is no prior conception of happiness with respect to which this choice can be evaluated. Yet, there might nevertheless be standards for the adoption of ends. This is particularly plausible when it comes to conflicting desires, insofar as one would seem to be rationally criticisable for forming a conception of happiness that cannot be pursued in a coherent manner.

Action-irrationality, on the contrary, is more straightforward. If an agent endorses a subset of his desires, then only these desires will be normative for the agent. Yet the remaining desires are still present and can still exert motivational force. Although the normativity of desires is dependent on endorsement, their motivational force is not dependent in this way. As a result, there can be desires that motivate the agent to perform actions that are contrary to what would be required to best implement the conception of happiness that the agent has endorsed. If the motivational force of those desires outside the endorsed subset can outweigh the force of those that have been endorsed, then the agent may well fail to implement his prudential commitments.

²⁸The endorsement approach conflicts with the epistemic as well as the metaphysical basis of our freedom, both of which are restricted to the choice of orderings of sources of incentives. Moreover, it involves a problematic account of pragmatic imperatives, cf. “Pragmatic imperatives and the value of happiness” (Bader: manuscript). In particular, these imperatives are to be understood in terms of wanting rather than willing and are hence based on inclinations rather than acts

different ends of self-love. Our inclinations set our (non-moral) ends. They determine what our happiness consists in and what we are to bring about. Freedom only comes in with respect to the choice whether or not to realise these ends, insofar as one can refrain from doing so when duty commands otherwise. In this sense, it is open to the agent to not adopt a maxim that includes these ends when the proposed action is impermissible.

This means that we need to distinguish two senses of ends. On the one hand, we have pre-incorporation ends. “An *end* is always the object of an *inclination*, that is, of an immediate desire to possess a thing by means of one’s action” (6:6 footnote; also cf. R6881 where Kant says: “that which immediately pleases us is our end”). These are the ends that are set by our inclinations and that function as the input into instrumental reasoning.²⁹ On the other, we have incorporation ends. These are the ends of our actions. They constitute the matter of the maxims that we have freely adopted and specify the states of affairs that are to be brought about by our actions. With this distinction at hand, we can see that the claim at 6:384-385 is concerned with incorporation ends that are the ends of our actions and have been made our ends by freely adopting the relevant maxims.

Which action we perform is a matter of free choice. This implies that which end of action we have is a matter of free choice. Yet, this does not imply that we have any choice amongst various ends of self-love. Instead, the end of action can either be the end that is involved in the practical rule put forward by instrumental reasoning (which operates with the ends set by inclinations), namely one’s happiness, or that of pure practical reason. There is thus no freedom in terms of end-setting. The (pre-incorporation) ends of self-love are given to us by inclination, i.e. they are a function of receptivity. The only question is whether we pursue these ends unconditionally or conditionally and hence whether in a particular choice situation we freely choose to pursue happiness, and thereby make one’s ends of self-love the (incorporation) ends of one’s action, or reject self-love in favour of duty (when the action proposed by prudence is impermissible).

of Willkür, given that they are assertoric imperatives that are based on the actual end of happiness that is of necessity had by every finite rational being. Moreover, instrumental reasoning needs to precede incorporation, and hence has to take as its input pre-incorporation ends, since we need practical rules (rather than ends) in order for pure practical reason to assess for universalisability, given that only practical rules but not ends have the right logical form for being universalised.

²⁹It is sometimes suggested that the endorsement account is necessary for dealing with conflicting desires. If desires conflict, then one cannot take whatever means are necessary for realising one’s ends. The suggestion, accordingly, is that the agent first has to form a (coherent) conception of happiness before instrumental reasoning is applicable. Willkür would in this way play a crucial role in forming our conception of happiness. However, one can simply adjudicate conflicts on the basis of the comparative strength of the desires. Conflicts can be adjudicated internally, so that the pursuit of competing inclinations is an instrumental matter that does not require the will to freely set or endorse ends.

4.2 Choosing means

A different way in which freedom might be thought to enter the prudential realm is not in terms of ends but in terms of means, i.e. freedom with respect to which actions we take to pursue our ends. In particular, there are cases in which prudence generates disjunctive imperatives, requiring the agent either to ϕ_1 or to $\dots\phi_n$. The disjunction is necessary for realising one's ends without any disjunct being necessary.³⁰ This type of situation arises when there are multiple courses of action that all lead to (or, at any rate, are expected to lead to) the same amount of happiness. This, in turn, can either happen if the same ends can be realised equally well in multiple ways, or if there are different sets of ends, only one of which can be realised, whereby the same extent of happiness would be realised in each case. It might be thought that such disjunctive imperatives open up room for choice within the prudential realm, in that an agent can freely choose which of the various options to pursue.³¹

However, we saw in section 2.3 that there is no freedom in the case of disjunctive maxims. The subsumption of action-tokens under maxims is done by judgement, which is a non-volitional faculty. Although there are different possible actions that can be subsumed under the maxim, it is not the agent's Willkür but his judgement that is responsible for subsumption. This means that, not only do we not have freedom in terms of setting ends, we also do not have freedom in terms of choosing amongst (permissible) means.

