

Anastasia Berg

# Kant and the Freedom to Do What We Want

**Abstract:** Even a morally good practical agent does not act solely from the recognition of the abstract demands of moral duty. Often, she acts to satisfy desires for particular ends that are not intrinsically moral. But if freedom, as Kant claims, consists in acting from universal principles one adopts from respect for the moral law, how can agents freely act to satisfy desires for particular ends? The standard answer to this question, the so-called Incorporation Thesis, is, I argue, unsatisfactory both as an interpretation of Kant and on philosophical grounds. I propose instead that, for Kant, the capacity to act freely for the sake of a particular, non-intrinsically moral end is not exhausted by the ability to step back, reflect and decide whether a desire is or can provide a reason to act. Rather, Kant shows, the place for the pursuit of particular ends is determined by practically rational agents' spontaneous constitution of their moral character, whereby they subordinate the pursuit of material, particular ends to the pursuit of formal, moral ones or vice versa.

## 1 Introduction

Let us begin with an apparently indisputable fact: a moral agent does not act solely from the recognition of her general moral duties. In performing a concrete action, she acts to realize particular ends, which may be given by, or traced to desires that are not intrinsically moral. In such cases, the explanation of the action will not be exhausted by the fact of the agent's recognition of her general moral duties. To get into view the kinds of case I have in mind, suppose I reach for a piece of cake. Such an act may be morally bad, as when I reach for someone else's piece of cake, having first made sure its rightful owners are distracted; it may also be morally permissible, as when I try a piece freely offered to me because I would like to find out what it tastes like; and it may even be morally worthy, as when I reach for the slice that seems best to me, among many, but only insofar as I recognize that, in addition to my hankering for some cake, it is my duty to sustain my life or offer support to my friend, the aspiring pastry chef. A moral theory ought to render the possibility of—as well as the differences between—these kinds of actions intelligible.

On the face of it, however, it is not obvious how the standard Kantian account of action could make sense of these ordinary intuitions, and in particular of the

idea that agents may come to act freely in pursuit of particular ends given by their desires. For if, as Kant claims, acting rationally and so freely, consists in acting from the recognition of principles (not to mention moral principles), it is not immediately clear how moral agents can ever come to act freely on ends picked out by contingent desires or, as we would say, do something simply because they want to.

The standard way of tackling this problem in the literature has been by means of the so-called “Incorporation Thesis” (IT). According to the IT, we act on particular desires only insofar as we evaluate them and deem them reasons to act by “taking them up” or “incorporating them” into principles of action. The IT has not only been uncritically adopted by the vast majority of Kant interpreters,<sup>1</sup> but has also widely influenced Kantian strands of contemporary practical philosophy—particularly through the work of Christine Korsgaard.<sup>2</sup> It is a testament to the popularity of the thesis that it has become hard for us to so much as recognize the question I am posing here as a real challenge.

However, I will argue that according to Kant, the freedom to pursue particular desires does not fundamentally turn on an encounter with particular desires which we stand back from, evaluate, and endorse or reject. Instead, we shall see, the freedom to pursue particular desires is grounded in the agent’s exercise of her agency in constituting her own moral character.

## 2 A Problem for the Kantian

Let us begin by bringing the problem into clearer focus. Human willing, according to Kant, is a capacity to be the cause of the reality of objects not just in accordance with principles (as is the case with respect to non-rational animal activity) but by means of the representation of principles. Consequently, if an action were per-

---

<sup>1</sup> With respect to Kant’s practical philosophy, no single interpretative position—apart perhaps from John Rawls’ “CI Procedure”—has enjoyed as much support as the IT. For examples of uses of the IT following its introduction by Allison (1991), see Wood (1999, 51–53), Hill (1991a, 86), Timmons (1994, 116–19), Johnson (1998), Morrisson (2005), Bojanowski (2007), Papish (2007, 129–130), Guyer (2008, 51–53 and 293–94), Schadow (2012, 112, n. 90, 180–186, and 291, n. 122), and Kang (2015, 96–101). Jens Timmermann (2003, 15) thinks Allison’s IT account requires supplementation by a more detailed account of the role of maxims in Kant’s theory of action and moral psychology but does not dispute the thesis itself. There are two notable exceptions: McCarty (2008) challenges the incorporation thesis on textual grounds and Dewitt (2018) flags Allison’s approach as generally unsatisfactory on philosophical grounds. I will return to both dissenting views below.

<sup>2</sup> Korsgaard’s account distinctly evokes the IT (see Korsgaard 1996b, 94; 2009, 105 and 115; and 1996a, 162 and 165). I will return to her account below.

formed not from the recognition of the validity of a principle, but simply and solely because someone felt like pursuing a certain end, it could not be an exercise of the will and would therefore not be free. Since freedom is necessary for the imputation of action, such actions—if they would be actions at all—could not be imputed to an agent. It thus seems to follow that acting on ordinary felt desires cannot be understood as an exercise of will and that therefore one cannot be held responsible for acting on one's desires.<sup>3</sup>

A Kantian account of the ability to act on desires must therefore explain how an agent can come to act on a contingent desire without relinquishing herself to the mercy of natural forces. This way of setting up the problem seems to call for the possibility of a transition from having an ordinary, contingent desire to having a principle that aims at that desire's satisfaction—a transition that would have to itself be an exercise of our freedom. Responding to the challenge, thus conceived, Henry Allison drew attention to a passage in Kant's late work, *Religion within the Bounds of Bare Reason*,<sup>4</sup> in which Allison took Kant to describe just such a transition from having a desire to freely adopting a principle of action. This transition guarantees, for Allison, that the agent's pursuit of that desire is itself an act of freedom. Here is the passage from Kant:

The power of choice [*Willkür*] has the quite peculiar characteristic that it cannot be determined to an action by any incentive [*Triebfeder*] except *insofar as the human being has admitted the incentive [Triebfeder] into his maxim [nur sofern der Mensch sie in seine Maxime aufgenommen hat]* (has made this a universal rule for himself, according to which he wills to conduct himself) (Kant, *R* 6:23–24).

I shall designate this passage the “incorporation requirement” passage, and, following Allison, refer to his interpretation of it as the Incorporation Thesis (IT).

According to Allison, an agent's actions can be “genuine expressions of agency” and therefore “imputable” in cases where an agent acts on desires, as long as the actions are “thought to involve an act of spontaneity on the part of the agent,

---

<sup>3</sup> It is important not to confuse the question just raised, which is the focus of this paper, with a different and distinct problem—namely, that Kant's equation of the will with practical reason leaves no room for the possibility of freely acting on a principle that opposes the moral law (the so-called Reinhold-Sidgwick objection). The question I address here is a different one: not how we can freely act on a non-moral principle, but how we can freely act to satisfy a desire.

