

Empty, Useless, and Dangerous? Recent Kantian Replies to the Empty Formalism Objection*

Fabian Freyenhagen

Like two heavyweight boxers exchanging punches, but neither landing the knock-out blow, Kantians and Hegelians seem to be in a stand-off on what in contemporary parlance is known as the Empty Formalism Objection. Kant's ethics is charged with being merely formal and thereby failing to provide the kind of specific guidance that any defensible ethical system should have the resources to provide. Hegel is often credited with having formulated this objection in its most incisive way, and a wealth of Kantian responses has been deployed to answer it. In this paper, I take up the objection as it appears in §135R of *Elements of the Philosophy of Right* in order to scrutinise the contemporary debate between the two camps. I propose that there are, in fact, three different, albeit connected objections and examine (what I take to be) the best Kantian replies to them. I will not adjudicate which of these replies is the most accurate interpretation of Kant's texts, nor trace the particular historical context in which Hegel takes up Kant's ethics, nor the way the Empty Formalism Objection fits into Hegel's wider system. This is partly because of constraints of space, and partly because many of the contemporary Kantian replies — for better or for worse — treat the Empty Formalism Objection as a self-standing philosophical problem, irrespective of its historical context or systematic place in Hegel's theory. My limited aim here is to show that, even if one grants — for argument's sake — the legitimacy of such a non-contextual approach, significant difficulties remain.

I. Empty Formalism in the Philosophy of Right

What we now call the Empty Formalism Objection finds, perhaps, its most pithy statement in Hegel's *Philosophy of Right*, about mid-way through the discussion of 'Moralität' (3rd Section, 'The Good and Conscience'). I will concentrate here on this statement, drawing on the rest of the book and Hegel's other works only when necessary.

Still, some minimal textual contextualisation of the Empty Formalism Objection within the *Philosophy of Right* is in order. Hegel's argument is at a point where it has been suggested that the good provides an objective standard by which the will should be guided (see §133). Specifically, the will faces the good as an obligation or duty (*Pflicht*)

which we are meant to obey for its own sake, since the good here is contrasted to all specific determinations of the will which have proved unsuitable guides for genuinely free willing in the preceding discussion and, hence, what remains is only the general abstract essence of goodness that we ought to respect as such. The question then becomes what this means concretely for our conduct (see §134). Here a problem arises: we require specific guidance in the form of detailed, substantial instructions and ends to orientate our wills in concrete situations, but such specific guidance is not contained in the mere idea of duty for duty's sake, and presumably can also not be derived from it (see §135). It is not that we have no idea at all what these instructions or ends could be — presumably drawing on the earlier discussion of abstract right and welfare, Hegel suggests that we should do what is right and work for the happiness of ourselves and others (see §134). Rather, the point is that even such abstract and general guidance as this cannot be derived and is not contained in the idea of duty for duty's sake; never mind the more specific guidance we need. Hence, this idea lacks any content and consists merely in an empty identity, forcing people to turn to their conscience for any specific guidance, which, however, has pitfalls of its own (which Hegel then goes on to discuss; see §§136-140; see also §141).

It is in this context that Hegel, in the remark to §135, turns to his critique of Kant's moral philosophy. He begins the remark by praising Kant: specifically, he credits him with having introduced the idea of autonomy, that is, the idea that morality requires the ability not just to reflect on and choose among one's impulses, inclinations, and desires (what Kant calls 'negative freedom'), but also the ability to be motivated by pure practical reason alone — in short, to act on the requirements of reason for the sake of these requirements alone and thereby to achieve unconditional self-determination in one's willing. This praise is followed by what we now call the Empty Formalism Objection. Kant is faulted for turning his important insight into a one-sided position, absolutising the moral standpoint. Doing so lands one with an empty formalism for the following reasons:

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

contain any principle apart from formal identity and that absence of contradiction already referred to. — The fact that *no property* is present is in itself no more contradictory than is the non-existence of this or that individual people, family, etc., or the complete *absence of human life*. But if it is already established and presupposed that property and human life should exist and be respected, then it is a contradiction to commit theft or murder; [...]. But if a duty is to be willed merely as a duty and not because of its content, it is a *formal identity* which necessarily excludes every content and determination. (PR §135R; translation modified).

This passage contains a whole barrage of criticisms.¹ While I will return to some of the more specific points later, it is useful here to distil the main elements of this passage, and I propose that it boils down to three general claims:

1. *No immanent doctrine of duties*: it is not possible to arrive at a doctrine of duties on the basis of the mere idea of duty for duty's sake or the formal identity of rational willing proposed in the categorical imperative.
2. *No criterion for testing potential duties*: even if candidate duties are provided from the outside, testing for whether there is consistency in rational willing (or for whether they can be willed as a universal law) does not provide a criterion for determining whether or not the candidate duties are genuine duties.
3. *False positives*: immoral acts could successfully pass the test for consistency in rational willing and the maxims involved could be universalised.²

These three criticisms are interconnected. For example, criticism 3 is a way of amplifying criticism 2: testing for consistency of rational willing is not just useless (as criticism 2 has it), but even dangerous.³ Also, as the discussion of these three criticisms will show, they form a kind of dialectic, with one leading to the other by way of certain Kantian replies.