As in the case of morality, affects and passions can make us deviate from what prudence demands by either bypassing or interfering with the higher faculty of desire. Whereas affects bring about actions via the lower faculty of desire, passions give rise to maxim-based actions and involve the higher faculty of desire. Passions prevent reflection and interfere with empirical practical reason. In particular, instrumental reasoning is prevented from considering all inclinations and only focuses myopically on the passions. As a result, one reasons and act as if one's entire happiness consisted in the satisfaction of one's passions, which leads one to act in ways that are detrimental to one's happiness (cf. 7:266, 7:410 & R6610).

4.3 Shaping inclinations

The discussion so far has operated on the assumption that the ends of self-love are simply given, that they are purely a function of receptivity. This assumption, however, is incorrect. Which actions one performs has an effect on what desires/ends one will have in the future. Our actions affect not only which desires are satisfied but also which desires we will have. Desires are to some extent

³⁰Kant accepts disjunctive imperatives at 29:1017.

³¹This would make room for choice without opening up room for practical irrationality, given that all the different options are tied in terms of strength, thereby not making it possible for the agent to pick the weaker of two incentives / the less prudent of two actions.

shaped by our actions. They are not entirely exogenous but partly endogenous. Prudence, accordingly, is not just a matter of calculating the best way of satisfying a given set of inclinations, but is also a matter of shaping one's inclinations.

Given that agents can influence and partly shape their desires, instrumental reasoning is not provided with a unique set of desires that stays constant no matter which action is performed. Since one's desires constitute the standard by means of which one is to assess actions from a prudential point of view, this implies that instrumental reasoning cannot utilise one and the same standard for evaluating different actions. There will not be a unique set of desires with respect to which one is to evaluate different actions, but instead different sets of desires corresponding to different actions. This means that the agent will be faced with shifting standards. Although the agent can evaluate the extent to which particular courses of actions will lead to the satisfaction of the desires that the agent will have if the relevant actions are performed, he cannot make comparisons across the different sets of desires.

One way in which one can ensure comparability is by normalising desire satisfaction. A 0-1 normalisation that treats the maxima and minima as equivalent (i.e. the complete satisfaction of desires is equally good across all sets and likewise for the complete non-satisfaction) allows for comparability across different possible life histories of one and the same person (as well as for interpersonal comparability).^{32,33} According to the normalised notion, all that matters is the extent to which one's desires are satisfied.³⁴ Someone who has few desires, all of which are satisfied, will be as happy as someone who has a large number of desires, all of which are satisfied. There is hence nothing that speaks in favour of having, say, more or fewer satisfied desires as such. Prudence then tells one to take that action which brings about the greatest satisfaction of desires. One should perform that action which ensures that the desires that are had, given that this action is performed, are satisfied to a greater extent than in the case of any other action

³²Hausman: 1995 argues that this is the only way in which one can achieve comparability.

³³For simplicity, the possible life histories are assumed to be of the same length, so that there is no need to decide between normalising temporal well-being and normalising lifetime well-being.

³⁴Kant needs a normalised notion of happiness in order to make sense of the proportionality of happiness and virtue that is part of the highest good. Virtue and happiness must have the same structure to allow for the relevant proportionality condition to be satisfiable. Since the notion of virtue has a maximum, the notion of happiness that is meant to be proportional thereto also has to have a maximum. This means that happiness has to consist in desire satisfaction rather than in pleasure and, moreover, has to be normalised so that the complete satisfaction of an agent's desires is construed as the maximal amount of happiness, independently of how many desires the agent happens to have. Whilst there is such a thing as the complete satisfaction of inclinations, since it is possible for all of one's inclinations to be fully satisfied, there is no such thing as the greatest amount of pleasure. Given that pleasure is an intensive magnitude that lacks an upper bound, the idea of the greatest pleasure is a deceptive concept that involves a hidden contradiction (cf. 2:32). This means that, whereas maximal happiness understood in terms of desire satisfaction is an ideal of the imagination, the greatest possible pleasure cannot be an ideal since it is contradictory to think of the greatest degree of something that lacks an upper bound.