<sup>4</sup> Contemporary interest in the incorporation requirement passage dates to the early sixties and, in particular, to John R. Silber's “The Ethical Significance of Kant's Religion” (1960, lxxix–cxxxiv). Rawls refers to the incorporation thesis in his lectures on Kant (2000, 294). Gerold Prauss (1983, 93–94) initiated interest in it in the German-speaking literature in the early eighties. The label by which it is known today—namely, the “Incorporation Thesis”—was given to it by Allison (1991).

through which the inclination or desire is deemed to be an appropriate basis of action” (Allison 1991, 39). While an agent has no control over what she happens to desire or how intensely she desires it, she nevertheless exercises control over her actions insofar as she is never determined to act directly by a desire. According to Allison, for the agent to act on a desire she must step back from it and evaluate it in light of “objective (intersubjectively valid) rational norms,” and can then accept or reject the desire “as sufficient reason for action” (Allison 1991, 39). A desire can be deemed a reason to act “only with reference to a rule or principle of action, which dictates that we ought to pursue the satisfaction of that inclination or desire” (Allison 1991, 40). On this account, when I have a desire, I consult the objective norms by which I determine myself to act, and then, as long as these norms do not preclude the pursuit of this desire, I find or articulate a principle according to which I *ought* to pursue the satisfaction of the desire. By virtue of adopting this principle I deem the desire a reason to act. If I happen to have a moral reason to pursue the end the desire specifies, all the better, but this is not necessary: that it is permissible to pursue the end the desire aims at is sufficient for the adoption of a principle according to which I ought to pursue it. For example, if I feel like having that slice of cake, I act freely insofar as I evaluate doing so in light of objective rational norms and, on the basis of these, deem my desire a reason for me to act by adopting a principle according to which I ought to pursue the desire in question.

The IT has become the canonical understanding of Kant’s account of the relationship practically rational agents bear to their inclinations, desires, and natural impulses. Interpreters are in near unanimous agreement that Allison’s reading of the passage captures a central Kantian commitment and that Kant was right to hold it—so much so that they rarely take the trouble to consider it in detail, instead often citing it in passing as a premise requiring no defense.<sup>5</sup> I will begin

---

5 See footnote 1 for examples of uncritical mentions and uses of the IT. Tamar Schapiro (2011) is an exception insofar as she defends Allison’s interpretation of Kant in some detail and argues it is a valuable contribution to contemporary philosophy of action. While embracing the IT in general, Marcia Baron (1993a and 1993b) queries whether it can be reconciled with Kant’s brief account of *Gebrechlichkeit*, the so-called Kantian version of weakness of the will. Iain Morriison (2005) responds to Baron and dismisses the threat of conflict between the IT and Kant’s account of weakness of the will. Importantly, both Baron and Morriison endorse the IT. An exception is Richard McCarty (2008), who argues that the IT is not textually supported. While I agree with McCarty’s general point that the incorporation requirement is concerned not with individual action but with moral character, I do not think, as McCarty does, that it has no application to the way we make everyday free choices. This is because I do not think that the free choice that the incorporation requirement is meant to apply to belongs to a separate noumenal world. Janelle Dewitt (2018,

my re-evaluation of the Kantian understanding of the relationship between practical reason and desire by demonstrating that the thesis is not textually supported and is philosophically unattractive.

### 3 Kant on the Principle of Incorporation

According to the IT, desire is incorporated into a maxim. This is Kant's term for a subjective principle of action, or the principle of action upon which an agent in fact acts.<sup>6</sup> Thus, when feeling a craving for some sweets I can act on this craving by reference to a principle like "when craving sweets, eat some." However, when we turn to Kant's own account of incorporation in the incorporation passage, we find that the kind of principle into which incorporation takes place is not an ordinary maxim, an ordinary principle of action, at all. Kant calls the maxim into which incorporation occurs the "supreme maxim" (Kant, *R* 6:36) and describes it as specifying not a particular course of action or even the requirement of the moral law but instead a kind of principle that constitutes the person's entire moral character, what Kant calls one's "*Gesinnung*."<sup>7</sup> This supreme maxim is a single, unifying rule that the power of choice "makes for itself" and is the "first basis [...] for the adoption of good maxims or the adoption of evil (unlawful) ones" (Kant, *R* 6:21). It, Kant adds, "can only be one" and it "applies universally to the entire use of freedom" (Kant, *R* 6:25). By adopting this supreme maxim an agent adopts a governing principle for the exercise of their will in its entirety. It thereby constitutes the agent's moral character at the highest level of generality.<sup>8</sup>

---

68) flags Allison's approach as generally unsatisfactory, but as I shall argue below, I do not think her proposed alternative sufficiently forecloses the dualistic worries she insightfully raises.

6 I follow Stephen Engstrom's (2010) interpretation of "maxims" according to which a maxim, a subjective principle of action, is just the principle of action upon which we do in fact act, while an objective principle of action is a principle upon which we should act. Cf. *G* 4:420 and *KpV* 5:76. When we act morally well, our maxim—our subjective principle of action—is an objective principle of action, whose authority we recognize.

7 I follow Julia Peters (2018) in leaving the term *Gesinnung* untranslated. As she notes, rendering *Gesinnung* as "disposition" (as is the case in both George di Giovanni's translation in *The Cambridge Edition of the Works of Immanuel Kant* and the earlier translation by Theodore M. Greene and Hoyt H. Hudson) is misleading, insofar as "disposition" can also refer to a behavioral disposition or tendency, a connotation absent from the German *Gesinnung*. "Attitude," Werner Pluhar's preferred translation, is closer to the German meaning.

8 It is possible to ask here whether the maxim referred to in the incorporation passage and into which a *Triebfeder* is to be incorporated is indeed the *Gesinnung*-constituting maxim, the single, supreme maxim which the power of choice makes for itself to guide all its exercises. To see that it must be so, it is important to, first, note that the discussion of incorporation occurs as a

This may come as a surprise against the background of the *Groundwork* and the second *Critique*, where no such supreme maxim is discussed. But although Kant mentions the need for this single, unifying principle for the power of choice for the first time in the *Religion*, the idea is implicit in his earlier work and, in particular, in the central thesis that the moral worth of an action depends on the principle on which an agent acts. In the *Religion*, Kant merely makes explicit an implication of this idea; namely, that in order to determine the moral worth of an action we must consult not just any principle of action, but the highest principle from which it is derived: the moral law, in the case of morally worthy action, or the principle of one's happiness, in all other cases.

To this point, that an action's moral worth is determined by the highest principle from which it is derived, Kant adds the claim that in a mature moral agent, this principle is singular and constant across all actions. The argument for this latter claim, which constitutes Kant's "rigorism," is in brief as follows: because the moral law commands universally, properly recognizing its authority means recognizing its authority universally, i.e., over all of one's actions. It follows that an agent cannot act only occasionally from genuine recognition of the moral law, for to allow oneself to occasionally deviate from the commands of the moral law is simply to refuse its authority as universally binding, which is to refuse its authority *tout court*. For whatever set of exceptional circumstances in which an agent would not follow the moral law, the highest principle an agent would act on would be "follow the moral law, except in that special set of circumstances in which you may choose to do otherwise," and so the observance of the moral law, at any and all points, would not be absolute but conditioned on that set of circumstances not obtaining. An agent is thus either committed to the supremacy of moral considerations in all matters, or, in effect, in none.<sup>9</sup> In the *Religion*, Kant titles this overarching commitment the *Gesinnung* or the *Gesinnung*-constituting maxim.

---

part of the discussion of "rigorism"—the way of thinking according to which the human being is "either morally good or morally evil" (*R* 6:22). Kant defends rigorism by arguing that a commitment to the law, which amounts to the incorporation and prioritization of the moral *Triebfeder*, cannot be partial, and that, conversely, any admission of an exception to the moral *Triebfeder* in the form of an incorporation or admission and prioritization of a sensible *Triebfeder* as prior to the moral law is wholly corrupting. In other words, Kant disputes by this very argument the idea that incorporation of a *Triebfeder* can be a local affair and aims to show why any seemingly local deviation is wholly determining of the subject. Second, at *R* 6:36, Kant speaks explicitly of the incorporation or admission of *Triebfedern* as into an agent's "supreme maxim." I discuss the rationale for these claims in what follows.

<sup>9</sup> The claim that maxims stand to one another in a hierarchical system has been defended in Beck (1996, 118), Korsgaard (1996a, 58), O'Neill (1989, 83–85), and Allison (1991, 91–94).