The criticisms are meant to be immanent ones — Hegel is trying to show how the Kantian position is problematic on its own terms and implodes when fully thought through (although it can also be rescued to some extent by being sublated in Hegel's own position). Hence, Kantians will have to recognise themselves in the characterisation offered. While I will suggest later that Hegel might well be accused of presenting an incomplete picture, the basic characterisation that has gone into the three criticisms above seems correct. Specifically, Kantians would have a hard time denying that the moral law, as Kant thinks of it, consists in a merely formal requirement of reason. It is

crucial to Kant's argument that, if there can be a moral law at all, it has to exclude all empirical ends and cannot be a 'material principle', since neither such ends, nor material principles have the required universality and necessity for lawfulness (G 4: 399-402, 419-421; KpV especially 5: 21-28). Hence, if there can be a moral law at all, then it must be formal in the sense of neither relying on empirical ends nor being a material principle. As such, it can only consist in the very idea of lawfulness itself — that is, in demanding that rational willing is consistent with itself and that one's subjective principle of action (one's maxim) can be willed to be a universal law. In short, Hegel latches onto the key move in Kant's argument for the categorical imperative, specifically the formula of universal law.⁴ Moreover, Kantians would also accept that all human action requires specific ends and guidance. Kant repeatedly admits that all actions are directed at ends (G 4: 427; MS, 6: 384f; see also KpV 5: 34) and our (lower-level) maxims involve specific content and specific ends, or at least specific action types done for specific types of ends (such as making false promises as a way to promote one's interests, or committing suicide when one expects more suffering than happiness from living on). Hence, the problem of empty formalism cannot be simply side-stepped by denying the background assumptions which lead to it.

Instead, Kantians have to tackle this problem head on. In what follows, I will discuss (what I take to be) the best contemporary defences against each of the three claims made by Hegel.

II. Kantian reply strategies

II.1 *An immanent doctrine of duties*

It is a controversial issue how, if at all, duties are derived within Kant's moral philosophy and what role, if any, the categorical imperative plays in this. I cannot here do justice to all the different interpretations. Instead, I concentrate on the three most promising avenues: (a) a side-stepping manoeuvre based on ascribing moral realism to Kant; (b) a reply that concedes that an immanent doctrine of duties is not possible, but denies that it was Kant's intention to provide one; and (c) a final, more elaborate defence, according to which there is a sense in which Kant's ethics does contain an immanent doctrine of duties after all.

II.1a

One recent suggestion has it that Kant is a moral realist who takes our moral duties as given, so that he neither needs nor attempts to offer a doctrine or derivation of duties.⁵ On this view, the categorical imperative is only meant to be a useful tool to counteract our tendency to rationalise away the fact that a specific moral duty applies to us,⁶ and its defence in Part III of the *Groundwork* is only concerned with assuring us of the obligatory nature of moral duties for finite rational beings like us.

Insofar as Hegel is commonly read as a moral realist himself,⁷ this interpretation of Kant's ethics would mean that there is actually no disagreement between Hegel and Kant when it comes to the impossibility of offering a doctrine of duties on the basis of the mere idea of duty for duty's sake. Moral realist or not, Hegel would simply have misunderstood Kant, although the Empty Formalism Objection might still be applicable to what we nowadays describe as Kantian constructivism (which is a form of moral anti-realism) as well as to any attempts to derive specific duties from this perspective. Moreover, perhaps Hegel's first criticism could be recast so as to put in doubt the usefulness of the categorical imperative for counteracting our attempts to rationalise away our obligations — this, however, would effectively collapse it into the second criticism (to which I return below in II.2). In this way, the overall thrust of Hegel's first criticism would have been sidestepped or at least diverted.

II.1b

Those who want to keep more distance from both Hegel and moral realism have to offer alternative strategies of reply. A familiar first move is to insist that the *Groundwork* should not be understood to do any more than it claims to: and that is to clarify and ground the highest principle of morality (G 4: 392). In the preface to this work, Kant states that he is not concerned with the application of the moral law (and thereby, one presumes, not concerned with arguing for or deriving specific duties). Such application would be the work of 'practical anthropology', not the 'metaphysics of morals' and even less of the preliminary study of its foundations (G 4: 388; see also 412; MS, 6: 217). In this sense, the examples of duties discussed in the *Groundwork* are just illustrations in the process of clarifying the highest principle of morality, not actually derivations of specific duties (such as the duty not to make false promises as a way to promote one's interests, or the duty to not commit suicide when one expects more suffering than happiness if one carries on living). Kant takes it to be uncontroversial that the duties used in the examples are accepted as 'actual duties' (G 4: 424) — no derivation is necessary to show this. He also accepts as given the customary division into self- and other-regarding duties of either a perfect or imperfect kind (G 4: 421; see also 423f), reserving his right to revise it later (G 4: 421n). The closest Kant comes to offering arguments for specific (ethical) duties is in the second part of the *Metaphysics of Morals*, the Doctrine of Virtue (*Tugendlehre*) (about which more below).

However, this first move merely shifts the problem, since the real question is not so much whether or not Kant attempts to provide or derive an immanent doctrine of duties in this or that book, but whether his moral philosophy has the resources to do so. One response to this question is to concede that if what is at issue is the possibility of deriving a doctrine of duty from the mere idea of duty for duty's sake, then Kant does not attempt it and in any case cannot provide it (nor could anyone else). Still, this concession does not damage Kant's ethics, since he does not attempt to offer an immanent doctrine of duties. In fact, Kant would himself reject such an immanent

doctrine as misconceived — unless we bring in knowledge about human beings, their needs and vulnerabilities, a doctrine of duties for human beings cannot be provided, and that is why such a doctrine and the specific duties we face fall within practical anthropology, not metaphysics of morals. Actually, even the discussion of specific duties in the *Metaphysics of Morals* is merely due to the fact that ‘a metaphysics of morals cannot dispense with principles of application and we shall often have to take as our object the particular *nature* of human beings, which is cognised only by experience, in order to *show* in it what can be inferred from universal moral principles’ (MS, 6: 217; emphasis in the original). In other words, the specific duties discussed in the Doctrine of Virtue are not derived merely from the idea of duty for duty’s sake, but owe their existence to practical anthropology.⁸

If one adopts this response strategy, then the decisive disagreement with Hegel is whether or not his second criticism is correct — that is, whether or not the mere idea of duty for duty’s sake, and thereby the categorical imperative, contains a criterion for testing maxims. The first criticism is correct as far as it goes, but it does not go far at all, since Kant was well aware that this mere idea did not contain a doctrine of duties. Before discussing the second criticism, one additional response to the first requires our attention.