(i.e. one compares the extent to which the desires D_L are satisfied in life history L with the extent to which the desires D_{L^*} are satisfied in life history L^*).³⁵

This account explains in what way reason is implicated in the shaping of desires, insofar as it is meant to ensure that one's desires are harmonious and form a systematic whole (cf. 6:58). To explain the role of systematicity we have to distinguish between the maximal and the total satisfaction of one's desires. Total satisfaction consists in every desire being fully satisfied, whereas maximal satisfaction consists in one's desires being satisfied to the greatest extent possible. These two notions come apart when desires fail to be co-satisfiable, in which case one cannot have all of one's desires satisfied. If not all desires are co-satisfiable, then one can only satisfy subsets thereof, with the maximal satisfaction consisting in the greatest subset of co-satisfiable desires being satisfied. Total satisfaction, accordingly, amounts to the maximal satisfaction of a set of co-satisfiable desires.

If we normalise with respect to total rather than maximal satisfaction, then this ensures that we should cultivate desires such as to achieve a greater degree of co-satisfiability.³⁶ A more harmonious set of desires allows for greater happiness, insofar as more desires, in particular a greater proportion of the total set of desires, can be satisfied. Reason, accordingly, has the task of bringing rational order and systematicity into our desires. By doing so, it can minimise conflicts and ensure greater co-satisfiability, thereby allowing one to get closer to total satisfaction.³⁷

Shaping desires to ensure greater consistency is an instrumental matter that involves considering how the acquisition, cultivation and suppression of desires affects one's prudential pursuits. Different actions lead to different total sets of desires, as well as different maximal sets of co-satisfiable desires and, finally, to different sets of actually satisfied desires and thereby different extents to which the agent's desires are satisfied. Prudence can require us to acquire or suppress desires because doing so allows us to satisfy our desires to a greater extent, insofar as the resulting set of desires allows for a larger subset to be co-satisfiable. If the agent's desires are co-satisfiable to a greater degree, then the maximal satisfaction of the resulting set will be closer to total satisfaction than that of other sets of desires that are co-satisfiable to a lesser degree. The normative significance of systematicity can thus be explained in terms of greater co-satisfiability allowing for a greater extent to which the total set of desires can be satisfied.

³⁵Cf. Bricker: 1980 for a helpful account of prudence in the face of changing desires.

³⁶Greater co-satisfiability only matters to the extent to which greater satisfaction is in fact achieved. Only when one ends up getting closer to total satisfaction as a result of more desires being compatible do we have reason to cultivate desires in such a way that they are co-satisfiable to a greater degree. Co-satisfiability thus only matters when it is combined with actual satisfaction.

³⁷Our desires are only partly endogenous. They are to a significant extent exogenous since the rational control that can be exerted over them is limited. This explains why the idea of happiness is an idea of the imagination (4:418) rather than an idea of reason. Instead of the whole preceding the parts, the idea of happiness is that of a "tolerable system" (5:73) that can be rendered more or less consistent.

Normalising desire-satisfaction makes it possible for motivation and reason to come apart, such that what one has most motivation to do can differ from what one has most reason to do. The more pleasure an action promises, the greater will be the motivation to perform that action. This is because prudential actions are motivated by the anticipation of pleasure. This is the sensible incentive that provides efficacy to practical rules proposed by prudence. The reasons for performing prudential actions, however, depend not on the amount of pleasure that will result but on the extent to which inclinations are satisfied.³⁸ The greater the extent to which inclinations are satisfied, the stronger the reason to take the action in question. In short, whilst the motivation for acting prudentially is based on the anticipated pleasure, the reason for this type of action is based on the resulting happiness.

As soon as it is possible for an agent's desires to vary from option to option, situations can arise in which the action that brings about the most pleasure is not the action that brings about the greatest happiness. When desires are held fixed, pleasure and happiness cannot diverge. Yet, when desires can vary in terms of their number, intensity or duration (i.e. in terms of their extensive, intensive and protensive magnitudes, cf. A806/B834), such that, say, one life history contains more desires than another, it becomes possible for that action which best satisfies the agent's ends to be distinct from that action from which the agent anticipates the greatest amount of pleasure. If one life history involves more desires than the other life history, such that more desires can be satisfied in absolute terms, yet a smaller proportion of desires is satisfied in the former case than in the latter case, then more pleasure will be anticipated from the former action despite the fact that it leads to less happiness. This is because the normalisation of desire-satisfaction ensures that happiness is a relative notion that is concerned with the proportion of desires that are satisfied, whereas pleasure is an absolute quantity that is not normalised. This type of disconnection makes it possible for the strongest incentives to speak in favour of prudentially sub-optimal actions. This might allow for inclinations to overpower us, making us take the option that makes us less happy but for which there is a stronger incentive.