We can therefore conclude that the incorporation maxim introduced by Kant in the *Religion* is the not the kind of thing proponents of the IT purport to find there: it is not a principle that “dictates” that we “ought” to pursue a particular desire or which deems a particular desire “sufficient reason” to act. It is a principle that organizes our most fundamental practical priorities by determining whether we are committed to the moral law absolutely or only conditionally.

But if the text is straightforward on this matter, why have interpreters insisted that incorporation concerns the adoption of an ordinary principle of action in the pursuit of an ordinary desire? Perhaps this is because of the interpreters’ conception of what *must* be the object of incorporation; namely, their conception of desire. This is the second way in which the IT interpretation deviates from Kant’s account of incorporation and the topic to which I will now turn.

## 4 Kant and the Object of Incorporation

As we have seen, proponents of the IT claim that the object of incorporation is a particular desire, which the agent evaluates, and thereby either accepts or rejects as a reason to act. This claim traces to Allison, who saw the IT as a corollary to his well-known “two-aspect” reading of transcendental idealism. I shall consider his account first.

### 4.1 Allison’s Conception of Desire

According to Allison, Kant’s famous distinction in the first *Critique* between appearances and things in themselves should be understood as a distinction between two “aspects” of one and the same underlying reality. The IT is the application of the two-aspect reading to the case of human agency. On this interpretation, we need not attribute to Kant the claim that, to quote Allison, “free agency occurs in a distinct ‘intelligible world’” or that “noumenal activities intervene in the causal order of the phenomenal world” (Allison 1991, 4). Instead, the contrast between freedom and nature is the contrast “between two ‘points of view’ or descriptions under which a single occurrence (a human action) can be considered” (Ibid.).

This framework justifies the attribution of a double character—both empirical and intelligible—to one and the same rational agent. The empirical perspective on human action conceives it as being causally determined according to natural laws. Here human action is to be explained in terms of the “familiar belief-desire” model, according to which “an agent’s empirical character, understood as a set of relevant beliefs and desires, functions as the empirical cause of the action” (Al-

lison 1991, 4). However, from the standpoint of the agent's intelligible character, the agent is not causally determined to act based on the set of their beliefs and desires but is "free," in the sense established by the IT; that is, from the standpoint of the agent's intelligible character the agent is free to adopt objective norms and in the light of those to incorporate their desires into principles of action—it is up to the agent to deem desires reasons to act.

Note, however, that while the intelligible and empirical characters are meant to provide two distinct perspectives on a single underlying reality, the term "desire" figures in both. In the empirical model, it is the sufficient cause of action: from the intelligible perspective, it is the thing that is evaluated and, on the basis of this evaluation, accepted or rejected as a reason to act. How are we to make sense of this double occurrence? To meet the criteria set by the two parts of the account, "desire" ought to mean something different in each, for otherwise what is the sufficient condition for action from one perspective is only a necessary one from the other. But if desire means something different on each side of the divide, it remains mysterious what it is that gets incorporated from the rational standpoint.

## 4.2 The Animal-Desire Conception of Desire

Implicitly addressing this worry, later interpreters sometimes characterize desire not as part of a simple causal explanation, but as the sort of thing that could play a part in an account of animal activity. Allen Wood, for instance, characterizes the faculty of desire in rational beings as the power to determine oneself to act on the basis of an "empirical impulse," where the account of this impulse "applies equally to rational beings and to brute animals" (Wood 1991, 51). Whereas for the mere brute the power of choice is determined by the sensuous, empirical impulse, rational beings are able to resist such impulses and only act on them if they decide to.

This animal-desire version of the IT, the version which characterizes human sensible desire with reference to the desire of brute animals, has been introduced into contemporary ethics. Its most well-known proponent is Christine Korsgaard, who attributes to Kant an account according to which desires are located outside of reason and must be approved of by reason if they are to be acted upon: "At the basis of every desire or inclination, no matter how articulately we can defend it, is a basic suitability-to-us, that is a matter of nature and not of reason" (Korsgaard 2009, 122) or "the reflective mind must endorse the desire before it can act on it, it must say to itself that the desire is a reason. As Kant puts it, we must *make it our*

*maxim* to act on the desire. Then although we may do what desire bids us, we do it freely” (Korsgaard 1996b, 94).<sup>10</sup>

In this account, human desire begins as a product of nature and only afterwards becomes distinctly rational, and thus, distinctly human: it is the reflective endorsement of human desire which distinguishes it from the desires of a non-rational animal.<sup>11</sup> Tamar Schapiro summarizes the point as follows: “When I have an inclination, the reflecting part of me is aware of the non-rational principle that shapes my inner animal’s way of seeing and responding to the world” (Schapiro 2011, 165). Crucially, although these interpreters diverge from Allison’s two-aspect reading of human action, they share in common with Allison the idea that desires are fundamentally extra-rational (it “is a matter of nature and not of reason,” the reflecting part is aware of “the non-rational principle”), and must be reflectively endorsed in order to be freely acted upon.

There are two fundamental problems with this kind of account of desire.<sup>12</sup> The first is the more familiar. Proponents of the IT claim that the sort of thing which, without modification, could either causally determine movement or naturally determine an animal to act, and whose occurrence therefore does not itself involve the exercise of rational capacities, can supply something that can be evaluated and deemed as a *reason* by the rational subject. But if desire can find itself at home in the conceptual habitat of explanation through empirical laws or natural teleology, it is far from obvious that desires are something agents could, from their own perspective, intelligibly deem as *reasons* to act.<sup>13</sup> No doubt a rational agent can re-

---

**10** The text in quotation marks is Korsgaard’s interpretation, not a direct quotation. See also Korsgaard (1996a, 162 and 165 as well as 2009, 105 and 115).

**11** While a non-rational animal’s “perceptions are its beliefs and its desires are its will,” in a human being perceptions and beliefs become the “objects of its attention,” we can “think about them” and, therefore, “our capacity to turn our attention on to our own mental activities is also a capacity to distance ourselves from them, and to call them into question” (Korsgaard 1996b, 92–93). Likewise, she later writes, “The first result of the development of this form of self-consciousness is liberation from the control of instinct. Instincts still operate within us, in the sense that they are the sources of many of our incentives—in fact, arguably, though by various routes, of all of them. But instincts no longer *determine* how we respond to those incentives, what we do in the face of them” (1996b, 116).

**12** For critique of the idea that Kant’s account of moral deliberation is best understood on the model of “reflective endorsement,” see Herman (2008). For criticism of reflective endorsement as inconsistent with a constitutivist reading of Kant’s practical philosophy, see Bagnoli (2009).

**13** This point echoes the one made by Warren Quinn in his influential essay “Putting Rationality in its Place” (1994). Quinn argued that we cannot conceive of desire merely as a brute impulse, or else we could never make sense of the way in which it rationalizes action. Desire, Quinn argues, must therefore be understood to involve more. Another way of making my point is that in Kant’s

spond to a causal force (e. g., a push or pull) or a physical impulse (e. g., nausea) in various ways, but evaluating and responding rationally to a force or impulse, or even finding reason to resign oneself to it, address it, or resist it is not the same as deeming *it* a reason. An agent who finds herself, say, pushed by a crowd, can observe the crowd's movement, ask herself whether it promotes or hinders her interests, and evaluate the general trajectory with approval or disapproval. She can try to resist it or resign herself to it and she can do all this on rational grounds: she can decide resistance would be too costly, or that the pushing can be accommodated without undesired consequences. She can even decide that she has a reason to get to the other side and on that basis consider the push salutary. But no act of resistance or resignation would amount to deeming the push of the crowd itself *a reason to act*.<sup>14</sup> My reason to act cannot be that I have been pushed. The same goes for a bout of hiccups, or an itch: I might find that I have reason to try and hasten the hiccups bout's conclusion, or scratch the itch, but a hiccup attack is not making a rational recommendation that I act in this way or that.