II.1c

A third strategy of reply is similar to the second, but less concessive. It consists in emphasising that some content and even some duties (albeit general) are contained in the mere idea of duty for duty’s sake, and that, although anthropological knowledge and perhaps even sociology are required to derive more specific duties from this content, this idea frames this derivation and, in this sense, Kant’s ethics contains an immanent doctrine of duties. Part of this defence relies on the thought that the very concept of duty only makes sense as applied to finite rational creatures and that, hence, knowledge about human beings (as finite rational creatures) is admissible, at least as long as it is merely descriptive and no normative ideas other than the form of the moral law (or, what comes to the same thing, the mere idea of duty for duty’s sake) is appealed to.⁹ In this sense, even the anthropological (and sociological) knowledge invoked is not really outside material — only other normative ideas, or substantive duties would be. Also, for the most general elements of the doctrine of duties even such knowledge would not be required: we can derive some content from the mere idea of duty for duty’s sake, specifically the objective end of humanity.

To substantiate this response, it is useful to begin by noting how Kant actually proceeds in the Doctrine of Virtue to arrive at the specific duties he puts forward. What is striking is that the formula of universal law, which Hegel seems to have in mind when making his Empty Formalism Objection, does not figure prominently. Instead, the formula of humanity features repeatedly in the reasoning,¹⁰ as does the idea of two obligatory ends, the perfection of oneself and the happiness of others (on which more below). Insofar as humanity as an end in itself is an objective end contained in the idea of

the categorical imperative itself and insofar as the two obligatory ends follow from this idea (at least in conjunction with some anthropological knowledge), Hegel's first claim could be rebutted and it could be shown that there is a doctrine of duties immanent in the mere idea of duty for duty's sake, after all. These are complex issues, so my discussion of them here is going to be a high-altitude sketch only.

Objective and obligatory ends

The key move in the transition to the formula of humanity in the *Groundwork* is the connection between the idea of the moral law and the will of a fully rational being (G 4: 426). Kant extends his analysis of what is involved in rational willing beyond what he has already said about it in Parts I and II, suggesting that all willing is directed towards ends (G 4: 427). However, since we are looking at rational willing as such, the ends in question cannot be empirical ends, from which, as we have already seen, the moral law is supposedly independent. Rather, we need to consider whether there are any ends which 'hold equally for all rational beings' (G 4: 427). Such ends would be what Kant calls 'objective ends'. They are *analytically differentiated* from subjective ends, that is, the particular ends for the sake of which finite subjects actually act, although in practice it can happen that a particular subject makes an objective end into his or her subjective end. Kant's formal principle of morality, the categorical imperative, has only to abstract from subjective (and thereby contingent) ends, but not from objective (and thereby necessary) ends.

In the *Groundwork*, Kant argues that humanity is an objective end. (By humanity he understands the rational nature of human beings, that is, their capacity for rational and autonomous willing).¹¹ The argument for this, in a nutshell, is the following (see, especially, G 4: 427f, 435, 437f, 440): If there is anything of absolute worth (that is, unconditional goodness), then it is an objective end, an end that holds for all rational beings. The only thing of absolute worth is a good will. (This premise is taken from Part I of the *Groundwork* (see 4:393f); and Hegel would presumably grant its truth for argument's sake within his immanent critique of Kant's ethics). The capacity for autonomous willing is required for a good will, since a good will is a will which aims at morality for moral reasons, and such reasons cannot be based on inclinations, since the latter are — by (their) nature — orientated towards self-love, which can differ from and even conflict with morality (G Part I and 4: 444; KpV 5: 22-27, 33ff). In fact, the capacity for autonomy is required for a good will, not as an independent pre-condition, but constitutively (G 4: 444) — willing autonomously consists of willing what is absolutely good, and *vice versa*. Consequently, the capacity for autonomy is an objective end, and, as the idea of humanity consists in this capacity, humanity is an objective end.¹² As such, we have to treat humanity in all our action always as an end in itself, not merely as a means — and that is exactly what the Formula of Humanity requires.

The problem facing Hegel is that Kant seems here to be doing exactly what Hegel denies is possible: to derive some content from the mere idea of good willing (which, in

the case of human beings, is the same as the mere idea of duty for duty's sake — as Hegel concedes, at least for argument's sake in PR (§§133-5). Admittedly, what is derived here is not yet some specific duty but 'only' an objective end. Still, this objective end plays an important role in generating specific ethical duties in the Doctrine of Virtue — for example, the arguments for the duty not to kill or maim oneself, the duty to oneself not to lie, and the duty against false humility rest centrally and explicitly on the claim that undertaking such acts would be to treat humanity in one's own person as a mere means (MS, 6: 423, 429f, 434-436). Similarly, duties to others, such as the duty of respect, are based on the same consideration as applied to others: we have a duty of respect because otherwise we would be permitted to treat others as mere means and disregard the dignity with which their humanity endows them (MS, 6: 462, 466).