As in the case of practical irrationality in the moral sphere, what makes irrationality possible in the prudential realm is that reasons and motivation can diverge. More precisely, the general explanation of the possibility of practical irrationality is that there can be motivationally undefeated actions that are such that it would be irrational to perform them, given that these are actions for which one has no reasons (which happens in the moral case) or only weaker reasons (which happens in the prudential case). We saw in the moral case that, even though there is no reason to do what is impermissible, one can nevertheless be motivated

³⁸Cf. R6881 "One should never say that one places one's end in pleasure; rather, that which immediately pleases us is our end, since pleasure is nothing but the relation between an end and our feeling."

to do it since, though not being good, the outcome is expected to be agreeable. That is, we have agreeableness without goodness and hence motivation without reason. In the prudential case it is possible that what an agent has most reason to do (and hence ought to do) is to ϕ whilst there is another action ψ for which the agent's motivation is stronger. The agent has stronger motivation for an imprudent action if it brings about more pleasure than the prudentially optimal action. The pleasure that is anticipated is greater yet the degree of happiness (understood in terms of desire-satisfaction) is lesser. When these two factors come apart, the agent can end up siding with pleasure rather than with happiness, performing an action for which there is greater motivation but less reason and thereby act irrationally. The agent is then not being responsive to the reasons that he has, failing to comply with the commands of instrumental reason by taking an option for which one has less reason. Practical irrationality consists precisely in this failure of being responsive to reasons.

5 Conclusion

We have freedom because we are hybrid beings that are both rational and sensible. We are affected though not necessitated in two different ways. Accordingly, we can choose between two different orderings of sources of incentives. The chosen ordering determines, on the basis of the permissibility of the practical rule that is put forward by instrumental reasoning, how incentives are incorporated and which practical rule one acts upon in a particular situation. There is hence no freedom in the prudential realm. There is neither freedom in terms of prudential end-setting, nor freedom in terms of the actions that are subsumed under one's prudential maxims. We only have freedom with respect to the choice between self-conceit and rational self-love. There is, consequently, only one type of practical irrationality from the point of view of morality, namely that of having the wrong *Gesinnung*. But there cannot be any irrationality or weakness of will in terms of implementing one's commitment. In the prudential sphere, by contrast, there is no irrationality with respect to setting ends. Ends are given to us by inclinations rather than being set by us. There might nevertheless be practical irrationality in terms of the pursuit of happiness, due to the fact that there can be actions that lead to more pleasure despite bringing about less happiness.³⁹

³⁹For helpful comments, I would like to thank the participants of the 'Kantian Freedom' conference at Simon Fraser University, the 'Kant, Metaethics and Value' conference at Trinity College Dublin, and the 'Kant and Freedom' workshop at Oxford.

References

- [1] ALLISON, H. E. *Kant's Theory of Freedom*. Cambridge University Press, 1990.
- [2] ALLISON, H. E. *Idealism and Freedom*. Cambridge University Press, 1996.
- [3] AMERIKS, K. *Interpreting Kant's Critiques*. Oxford University Press, 2003.
- [4] BADER, R. M. Kant's theory of the highest good. In *The Highest Good in Aristotle and Kant*, J. Aufderheide and R. M. Bader, Eds. Oxford University Press, 2015.
- [5] BARON, M. Freedom, frailty, and impurity. *Inquiry* 36, 4 (1993), 431–441.
- [6] BECK, L. W. *A Commentary on Kant's Critique of Practical Reason*. University of Chicago Press, 1960.
- [7] BRICKER, P. Prudence. *The Journal of Philosophy* 77, 7 (1980), 381–401.
- [8] FRIERSON, P. *Kant's Empirical Psychology*. Cambridge University Press, 2014.
- [9] HAUSMAN, D. The impossibility of interpersonal utility comparisons. *Mind* 104, 415 (1995), 473–490.
- [10] HILL, T. *Virtue, Rules, and Justice: Kantian Aspirations*. Oxford University Press, 2012.
- [11] JOHNSON, R. Weakness incorporated. *History of Philosophy Quarterly* 15, 3 (1998), 349–367.
- [12] KANT, I. *Kants gesammelte Schriften*. Reimer/de Gruyter, 1900.
- [13] MCCARTY, R. *Kant's Theory of Action*. Oxford University Press, 2009.
- [14] MORRISSON, I. On Kantian maxims: a reconciliation of the incorporation thesis and weakness of the will. *History of Philosophy Quarterly* 22, 1 (2005), 73–89.
- [15] TIMMERMANN, J. *Sittengesetz und Freiheit*. de Gruyter, 2003.