Proponents of animal-desire may respond that my awareness of an extra-rational desire, or the operation of an extra-rational principle, need not operate analogously to the observation of a brute event. An animal's desire is not a mere event but is itself goal- or end-directed. All else being equal, an animal's particular desire belongs within a system of ends that jointly constitute the animal's self-preservation, as an individual and as a member of its species. As such, an animal's desire could be understood to make a suggestion or a "proposal"<sup>15</sup> concerning what it would take to, say, maintain the practical agent's preservation.<sup>16</sup>

I do not think this response is satisfactory. First, it cannot be assumed that a contingent, sensible desire could play the role of a proposal with respect to the agent's reason any more than a brute fact would. At first glance this may seem counterintuitive. Let us consider one of Schapiro's central examples: thirst.<sup>17</sup> Surely, we could justify drinking water by saying something like "I drank because I was thirsty." But while it is true that the feeling of thirst may indicate to us that we are in physical need of water, if when thirsty I proceed to obtain and drink water, it is not the physical indicator of that need that I deem a reason to act, it is rather my

---

system, as long as desire is located in nature, it can never be conceived along the lines that interpreters have wished it would—an evaluative perspective that can rationalize a particular action.  
 14 While some philosophers today are happy to admit of a sense of "rationalization" so thin that the very presence of a non-rational desire can rationalize acting on it, such a thought is antithetical to Kant.

15 See Schapiro (2009, 2014).

16 This is the strategy adopted by Schapiro (2009, 251).

17 See Schapiro (2009).

dependence on water for survival and the fact that I may have consumed an insufficient amount of it, of which thirst is usually reliable indicator. Likewise, my “pathological”—as opposed to practical—desire to check my e-mails points in the direction of a concrete action, but in the good case if I decide to check my e-mails it is not the pathological urge to revisit the inbox that I deem a reason to do so. To have a pathological urge just means to have an unreasoned, usually unproductive and potentially even destructive, inclination. This is why a pathological urge can be gratified, but not, standardly, rationalize action. I say “standardly” because there is of course a sense in which the itch and the pathological desire to check one’s emails *can* be deemed as reasons to act: this is the case when I decide to succumb to the urge to scratch or scroll because I have no other way to alleviate the discomfort entailed by my failing to do so. But in such a case, the thirst or urge do not “recommend” the end in any way, i. e., they do not reveal anything of value about it, except as a means to alleviate pain, to make the nagging stop. This, however, is not what we standardly take our desires to mean to us and, subsequently, neither is it standardly what we take ourselves to do when we choose to act on them.

The animal-desire interpretation of the IT encounters a second challenge. It is not clear what would afford contingent desires the status of proposals. The intuitive answer, that, like thirst, desires indicate what is necessary for our survival, or wellbeing, is hard to sustain. This may be the case for actual non-rational animals, of course, as long as they are placed in their natural habitat. But ordinary practical agents’ contingent desires are hardly restricted to those that can so much as purport to promote the agent’s “animal wellbeing”: contingent desires may be for ends that it would not make sense for sub-rational animals to have—for example, intellectual and artistic desires or the desire to check one’s e-mails compulsively—and they could be for ends that downright contradict the ostensible aim of self-preservation—for example, a willingness to sacrifice one’s material wellbeing for the sake of an abstract value. Here, the animal desire proponent may reply that while the desire may aim at ends that it would not make sense to attribute to animals, it nevertheless maintains the structure of or continues to function as animal desire. But if the ends of our contingent desires are in no way intrinsically linked to any intelligible notion of our wellbeing (material, physical, biological, etc.), if they are, in other words, from the standpoint of rationality *purely* contingent, it is hard to see why practical reason would acknowledge their authority to make so much as a suggestion concerning what should be done, let alone crown it a reason to do so. There is no reason, after all, empirical or otherwise, to assume the majority or even a large enough share of our contingent desires point to what is necessary for our survival, or even to what would in some way promote our objective well-being. Whether absolutely or in large part, contingent desires are so shap-

ed by culture and society as to render their connection to our objective, rational good extremely tenuous.

### 4.3 The Pro-Tanto Reasons Conception of Desire

A similar problem arises for another proposal for how to conceive of contingent desire; namely, as the source of “pro-tanto reasons.” According to this suggestion, the aim of the IT is to give an account of how different kinds of pro-tanto reasons are turned into all-things-considered reasons, that is, reasons on which we act. A desire gives rise to a pro tanto reason to act, but in order to act on it, I must first incorporate the pro-tanto reasons into a maxim. The advantage of this view over the animal desire view is that it apparently allows us to distinguish between desires that offer pro-tanto reasons and those that do not and that, according to it, no extra-rational reasons are turned into reasons for action. Rather, pro-tanto *reasons* for action are turned into all-things-considered reasons for action. As to modification of the incorporation thesis, the solution is wanting, for it would still leave it unclear how contingent desires, as such, can come to constitute pro-tanto reasons for us.<sup>18</sup> Alternatively, if desires are nothing but pro-tanto reasons, then we need an explanation of what justifies this equivalence between contingent desire and a pro-tanto reason, along the lines of our objections to the animal-desire model, and the idea that a sensible, contingent desire may be deemed, or, as in this case, just is a pro-tanto reason to act.

Importantly, however, none of this is to suggest that explanations of action that refer to “desires” are inappropriate in a Kantian framework. Rather, my aim is to show that the available interpretations of desire in the Kantian framework in general and in particular in the context of the IT readings cannot render such a practice intelligible. For on the IT understanding of what “desires” are, we could at best think of desires as proposing objects about which agents could ask themselves whether, *independently* of their desires, they have reasons to pursue them.<sup>19</sup>

---

<sup>18</sup> See Engstrom (2009) for an account of this general shape.

<sup>19</sup> The worry raised here echoes similar concerns that have been raised with regard to Kant’s theoretical philosophy and in the philosophy of perception more broadly. John McDowell (2013, 256) argues against the idea that if the exercise of sensibility is understood as independent of the exercise of rational capacities, it cannot make objects available to cognition in a way that renders intelligible our capacity to make empirical judgments about the world based on our perceptual experience. If one is to attribute an epistemic role to sensory experience—one that attributes to experience the capacity to warrant or justify belief—one must therefore characterize it as a capacity whose exercises are *essentially* able to stand in rational relations. This specifically precludes the

## 4.4 The Problem of Psychic Alienation

A more general problem with the IT conceptions of desire is the kind of alienated relationship it posits between an agent's desires and her rational capacities. One interpreter who does criticize the IT, Janelle DeWitt, focuses on the psychic alienation that it implies. Referring to the assumptions underlying the IT, De Witt writes as follows:

Though this “distinct natures” view is able to give an explanation for how reason can generate a form of agential control (via the incorporation thesis), it unfortunately seems to come at the high price of a unified psychology. Because the subject's two natures are *entirely* distinct, reason can *only* exercise control as an alien force external to his lower nature. The result of this, however, is a deep rupture between fundamental parts of the subject's psychology—of who and what he is—because he can identify with only one of his two natures at any given time (DeWitt 2018, 68).

Dewitt's dissent is rare and insightful: the IT proponents cannot avoid the threat of saddling Kant with a conceptually dubious and psychologically unattractive dualism.