Turning to the two obligatory ends of one's own perfection and the happiness of others, Kant's argument here too is probably best constructed as relying heavily on the formula of humanity, at least in the case of the first obligatory end. To say that one's own perfection is an obligatory end is to say that we have a duty to develop and protect our capacities. Specifically, this means that we have a duty to preserve our body, to develop our moral capacity to act, to cultivate our natural powers and faculties as well as moral predispositions. The argument in support of the thesis that this is an obligatory end is an extension of the argument that humanity is an end in itself. Thus, if our capacity for autonomy is an objective end because it is required for there to be absolute worth, then whatever makes this autonomy possible is also an objective end. Protecting and developing our capacities is required in order to sustain and actualise autonomy, and, hence, our own perfection is obligatory as an end. Admittedly, this judgement relies on some knowledge of human beings — that they are born in such a way that their capacities need to be developed, that they have vulnerabilities and needs which make it necessary to protect their capacities, and so on. Still, the knowledge required here is of a very general nature and, at least in part, just an extension of the very idea of a finite, embodied rational agent.

Why is the perfection of others not also an obligatory end? Well, Kant thinks that we cannot directly promote the perfection of others — for perfection requires that agents set their own ends in the light of their understanding of duty, and no one can do such end-setting for someone else (MS, 6: 386). Still, we can and should help others to be *in the position* to perfect themselves — and this thought leads us to the second obligatory end, the happiness of others: since we are sensuous creatures, we would have difficulties working towards our own perfection if we were unhappy, and to avoid the latter we require the help of others.

If the happiness of others is an obligatory end for us, then this means that we have the duty to promote the ends of others, at least as long as these ends are morally permissible. The argument in Kant's text builds more on the idea of the formula of universal law and thereby presents an even more direct counterexample to Hegel.¹³ If I take the maxim of self-love and test it by way of the categorical imperative, then I realise

that universalising the maxim requires me to include the happiness of others among my ends (MS, 6: 393; see also KpV 5: 34). I can only rationally will to promote my own happiness, if I also accept the happiness of others as a reason for action.¹⁴ One might think that it follows from this that we have a general duty to promote everybody's happiness, including our own. However, for Kant it cannot be a duty to do something that we are going to do anyway, independently of our having a duty to do it. Consequently, promoting one's own happiness cannot be a duty, since it is an end every human being has 'by virtue of the impulses of his nature' (MS, 6: 385f; see also 387, 451). Thus, we naturally tend to promote our own happiness anyway, but this is permissible only if we are willing to promote other people's happiness, too.¹⁵ Hence, the happiness of others is an obligatory end for us and so is the duty to promote their ends (subject to the constraint that the ends they pursue to attain happiness are morally permissible ends). Again, it is just certain general facts about human nature (such as that they naturally seek their own happiness) play a role in this argument.

Kant uses these two obligatory ends to derive some of our more specific ethical duties in the Doctrine of Virtue, particularly the duty to search one's conscience, the duties to develop and increase one's natural and moral perfection, the duties of beneficence, and the duty to oneself regarding non-rational beings (see especially MS, 6: 441, 443, 445-447, 451-453). To take the latter as an example, Kant argues that we have a duty not to destroy natural beauty wantonly and not to be cruel to animals because doing either of these acts weakens or undercuts our work towards moral perfection. Specifically, it blunts dispositions that are important for morality, such as, presumably, our sensitivity to the suffering of sensuous creatures, which human beings are, after all, as well.¹⁶

Insofar as both obligatory ends are clearly linked to the categorical imperative, this imperative frames the derivation of duties, while it is not by itself sufficient for the purpose of deriving specific duties. Instead, we need to draw on the specific nature of human beings and their circumstances — that they have needs which they require the help of others to satisfy, that they interact in various ways, that their capacities need to develop, that they are vulnerable, that they naturally seek happiness (the satisfaction of their desires), etc. Nonetheless, no purported or candidate duties are brought to the derivation from the outside, nor any other normative ideas beyond what is contained in the concepts of duty for duty's sake, good will, and rational beings with wills. In this sense, one could argue that Kant's ethics contains an immanent doctrine of duties. While based on a formal principle of morality, it is not devoid of moral content, but includes a way to derive specific duties by reference to this principle and the general facts about human beings relevant to its application — or, at least, this is what Kantians would argue.¹⁷

There are various ways in which Hegel or Hegelians might respond to these arguments, such as doubting the Kantian thesis that moral reasons cannot be based on inclinations.¹⁸ Instead of following up the wider debate between Kantians and Hegelians

that such a challenge to a key Kantian premise would entail, I want to take up a more direct Hegelian rejoinder. Even if Hegelians were to accept for argument's sake that some ends can be derived from fairly minimal premises within Kant's ethics, they would argue that the ends in questions are too general and vague to get to the kind of specific duties and guidance that would be required for actual ethical practice.¹⁹ What it is to respect another person will vary enormously from one age to the next and from one society to the next — it will be different in twelfth century Japanese society and twenty-first century Sweden, and it will be different in contemporary Cairo and contemporary Kaliningrad. Similarly, the duty to help those in need will vary enormously with context, notably with institutional context.²⁰ Is there a state-run and coercively enforced redistributive mechanism in place to help those in need? If so, does the duty to help others extend merely to compliance with this mechanism or go beyond it? Are there perhaps two duties here (one to comply with just institutions and one of charity)? And what if there is no such mechanism or one that is problematic in various ways? Will not the duty to help those in need change quite fundamentally in this different context? And if so, in what way are we still talking of one and the same duty that is merely applied differently in different contexts?