Nevertheless, despite its merits, DeWitt's own solution does not, I think, fully take the measure of her own concerns. She proposes that we can avoid attributing to Kant “a fractured sense of what it is to be human,” by showing how in his system a person is able to “take on the perspective of his rational *and* animal natures at the same time” (DeWitt 2018, 68). This, however, raises the question of how to characterize this third, “human” perspective, as DeWitt aptly names it, from which one is supposed to adopt both perspectives, of one's rational and animal natures, at the same time. If adopting both perspectives at one and the same time is a matter of having both sets of separate concerns simultaneously, we seem to get ourselves back to square one. Perhaps, then, it is a matter of adopting a third perspective—one that allows us to evaluate the two others. But what kind of perspective or stance is adequate to reflect upon and decide between the two original perspectives? If it is to genuinely constitute a different perspective, it cannot itself be either a rational perspective or an animal one. But then what perspective is it? And how does it allow the agent to adjudicate between the claims of reason and sensibility? To quote Korsgaard's objection to a similar model, “if the person identifies

---

possibility that experiences understood in merely causal terms can play any epistemic or justificatory role in our thinking. See also Boyle (2016) for related thoughts about the application of this form of worry about our perceptual capacities to the case of human desire and action.

neither with reason nor with passion, then how—on what principle—can she possibly choose between them?” (2009, 123). DeWitt leaves the question open.

Objections to Kant’s practical philosophy along these lines are not new. Most interpreters have learned either to resign themselves or defend Kant’s account as a prescient characterization of our cursed fate. However, there is another worrying consequence of this account of contingent desires, to which interpreters have not yet paid sufficient attention.

## 4.5 The Problem of Particular-Moral Ends

Every action, moral or not, aims at a particular, material end. For example, in order to generally “keep my promises,” I must return particular debts at particular times. This was clear to Kant. In the second *Critique*, he writes that it is “undeniable that every volition must also have an object and hence a matter” (Kant, *KpV* 5:34).<sup>20</sup> The fact that an action aims at a particular end or has a matter does not make that action morally bad. The moral worth of an action depends on whether or not this matter is “the determining ground and condition” of the action (Kant, *KpV* 5:34). In other words, it is only if one’s action is dedicated *solely* to securing some particular end—pursuing it as if it is worthy in itself—that it loses moral worth.

Kant moreover held that merely formal principles cannot provide matter for particular actions. The source of “matter”—for Kant, the source of particular ends—is sensible desire.<sup>21</sup> It follows that the choice of specific objects of pursuit, in *all* our actions, is determined by what we sensibly desire.<sup>22</sup> We have seen that most IT

---

<sup>20</sup> See also Kant, *R* 6:4–5.

<sup>21</sup> See Kant, *KpV* 5: 33–34, 109.

<sup>22</sup> The assumption that sensible, extra-rational desire is necessary to provide objects for the will is ubiquitous. Even recent interpreters like Carla Bagnoli, Dewitt, and Jessica Tizzard, who explicitly claim to offer non-dualistic accounts of the interaction of our spontaneous and receptive practical capacities, still maintain this hallmark of the IT: the assumption that reason cannot supply the so-called matter of the will—particular possible ends to be pursued—and that extra-rational input is necessary for every action. Bagnoli writes, “Respect [...] explains how moral reasons drive us to action. It does so not by adding to the many incentives that rational agents review in deliberation, but by constraining and ranking such incentives” (2011, 76). In other words, reason does not provide its own ends to pursue but only limits and organizes hierarchically given extra-rational ends. Tizzard means to show how reason can structure sensibility through and through and to this end argues that moral respect is a form of practical sensibility analogous to the forms of intuition, space, and time. However, on her account the agent’s orientation towards particular ends still requires the contribution of given, extra-rational sensible desires. Practical reason requires “sensible

proponents conceive of sensible desire as essentially having its source outside of reason, an object in the empirical world, a physical urge or something akin to animal desire. For this view, reason may constrain and structure such extra-rational matter, but the matter on which reason exerts its influence must come from outside it. However, taking these two points together yields a highly worrying consequence: namely, that even if an agent recognizes that her duty requires generally keeping her promises, for her to keep her particular promise by performing a particular task, that agent would have to experience, over and above her recognition of the formal demand to keep promises, an extra-rational sensible desire for the concrete ends that constitute a realization of that duty. In other words, if the matter of the will were extra-rational, concrete realizations, even of one's moral commitments, would essentially depend on a supply of extra-rational matter.

This saddles Kant with a claim that is deeply at odds with our moral self-understanding. When an agent recognizes their moral obligation, i. e., when they realize they *should* or *ought* to do something, they not only recognize what they should do in general terms—for example, that they should keep promises or help friends in need—they are also able to recognize specific concrete courses of action as the morally right thing to do. In particular, an agent can direct herself towards fulfilling a particular end in the course of performing her duty without needing to have, in addition to her capacity to recognize duties in their general form (keep promises or help friends), an extra-rational, *non-moral* desire for that particular end. I need not “feel like” transferring some money, independently of being aware of my duty to repay a debt, in order to do the morally right thing *in concreto*. As the Kantian ethicist would presumably want to put it: I should be able to do the right thing completely independently of whatever it is I anyway happen to feel like doing.

---

deliverances to identify and seek out particular objects” (Tizzard 2018, 632). DeWitt’s account comes closest, certainly in its ambitions to the one proposed in this paper. She argues that it is the “extra-rational” conception of feeling and desire that is the source of the problem with the IT account and proposes an account of feelings and desires as partially constituted by reason in order to resolve the problem. However, on Dewitt’s account, the performance of action depends on extra-rational matter. Actions, according to her, require “the adoption of an empirical object or state of affairs as the matter of my will,” but the sort of feelings reason engenders, so-called “formal” feelings “cannot provide this type of matter.” The higher formal feelings “can only motivate me to adopt general principles of good-willing, such as honesty, and this activity is confined to the will itself.” For matter, “the lower feelings, as material feelings” must be consulted (Dewitt 2014, 55–56). This implies that in order for me to exercise my duties in concrete circumstances I must have non-rational feelings for the ends that might realize my formal ends. This is the element of her account that I worry belies her ambitions..

In fact, although it is often overlooked, the idea that a particular end can be pursued *from* recognition of one's duty features explicitly in Kant. In the Preface to the First Edition of the *Religion*, Kant writes,

although morality does not on its own behalf need a presentation of a purpose which would have to precede determination of the will, yet it may well have a necessary reference to such a purpose, namely not as the basis of the maxims adopted in conformity with those laws but as these maxims' necessary consequences.—For without any reference to a purpose, no determination of the will can take place in a human being at all [...]. Without this purpose, a power of choice that, to a projected action, adds in thought no either subjective or objectively determined object (that it has or should have), [hence being] instructed indeed as to *how* it is to operate but not *toward what*, cannot be adequate to itself (Kant, *R* 6:4)

Crucial to note here is that Kant not only asserts that every action, moral or otherwise, must have a “necessary reference” to a particular purpose, but that the pursuit of material ends can itself come in one of two varieties: The material object of an action can be merely “subjectively” determined, as when an agent pursues an end only because it would contribute to their own satisfaction, but the particular, material end can also be “objectively” determined, as when that end is an effect of a morally required action. Crucially, ends that are both material and objective emerge as the adopted formal principles' “necessary consequences.” In other words, we pursue them solely *because* they are the effects of those actions that constitute the concrete realization of our moral duties. They are not extra-rational but proceed from the agent's recognition of what her duties demand of her, right there and then. Thus, the IT's assumption that the object of desire is extra-rational is not only philosophically unattractive; it renders it impossible to find room for the central phenomenon Kant articulates here: that of pursuing particular ends through the recognition of duty.

## 5 Rethinking Incorporation

To get Kant's actual account of incorporation into view, we can begin by asking: if what gets incorporated is not or is not grounded in an extra-rational desire, what is it? The first thing to note is that in the original incorporation requirement passage Kant does not speak of *Begierde* or *Neigung*, i. e., what we standardly translate as “desire,” or “inclination,” respectively, at all; the object of incorporation is rather what he calls a *Triebfeder*. Here it is again, for reference:

The power of choice [*Willkür*] has the quite peculiar characteristic that it cannot be determined to an action by any incentive [*Triebfeder*] except *insofar as the human being has admitted the incentive [Triebfeder] into his maxim [nur sofern der Mensch sie in seine Maxime auf-*

*genommen hat*] (has made this a universal rule for himself, according to which he wills to conduct himself) (Kant, *R* 6:23–24).

What is incorporated into the singular, most general *Gesinnung*-constituting maxim is a “*Triebfeder*.”<sup>23</sup> What is a *Triebfeder*? Kant defines it as the “subjective ground of desire” (Kant, *G* 4:427) or “subjective determining ground of the will” (Kant, *KpV* 5:72). But what is it that “determines the will”? *Triebfeder* is most commonly translated as “incentive.” This is somewhat misleading, however, because, as other interpreters have noted, Kant does not have in mind an object or circumstance that incites the agent to act, like a tax incentive.<sup>24</sup>

Instead, interpreters have proposed that a *Triebfeder* is best thought of as something like “the dynamic or conative factor in willing” (Beck 1996, 216; cf. 90) or “the driving or propelling force from which action or effort springs” (Herrera 2000, 395), as opposed to the cognitive or representational factor. Nevertheless, this reading misses the implication of a key part of Kant’s account in the second *Critique*: namely, the idea that when the agent acts from consciousness of the moral law, the moral law *itself* is the *Triebfeder* of the human will (Engstrom 2010, 93).<sup>25</sup> Thus, Stephen Engstrom writes, in order for the moral law to move you, the objective principle must become the subjective principle and thereby become itself *Triebfeder*: In other words, the *Triebfeder* is not a conative “something” that motivates the agent to adopt a certain cognitive principle, it is rather a principle of action—in the good case, the moral law—as *actually motivating* the agent (Engstrom 2010, 93). It is, in other words, the principle insofar as it is adopted as a maxim.

---

23 This is not of course to suggest that Allison or his followers’ error was one of translation. My aim is to draw attention to the fact that by attending to the original term we might find reason to deem its translation by the word “desire,” something that perhaps was meant to align Kant’s position with contemporary philosophical ones, inadequate. In other words, the question is not whether the object of incorporation is a desire or a *Triebfeder* but rather what a *Triebfeder* is.

24 This is the sense in which Korsgaard, for example, interprets it: “An incentive is a motivationally loaded representation of an object” (2009, 104–5). Earlier, however, she claims more simply that “incentive” is just Kant’s “own language” for “desire or impulse” (Korsgaard 2009, 72). The translation of *Triebfeder* by incentive is persuasively objected to by Larry Herrera (2000, 395, n. 1) and Engstrom (2010, 91–93). I leave the term untranslated.

25 Indeed, while in the *Groundwork*, Kant uses the term *Triebfeder* exclusively with reference to non-moral activity (see Kant, *G* 4: 400, 404, 407, 411, 412, 419, 425, 427, 431, 439, 441, 444, 449, and 461), from the second *Critique* onward, he employs the notion also to express what it is for reason to be practical in a finite rational being: *the moral law itself* must serve as the will’s *Triebfeder* (*KpV* 5:72).

Engstrom was only concerned with the moral *Triebfeder*, but I propose to extend the suggestion that the moral *Triebfeder* is the moral principle as adopted by the agent to the interpretation of the role of both kinds of *Triebfeder*. In so doing, we can interpret the distinction between the moral and non-moral *Triebfedern* against the background of Kant's distinction between moral and non-moral *principles* of action. Consequently, whereas proponents of the IT distinguish non-moral motivation from moral motivation by pointing to non-moral desires' extra-rational source, I suggest that we should understand the contrast between non-moral *Triebfedern* and moral *Triebfedern* as the contrast between acting on *material* (instrumental) principles and acting on *formal* (moral) ones. Accordingly, to admit non-moral *Triebfedern* into one's "supreme maxim" is to admit instrumental principles of action into one's will.

Thinking of moral and non-moral *Triebfedern* along the lines of the distinction between formal and material *principles* is consonant with an oft-neglected strand in Kant's writings prior to the *Religion*. Throughout his critical works, Kant contrasts the kind of action that is performed from consciousness of the laws of reason, or "formal" principles, not with action from empirical or animal causes or sources, but with the sort of action that is directed at particular ends in consciousness of instrumental or "material" *principles* (Kant, *G* 4:428).

In the opening section of the second *Critique*, Kant claims that the difference between the higher and lower faculties of desire is just the difference between the kinds of principle upon which a person can act: material or formal. What is more, Kant makes explicit that both material and formal principles and so both the higher (moral) and lower (sensible) faculties of desire are both "always a product of reason" (*KpV* 5:20). The difference lies in *how* they determine the agent to act: material principles specify what an agent must do to attain a particular effect. Formal, moral principles specify not the effect of an action, but its form, its suitability for the giving of universal law (see Kant, *KpV* 5:21). What is crucial is that whether the agent exercises the lower or higher faculties of desire—namely, whether she acts on a non-moral material principle or a moral-formal one—the agent is acting on a representation of principle and therefore exercising a *rational* capacity to act.

Further evidence for the claim that the contrast between the moral and non-moral action is not one between reason and nature but between formal and material ends and principles is provided by Kant's extensive list of possible non-moral ends an agent might pursue. Kant explicitly states that the representations of ends an agent can pursue in accordance with material principles can have their origin "in the understanding" or even in "reason" (*Ibid.*) and includes things like reading "an instructive book," enjoying "a fine speech" or "an intellectual conversation," experiencing the "joy" in benefitting "a poor man" (*Ibid.*) even the "consciousness of our strength of soul in overcoming obstacles opposed to our plans," or "cultivat-

ing our talents of spirit” (Ibid.) All of these make sense only as the ends of a rational being, and moreover, could, in principle, be pursued out of recognition of moral duty. But if pursued for their own sake, Kant says, as though they were good in themselves, the actions aiming at such ends would lack moral worth.

Kant’s central point here is that *no* particular end is good or bad in itself. Therefore, even what starts as a commitment undertaken for moral reasons can ossify into an action performed simply for the sake of a particular end that is pursued merely for its own sake, in blindness to the requirements of duty. A singular focus on a particular purpose (e.g., scholarship) or the fetishizing of a particular kind of action for its own sake (e.g., political activism) may both be as opposed to the morally good conduct of an upright agent as would be the prioritization of the simplest forms of bodily pleasure.