In reply, Kantians might begin by conceding some element of contextuality, although they would probably maintain that this is really only an application issue.²¹ It might well be true that what it is to treat someone with respect and as end in itself will vary from one age to the next and from one society to another, but the general duty of respect is invariant, and any differences are just about specifying its detailed implications or the practical rules of implementing it. Also, Kantians would present the following counter-argument to the Hegelians. If one makes moral norms and duties too context-specific, one opens the floodgates to relativism and, perhaps more importantly for a Hegelian, one cannot demonstrate that the specific practices and institutions are rational. To avoid this, we have to be able to show how the specific moral duties fit with the demands of reason — and the categorical imperative is meant to show just this. It brings out the choice-worthy aspects of maxims (and practices), and without it, we fall into either dogmatism or relativism — neither of which are suitable for Hegel's aim of *rational* reconciliation. As Ameriks puts it:

the crucial point here is that his superior concreteness can be a real advantage only if the content is *correct*. At this point, rather than lapsing into social relativism, as he is too often still charged with having done, Hegel tends to fall back on bounds set by pure morality: our commitment to *Sittlichkeit* is to be restrained by a Kantian respect for man as an end in himself. In this way Hegel can properly criticise ancient slave culture as well as the similarly objectionable *Sittlichkeit* of some modern societies, but only at the price of relying on what seem to be the very abstract principles he meant to transcend. (Ameriks 2000: 314)

In sum, Kantians are happy enough to concede that we need what Herman has called ‘middle theory’ — which includes not just general anthropological knowledge but also knowledge about the particular social context — in order to arrive at specific duties.²² Still, if we gave up on framing this derivation of duties by the categorical imperative (and the objective and obligatory ends), then we would give up on rational grounding altogether.

At this point, Hegelians respond by arguing that the categorical imperative does precious little work in the actual derivation of duties. Instead, when Kant in his time and Kantians nowadays say that they derive duties, they in fact just fall back on the ethical life that surrounds them (or the remnants thereof).²³ Rather than holding social practices to an independent critical standard, they, perhaps unwittingly, rely on contextually given norms and turn into defenders or, at most, reformers of the *status quo* (as they see it). Presumably, this reassertion of the basic objection is partly due to the fact that Hegelians think that the categorical imperative does not actually provide a workable criterion with which to test maxims or candidate duties. And in this way the debate about the first criticism by Hegel merges with the debate about his second criticism, to which I will now turn.

II.2 *A criterion for testing the candidate duties*

Kant suggests at one point that an important function of the categorical imperative is to counteract our tendency to rationalise our past behaviour and future actions in such a way as to let us get away with breaches of what we — at least in the abstract — already know is morally wrong (G 4: 405). The categorical imperative can do this counteracting work because we can use it to test our maxims and, presumably, also any purported duty in order to determine whether or not they are genuinely morally permissible or required. It would thus be a serious blow if Hegel were right that the categorical imperative is unsuitable for this purpose because it does not actually contain a criterion that rules in genuine maxims (and candidate duties) and rules out the problematic ones. Moreover, it would be a blow to Kantians, whichever of the three reply strategies sketched above they adopt.

Hegel’s criticism is specifically directed against the formula of universal law, that is, the categorical imperative to ‘[a]ct only on that maxim through which you can at the same time will that it should become a universal law’ (G 4: 421). In effect, Hegel says that this formula does not go beyond the mere idea of duty as the lack of contradiction. There is truth in this assertion insofar as Kant would presumably admit that the formula of universal law is about whether or not one can will *without contradiction* that one’s maxim become universal law. Specifically, for Kant, there are two kinds of cases of contradiction at stake: (a) one cannot will the maxim to be a universal law because its universalisation cannot be conceptualised (that is, there is what is nowadays called a ‘contradiction in conception’); or (b) the maxim is conceivable as universal law, but its universalisation can

still not be willed by rational beings (that is, there is ‘a contradiction in willing’).²⁴ A strict duty is one where universalising a maxim that was contrary to it generates a contradiction in conception, while a wide duty is one where universalising a maxim contrary to it would avoid this problem but instead generate a contradiction in willing.²⁵

Hegel does not discuss contradiction in willing, perhaps because he thinks that if the contradiction in conception test fails, the whole enterprise is doomed — after all, strict duties would then not be accounted for. What he does discuss (and repeatedly so) is a variant of the false promising example from the *Critique of Practical Reason*, the Deposit Example (KpV 5: 27; see PS ¶437; PR §135R). In this example, Kant imagines that someone has placed a deposit with another person but has subsequently died and left no record of it. For the person with the deposit, this would be an occasion to act on the maxim ‘to increase my wealth by every safe means’, but this person might test whether or not this is a permissible maxim by asking whether or not the maxim could be willed as a universal law. Kant says it could not be so willed, since in a world where it was a universal law that people did not return unrecorded deposits when it is safe to keep them, ‘there would be no deposits at all’ (KpV 5: 27) or at least no unrecorded ones. One might object here to specific aspects of the example or to Kant’s reasoning, but Hegel presses a deeper point: without the premise that property is a morally permissible institution and theft morally problematic, the mere fact that someone does not return a deposit to his or her owners (or to the heirs) cannot be shown to be morally problematic, for there is nothing with which this act or the underlying maxim could come into contradiction. To recall the passage from the *Philosophy of Right*, not denying the right to existence of a particular group, nor rejecting such a right for humanity as whole, nor abolishing the institution of property leads to a contradiction, unless we have already assumed what would need to be demonstrated: that we have a duty to preserve or at least not endanger human life and a duty to respect people’s property. Without these presuppositions, neither a contradiction in conception, nor a contradiction in willing could be generated.

In reply, Kantians could argue that the Hegelian worry is misconceived. Specifically, Hegel seems to assume that we would first need to show that promising or property are justifiable institutions in order to show that acting contrary to them is morally problematic, but this is beside the point because what is at issue is a contradiction in conception, not a contradiction in willing (where the rational acceptability of institutions might play a role). The contradiction in the case of the maxim of false promise-making is that the maxim both relies on the existence of an institution (here promise-making) and, if adopted as a universal law, makes its existence impossible. This suffices to strike the maxim down, independently of the merits or demerits of the institution in question. A justification of the institution is neither assumed nor necessary, since the agent, in invoking the institution, already accepts it in a sufficient sense: he or she cannot both want to use it and undermine the possibility of its existence.²⁶ Admittedly, such a contradiction does not show that the institution in question (say

promise-making or private property) is morally permissible. Still, showing this is necessary in order to defeat the maxim.

However, Hegelians could respond by arguing that it is misleading to evaluate an institution and the maxim making use of it wholly independently of each other — such a strict division of labour is likely to generate (moral) blind spots. Basically, the contradiction in conception test might be good at showing that free-riding is impermissible, but even if it were good at showing this, it could produce the wrong result, since not all free-riding or non-compliance with an institution is obviously impermissible. Consider the following example: Under the so-called ‘Jim Crow laws’, non-whites were prohibited in many Southern US states from sitting in the front seats of buses long into the 1950s. This requirement would seem to run counter to the moral egalitarianism of Kant’s ethics — even deserving to be resisted by permissible means (such as civil disobedience). Indeed, there were black citizens, such as Jackie Robinson in the 1940s or Rosa Parks in the 1950s, who decided to oppose this practice by such means, and Kantians would surely want to vindicate their behaviour as not just permissible but even admirable. Yet, their behaviour would fail the categorical imperative on at least some descriptions of it. Thus, if Jackie Robinson or Rosa Parks had made it their maxim to always sit in the front when taking the bus, this maxim would have generated a contradiction in conception — we cannot conceive of a world in which everyone acts on this maxim, since it is in the nature of buses that they can only accommodate a finite number of passengers in the front seats. Testing for a contradiction in conception seems to produce here a false negative, that is, it excludes something as a violation of a strict duty that is no such violation.²⁷

Kantians would tend to say that there is something wrong *not* with the test provided by the categorical imperative, but rather with the *formulation* of the maxim — for example, the maxim should not be described as always sitting in the front when taking the bus, but as acting against unjust institutions by legally permissible means. Still, in order for this reply not to be merely *ad hoc*, Kantians would have to provide clear limits and guidance on maxim formulation, and to do so without smuggling in substantive ethical content. In this way, the discussion has to move on to what Kantians can reply to Hegel’s third criticism.

II.3 False Positives (and Negatives)

There is more to Hegel’s third criticism than is betrayed by the brief remark in the *Philosophy of Right*. When we turn to the *Phenomenology of Spirit*, we can see Hegel making a similar point: ‘Just because the criterion [of rational consistency] is a tautology, and indifferent to the content, one content is just as acceptable to it as its opposite’ (§430). More specifically, it turns out that the objection here is, at least in part, related to what is sometimes called the *Act-Description Problem*. As Hegel argues in the context of the discussion of the Deposit Example (§437; see KpV 5: 27 and above), nothing prevents me from stopping to describe the object as deposit or the rightful property of someone

else and viewing it, instead, as ‘my rightful property’ or ‘an unowned object’ — or so Hegel claims:

If I should keep for myself what is entrusted to me, then according to the principle I follow in testing laws, which is a tautology, I am not in the least guilty of contradiction; for then I no longer look upon it as the property of someone else: to hold on to something which I do not regard as belonging to someone else is perfectly consistent. Alteration of the *point of view* is not contradiction; for what we are concerned with is not the point of view, but the object and content, which ought not to be contradictory. Just as I can — as I do when I give something away — alter the view that it is my property into the view that it belongs to someone else, without becoming guilty of contradiction, so I can equally pursue the reverse course. (PS ¶437)

The crucial point in this argument is that changing my description of the object in question is not contradictory and, hence, not excluded by the testing procedure. In this sense, it seems again as if the criterion provided by rational consistency (or willing my maxim as universal law without contradiction) is not sufficient as a criterion and might not exclude immoral acts and maxims.

One way to capture Hegel’s point somewhat differently is to say that Kant faces a dilemma when it comes to the input into the test that the categorical imperative supposedly provides: if the moral law consists only in the requirement that we should not act other than on a maxim that could be a universal law, then any maxim can be made to conform to the moral law if suitably described (the Act-Description Problem);²⁸ if, on the other hand, there are constraints on what kinds of maxim (or act descriptions) are the right ones to be tested, then Kant’s ethics is not formal, but contains substantive constraints after all.²⁹

It might help to consider an example. A man we might call ‘Eric Cantona’ is leaving the football pitch in anger and while walking towards the changing room a fan from the opposite team insults him.³⁰ Let us grant that kicking the fan in response is morally impermissible and see whether the Kantian testing procedure tracks this. What if Cantona’s maxim were the following: ‘Whenever I, a French Player with the number 7 of the football Club, Manchester United, am insulted by the Crystal Palace fan Matthew Simmons on the 25th January 1995 in Crystal Palace’s football stadium, I will kick the amassed matter at space x?’ Here, it seems very difficult, if not impossible, to claim that there is a contradiction in conception or in willing — for I can consistently conceive and perhaps even rationally will a world where everyone who is the French Player with the number 7 of the football Club Manchester United, is permitted to kick the amassed matter at space x when they are insulted by the Crystal Palace fan Matthew Simmons on the 25th January 1995 in Crystal Palace’s football stadium. As there is only going to be —

at most — a one-off application of this maxim, even if it were universally adopted by everyone, it seems that it does not undermine any important institutions or in other ways run counter to many of the usual things Kantians say would make passing the universalisability test impossible. Moreover, the description is, at once, specific and vague, and in such a way that none of the important moral elements which one would normally flag up come into view.