But if having “sensible” desire is not a matter of having ends from a particular source (say, sensible nature) but is instead about the way an end motivates, i.e., about the sort of principle an agent may act upon, why call it “sensible”? Is the price of the interpretation proposed here ignoring the significance that Kant attributes to our being sensible beings? Not at all, though it does mean that we might want to modify our understanding of what “sensible,” in sensible desire, really means. According to Kant, it is a condition of choosing immorally that we have a sensible nature. A will that is not “affected by needs and sensible motives,” Kant says, would be a “holy will” that “would not be capable of any maxim conflicting with the moral law” (*KpV* 5:32). But we should be careful not to import our own ideas about what “sensibility” means into our interpretation of Kant. The form of human practical sensibility for Kant is feeling—in particular, pleasure and displeasure. Kant holds that pleasure and displeasure are forms of awareness of the agreement or disagreement of a representation with a power of the mind. In pleasure and displeasure, we sense whether the objects of our representation harm or benefit us *qua* possessors of different capacities. Some of these capacities are bodily, but not all of them are. Our fundamental mental capacities are rational: the capacity to reason theoretically and practically. A rational, sensible subject thus registers *sensibly*, i.e., through feeling, how objects around them stand to benefit or harm them *qua* *possessor of rational capacities*, no less than she registers how objects benefit or harm her *qua* *possessor of physical powers* and therefore insofar as she is not materially self-sufficient. Thus, for Kant, pleasure is to be had in anything that can benefit us, promote our goals, support our activity, and of which we can be aware. Kant’s general account of feeling thus makes intelligible the claim that humans can take pleasure, as in Kant’s examples, in distinctly rational ends like intellectual conversation or book reading or the cultivation of one’s spiritual talents. Sensibility, in the practical realm, is therefore not a faculty whose principles are merely empirical causal laws or the laws of animal sentience

and desire. Practical sensibility can reveal our attitude toward particular ends, whatever they may be. As a consequence, we see we must expand our conception of sensibility—what we are practically receptive to—so that we include the whole gamut of practical ends that practically rational beings plausibly pursue.<sup>26</sup>

According to the account I have presented here, what *morality* constrains, is thus not our extra-rational “nature,” but the rational capacity to value—and so act for—*particular* ends. The most important consequence of this account is that it allows us to recognize that the primary locus of the opposition to morality is the capacity to act on principles which guide our action towards particular ends. In other words, the opposition to morality is *internal* to practical reason.

Having clarified the distinction between the two kinds of *Triebfedern* as the distinction between the sort of principles an agent may come to act upon, we can finally identify the proper object of incorporation in Kant’s account. In light of the association of the sensible and moral *Triebfedern* with material-particular and formal-moral principles, respectively, we see that when Kant speaks of *Triebfedern* to be incorporated into the general maxim of one’s will, we need not take him to mean the incorporation of extra-rational desires characterized by their non-rational form of motivation. Rather, we ought to recognize that claim as referring to the admission, into our highest maxim, of two possible kinds of principles upon which we might act: formal-moral principles and those aiming at particular concrete ends.

To see what all of this means we must consider another key element in Kant’s account of the incorporation of *Triebfedern* into the *Gesinnung*-constituting maxim, a feature that finds little room in the interpretation given by the IT’s proponents and one which our interpretation puts us in a unique position to understand. A few pages after the “incorporation requirement passage” Kant claims that human beings admit “both incentives [*Triebfedern*] for determining the will” into the supreme maxim, and that the distinction between good and evil character lies not in which *Triebfeder* is admitted but in the “subordination (in the maxim’s form): which of the two he makes the condition of the other” (Kant, *R* 6:36). Thus, Kant claims that, without exception, *both* moral and non-moral *Triebfedern* are taken up into the *Gesinnung*-constituting maxim and are ordered by it.

On the standard reading, this could only mean that the agent is saddled with two sources of motivation, which are always in potential conflict with one another, with one source simply given priority over the other. In particular, it means that

---

26 Cf. Kant, *KpV* 5:9, n.; *MS* 6:211; *KU* 5:220, 204, 230–233, and 278; and *Anth* 7:321. For a good exposition of this aspect of Kant’s account, see Newton (2017). I discuss Kant’s faculty of feeling at greater length in Berg (2021).

the good agent commits herself to pursue her contingent, non-rational desires only as long as these do not conflict with her commitment to the moral law. She is psychologically divided, and her moral activity depends on extra-moral matter to be supplied to her by her desires. By contrast, by interpreting the distinction between *Triebfedern* as the distinction between practical *principles*, we can recognize an altogether different motivation for admitting both kinds of *Triebfedern*: both kinds of principles are admitted because both are necessary for practical agency. No action is merely formal; all action is always directed at a particular purpose, even moral action.

We can also recognize an altogether more attractive notion of subordination, one which points to the hierarchical organization of one's practical priorities: the moral agent always acts on both moral and non-moral principles. Morally good actions stem from a commitment to never let the pursuit of a particular purpose from which one expects gratification, even one that is derived from moral commitments, take priority over commitments to formal and unconditional principles of action; that is, over her commitment to the moral law. Morally bad actions trace to the agent's subordinating the commitment to act from consciousness of the moral law to particular, material goals. This need not involve a stable commitment to the satisfaction of one's every whim; it suffices that an agent makes their moral conduct dependent on the absence of particular circumstances that would, if they were to obtain, constitute a sufficient reason to excuse oneself from doing what she ought to. Thus, according to Kant's incorporation requirement, our freedom to act on desires is grounded not, in the first place, in our responsibility for the evaluation of concrete desires, but in our set of practical priorities: we are responsible for the way in which we organize our practical life as a whole.

Furthermore, this interpretation of the objects of incorporation, *Triebfedern*, as the capacities to act from the recognition of two different kinds of principle, makes room for the idea that particular ends can be pursued not merely insofar as they do not oppose the moral law but also from genuine recognition of one's moral obligations. As we have seen, the IT rendered problematic the possibility of a particular end being pursued from recognition of duty. Conceiving of incorporation in terms of the ordering of principles of action allows us to accommodate the intuition that acting to realize one's duty does not require matter to be supplied by an extra-rational sensible desire, and that an agent's recognition of what they should do can find expression *in concreto*. For according to the interpretation advanced, the revolution that rationality brings to our practical receptive faculties makes it intelligible that the agent could "sensibly desire" an end; that is, pursue an end according to an instrumental principle, which is objectively determined, and thus whose value is moral. An agent, in other words, can act for particular purposes that themselves reflect particular morally worthy ends to pursue.

This means reason can direct itself to particular purposes of its own accord. Reason's form can provide for its own matter: its own objectively determined ends.

## 6 The Freedom to Do What We Want

According to Kant, to freely do something because I want to do it, is not a matter of rationally evaluating a particular desire, with which I happen to find myself contingently saddled, and in light of such evaluation deeming it a reason to act. It is a matter, rather, of pursuing a particular end, in which I do or expect to find pleasure. In pursuing such a particular end, I exercise a will whose fundamental organizing principle and thus shape, its *Gesinnung*, I have determined and for which I can therefore be held responsible. In this self-organization I determine the priority relation between the pursuit of material ends according to instrumental principles and moral ends according to formal ones. It follows that in pursuing a particular end I may be pursuing an end derived from a formal, moral principle or I may pursue the satisfaction of a simple bodily need. In a sense, I may do either from the recognition of duty or simply for its own sake (for, as I have shown above, even a particular end derived from general moral principle can in turn be pursued as if it is good in itself). No transition from having the desire to the having of a principle is required since to have a particular desire is to have a particular end and a principle of instrumental action aiming at its realization. This principle does not specify that one ought to pursue it and therefore that one has reason to pursue it, it specifies rather how one would go about achieving it. Whatever the moral worth of the action, I am free in acting on such desires because, having admitted both kinds of *Triebfedern*, both kinds of principles or both kinds of action-determination, into my highest maxim, I have freely determined how to subordinate the one to the other.