There are a number of things Kantians say in response to the Act-Description Problem, not all of which are clearly compatible with each other. I will here concentrate on the most important and generally accepted line of response. Testing maxims for conformity with the Categorical Imperative is first and foremost a test to assess maxims for action from the deliberative, first personal perspective (rather than a third-personal assessment of rightness of action). The first-personal perspective brings certain constraints on maxim-formulation with it. The maxims need to be at least minimally rational, where this means that they must conform to the hypothetical imperatives in question (rules of skill (technical) and counsels of prudence (pragmatic)) — otherwise they fail as principles of willing and action, quite independently of their moral permissibility. The maxims to be tested also have to be those that the agent sincerely believes he or she actually acts on. Finally, we are only interested in action-types, not specific acts — that is, we are interested in, for example, deception-for-personal gain, not the specific time of day this takes place. These three constraints are not substantive in a way that takes anything away from the formalism of Kant's ethics, since they are part of the very idea of maxims as subjective principles of action and the particular first-personal use to which we put the categorical imperative — if we want to know whether the behaviour we intend to undertake and the specific reasons why we undertake it are permissible, we had better make sure that we test the relevant maxim, that is, one which sincerely describes what we are up to, which involves the required means, and which is sufficiently general to allow for whatever specific variation is required to then put it into practice. These constraints are not independent moral constraints, but constitutive norms for the very process of testing our maxims for moral permissibility.³¹

Let us return to the Eric Cantona Example and look specifically at the second constraint mentioned: sincerity. If the maxim is formulated in such a specific way as my example above was — ‘Whenever I, a French Player with the number 7 of the football Club Manchester United, am insulted by the Crystal Palace fan Matthew Simmons on the 25th January 1995 in Crystal Palace's football stadium, I will kick the amassed matter at space x’ — then it must be part of the intention that it stands and falls with this specificity. Thus, Cantona was not permitted to kick the Crystal Palace fan on his adopted maxim, if he had just found out that he was in fact not French; or if the fan he kicked actually sat on seat y (not x); or if he was mistaken about the date he was acting on. Yet, it would seem that Cantona would have proceeded to kick the fan, even if he had discovered these facts. Hence, it seems that the maxim he ascribed to himself is not the maxim he did, in fact, (set to) act upon, as he would have admitted, if pressed. The

actual sincere maxim would have depended not so much on specific details as the imagined one did, and Cantona could have known this. Yet, this means that the imagined one was insincere and as such unsuitable for testing it as a maxim for action that he was actually going to adopt. In this way, a lot of maxims, which look like they would pass the test of the Categorical Imperative because they are too specific or idiosyncratic, would be disallowed even before this test is applied.³² Consequently, a significant first step towards overcoming the Act-Description problem is to recognise that the maxims to be tested have to meet the requirement that they are sincerely held and instrumentally rational (as well as stated in terms of action-types).³³

Moreover, Kantians would also disagree with Hegel's claim that altering one's point of view (say, from accepting that an object is Smith's rightful property to claiming it as my own rightful property) involves no contradiction in rational willing (PS ¶437). After all, at least if I intentionally alter my point of view, then I also act on a maxim — say the maxim, 'Whenever it is to my advantage, I will re-describe things, actions or maxims so as to make my behaviour and willing appear to be compatible with the categorical imperative'. One could argue that this maxim would itself not pass the categorical imperative, and to act on it would display insincerity not in the practically self-defeating sense of not really testing the maxim I genuinely want to act on, but in the direct moral sense of aiming at deception.

Still, there is something unsatisfying about the Kantian reply strategy, for the problem is not just that the maxim tested was not sincerely held or might have involved an impermissible deception strategy, but that — even if it had been sincerely held in all its specificity and was not adopted to deceive — the action it proposed is morally problematic: assuming that the deposit really ought to be returned and that the fan really ought not to be kicked, then this part of it has been lost track of in the response considered so far to the Act-Description Problem. The categorical imperative would only be suitable to pick up on the second-order maxim of deception, but absent such a maxim or insincerity, the presence of either too much specificity or too much vagueness (or both) in the description means that the testing process would not pick up on the morally salient features and would not yield the result of excluding what is morally impermissible and allowing what is permissible. Similarly, if we think back to the case of civil disobedience against racial segregation, it seems as if the maxim, 'I will always sit in the front seats when taking the bus' could be sincerely held (and is instrumentally rational and formulated in terms of action-types) but would still come out as impermissible on that formulation. Also, the more general substitute — acting against unjust practices by permissible means — is too abstract to guide us in specific cases — Rosa Parks also needs to know whether or not sitting in the front of the bus is one such permissible means. In other words, the categorical imperative fails to exclude impermissible and to pass permissible actions, unless they are sincerely and correctly described in the first place, but doing the latter turns out to be very difficult indeed. Often quite natural

descriptions will be such that they yield false positives or negatives if tested for moral permissibility by way of the categorical imperative.