This interpretation in turn leaves us with an important question, which has been obscured by the proponents of the IT: how should we understand the idea that agents freely determine their own practical priorities (i. e., their moral character)? It is clear that the freedom we exercise over the constitution of our character cannot be an instance of an ordinary exercise of agency—for it is not the result, at least not in any obvious way, of an ordinary act of the will—and therefore good moral character is not something we can simply decide to bring about. Our character, in other words, is not another thing that we can pursue. How, then, can we be held responsible for its structure? Whether Kant has the resources to answer *this* puzzle is an important and difficult question, one which I will have to take up elsewhere. I take it to be one of the major virtues of my account, however, that it allows us to see this puzzle clearly.

## Abbreviations

### Kant's Works

In the case of the *Critique of Pure Reason*, I follow the standard practice of referring to the 1781 (A) and 1787 (B) editions. For all other texts, citations are given as abbreviation, volume number, and page number as per the *Akademie Ausgabe* (AA), *Kants Gesammelte Schriften*, edited by the *Königlich Preussische Akademie der Wissenschaften* (29 Vols.; Berlin: de Gruyter, 1900–). Unless otherwise stated, all translations come from *The Cambridge Edition of the Works of Immanuel Kant*, edited by Paul Guyer and Allen Wood (Cambridge: Cambridge University Press, 1992–).

### References

- Allison, Henry E. (1991): *Kant's Theory of Freedom*. Cambridge: Cambridge University Press.
- Bagnoli, Carla (2009): "Review of *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*, by Christine M. Korsgaard". In: *Notre Dame Philosophical Reviews*.
- Bagnoli, Carla (2011): "Emotions and the Categorical Authority of Moral Reason." In: Bagnoli, Carl (Ed.): *Morality and the Emotions*. Oxford: Oxford University Press, 62–81.
- Baron, Marcia (1993a): "Freedom, Frailty, and Impurity." In: *Inquiry* 36. No. 4 (December 1), 431–441.
- Baron, Marcia (1993b): "Henry Allison on Kant's Theory of Freedom." In: *Dialogue: Canadian Philosophical Review/Revue Canadienne de Philosophie* 32. No. 4, 775–781.
- Beck, Lewis White (1996): *A Commentary on Kant's Critique of Practical Reason*. 2<sup>nd</sup> ed. Chicago: University of Chicago Press.
- Berg, Anastasia (2021): "Kant on Moral Respect." In: *Archiv für Geschichte der Philosophie* 103. No. 4, 730–760.
- Bojanowski, Jochen (2007): "Kant und das Problem der Zurechenbarkeit." In: *Zeitschrift Für Philosophische Forschung* 61. No. 2, 207–228.
- Boyle, Matthew (2016): "Additive Theories of Rationality: A Critique." In: *European Journal of Philosophy* 24. No. 3, 527–555.
- DeWitt, Janelle (2014): "Respect for the Moral Law: The Emotional Side of Reason." In: *Philosophy* 89. No. 1, 31–62.
- DeWitt, Janelle (2018): "Feeling and Inclination: Rationalizing the Animal Within." In: Sorenson, Kelly and Williamson, Diane (Eds.): *Kant and the Faculty of Feeling*. Cambridge: Cambridge University Press, 67–87.
- Engstrom, Stephen P. (2009): *The Form of Practical Knowledge: A Study of the Categorical Imperative*. Cambridge: Harvard University Press.
- Engstrom, Stephen P. (2010): "The Triebfeder of Pure Practical Reason." In: Reath, Andrews and Timmermann, Jens (Eds.): *Kant's Critique of Practical Reason*. Cambridge: Cambridge University Press, 90–118.
- Guyer, Paul (2008): *Kant on Freedom, Law, and Happiness*. Cambridge: Cambridge University Press.

- Herman, Barbara (2008): *Moral Literacy*. Cambridge: Harvard University Press.
- Herrera, Larry (2000): "Kant on the Moral Triebfeder." In: *Kant-Studien* 91. No. 4, 395–410.
- Hill, Thomas E. (1991a): *Dignity and Practical Reason in Kant's Moral Theory*. Ithaca: Cornell University Press.
- Hill, Thomas E. (1991b): "Kant's Argument for The Rationality of Moral Conduct." In: Hill, Thomas E.: *Dignity and Practical Reason in Kant's Moral Theory*. Ithaca: Cornell University Press, 77–122.
- Johnson, Robert N. (1998): "Weakness Incorporated." In: *History of Philosophy Quarterly* 15. No. 3, 349–367.
- Kang, Ji-Young (2015): *Die allgemeine Glückseligkeit: Zur systematischen Stellung und Funktionen der Glückseligkeit bei Kant*. Berlin: de Gruyter.
- Korsgaard, Christine M. (1996a): *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.
- Korsgaard, Christine M. (1996b): *The Sources of Normativity*. Cambridge: Cambridge University Press.
- Korsgaard, Christine M. (2009): *Self-Constitution: Agency, Identity, and Integrity*. Oxford: Oxford University Press.
- McCarty, Richard (2008): "Kant's Incorporation Requirement: Freedom and Character in the Empirical World." In: *Canadian Journal of Philosophy* 38. No. 3, 425–451.
- McDowell, John (2013): "Avoiding the Myth of the Given." In: McDowell, John: *Having the World in View: Essays on Kant, Hegel, and Sellars*. Cambridge: Harvard University Press, 256–272.
- Morrisson, Iain (2005): "On Kantian Maxims: A Reconciliation of the Incorporation Thesis and Weakness of the Will." In: *History of Philosophy Quarterly* 22. No. 1, 73–89.
- Newton, Sasha (2017): "Kant on Animal and Human Pleasure." In: *Canadian Journal of Philosophy* 47. No. 4, 518–540.
- O'Neill, Onora (1989): *Constructions of Reason: Explorations of Kant's Practical Philosophy*. Cambridge: Cambridge University Press.
- Papish, Laura (2007): "The Cultivation of Sensibility in Kant's Moral Philosophy." In: *Kantian Review* 12. No. 2, 128–146.
- Peters, Julia (2018): "Kant's Gesinnung." In: *Journal of the History of Philosophy* 56. No. 3, 497–518.
- Prauss, Gerold (1983): *Kant Über Freiheit Als Autonomie*. Frankfurt am Main: Klostermann.
- Quinn, Warren (1994): "Putting Rationality in Its Place." In: Quinn, Warren: *Morality and Action*. Cambridge: Cambridge University Press.
- Rawls, John (2000): *Lectures on the History of Moral Philosophy*. Cambridge: Harvard University Press.
- Schadow, Steffi (2012): *Achtung für das Gesetz: Moral und Motivation bei Kant*. 1<sup>st</sup> ed. Berlin: de Gruyter.
- Schapiro, Tamar (2009): "The Nature of Inclination." In: *Ethics* 119. No. 2, 229–256.
- Schapiro, Tamar (2011): "Foregrounding Desire: A Defense of Kant's Incorporation Thesis." In: *Journal of Ethics* 15. No. 3, 147–167.
- Silber, John R. (1960): "The Ethical Significance of Kant's Religion." In: Kant, Immanuel: *Religion Within the Limits of Reason Alone*. Theodore M. Greene and Hoyt H. Hudson (Trans.). New York: Harper & Row, lxxix–cxxxiv.
- Timmermann, Jens (2003): *Sittengesetz und Freiheit: Untersuchungen zu Immanuel Kants Theorie des freien Willens*. Berlin: de Gruyter.
- Timmons, Mark (1994): "Evil and Imputation in Kant's Ethics." In: *Jahrbuch für Recht und Ethik* 2, 113–141.
- Tizzard, Jessica (2018): "Kant on Space, Time, and Respect for the Moral Law as Analogous Formal Elements of Sensibility." In: *European Journal of Philosophy* 26, 630–646.

Wood, Allen W. (1999): *Kant's Ethical Thought*. Cambridge: Cambridge University Press.