Herman suggests a reply to worries such as these when she accepts that human beings not only need to know the categorical imperative but also need to be educated in rules of moral salience to realise that they are in a moral situation and identify the features that require moral deliberation.³⁴ This represents a break with what Kant says in his more optimistic moments about the categorical imperative: namely, it is a break with his claim that everyone could know (or, at least, be assured of) what is morally required of them just in virtue of the compass the categorical imperative provides (G 4: 404; see also KpV 5: 27, 35-37).³⁵ Still, there are other passages where Kant acknowledges that certain sensibilities and the capacity of moral judgement are required as background conditions of moral agency (MS, 6: 399-402), and in some places he even admits that we require experience in sharpening these sensibilities and capacity to judge in order to recognise what we have to do in specific situations and be motivated to act accordingly (see, most notably, G 4: 389).

However, if we need rules of moral salience, or moral sensibilities to correctly describe the situation and to make correct use of the categorical imperative, then one might again question whether the categorical imperative still does any useful work. If I have described the situation correctly so as to bring the morally salient features into view and I am generally disposed to act morally, then why does this not suffice by itself to settle what I have to do? Why do I still need to consult the categorical imperative? In fact, many of the paradigm examples of morally right behaviour — such as the actions of those who helped Jews to escape persecution and death during the Nazi reign — did not involve any appeal to the categorical imperative or much reflection at all, but simply the perception of the need of others and the danger they were in.³⁶ On the other hand, knowledge of and appeal to the categorical imperative did not stop people from carrying out some of the most paradigmatic evil acts — Eichmann famously claimed to have always tried to live according to the requirements of the categorical imperative and quoted it in his defence.³⁷

In reply, it could be argued that one reason why testing my maxim by way of the categorical imperative might still be required, or, at least, useful is that even if I have the morally salient features clearly in view, I might not yet know how to adjudicate between them, and the categorical imperative test could help with this task, at least once the situation is described correctly. Moreover, as we have mentioned already, testing our maxims by way of the categorical imperative is an important counterweight to our tendency to rationalise away what we know is morally required. Also, the mere fact that the testing procedure has been misunderstood and misapplied (as, Kantians would claim, happened in Eichmann's case) does not yet show that it is always useless or dangerous to employ it. And Kantians do not claim that we would, in all cases, actually have to go through the procedure, but merely that we could and that doing so would normally yield the right results — although there is no iron-cast guarantee that it will do so, since the

proper use of the procedure requires us to have the right kind of practical orientation and sensibilities to guide this process.

Still, even if it were true that the categorical imperative is not completely redundant, there are other worries to which this revised view of Kantian ethics gives rise. Firstly, if Kant's optimism is indeed unfounded and has to be given up in the way Herman suggests, the issue of culpability becomes much trickier, since now luck plays a much bigger role in whether or not the individual is fully capable of describing the situation and using the categorical imperative correctly. There are related repercussions, such as the fact that the alleged advantage of the categorical imperative over the counsels of prudence — that the requirements of the former, but not of the latter, are always 'quite easily seen and without hesitation by the most common understanding' as well as 'within everyone's power at all times' to satisfy — disappears, and with it also goes the reasonableness of the moral demands to 'command compliance from everyone, and indeed the most exact compliance' (KpV 5: 36).³⁸

Secondly, one other important lesson from this discussion is that the earlier Kantian argument that the categorical imperative (or the objective end of humanity) could and should be used to test for the permissibility of social practices is now shown to rest on presuppositions that call the viability of its employment for this purpose into question. How could the categorical imperative provide a critical standard of this sort, if it actually turned out that using it correctly relied on having been brought up to recognise morally salient features? Presumably, if social practices are morally problematic, the moral upbringing and education in that society may well be problematic too. If so, the purported advantage of abstract morality over concrete ethical life — that the former, but not the latter, offers us the resources to evaluate our social practices on the basis of a standard independent from them — evaporates, at least for those within these practices.³⁹ And if the social world contains some non-corrupted normative resources and upbringing after all, then no appeal to context-transcendent standards will be required.

In sum, something of Kant's enlightenment optimism — the optimism that we can all use our own reason unassisted, whatever our upbringing or social context — is required to keep the Kantian and Hegelian positions apart, but one could reasonably disagree as to the viability of this optimism, and even Kantians seem to have given up on it. To close with the boxing analogy with which I started, it seems as if the result is a split decision points victory for the Hegelians.

Fabian Freyenhagen
University of Essex
f.frey@essex.ac.uk

Notes

* This paper substantially overlaps with my, ‘The Empty Formalism Objection Revisited: §135R and recent Kantian responses’, in T. Brooks (ed.) (2011), *Hegel’s Philosophy of Right: Essays on Ethics, Politics, and Law*. Oxford: Blackwell. For comments and criticisms on earlier drafts, my thanks go to audiences at talks in Newcastle, Stirling, and Dublin, as well as at the Hegel Society of Great Britain and UK Kant Society 2010 Joint Conference in Oxford. Special thanks are due to Thom Brooks, Rowan Cruft, Katerina Deligiorgi, Wayne Martin, David McNeill, Bob Stern, Lea Ypi, and the fellow members of the Cambridge *Forschungskolloquium* (Manuel Dries, Martin Eichler, Raymond Geuss, Michael Hampe, Richard Raatzsch, Jörg Schaub, and Christian Skirke).

¹ In fact, Hegel goes on to point to the *Phenomenology of Spirit* for a discussion of further antinomies that the Kantian position gives rise to (PR §135R, referring to PS ¶¶596-631).

² One might ask here by what means we are to ascertain that the maxims are immoral, given that they purportedly pass the categorical imperative. One way to argue here is that Hegel (rightly) takes Kant to be trying to vindicate common sense morality, so that it would be a problem, if a maxim that is immoral by the lights of common sense would pass the categorical imperative (or if one that was seen as permissible failed to pass it).

³ One could read Hegel’s passage as entailing the stronger claim that the categorical imperative contains *no criterion whatsoever*, not merely that it contains no criterion for testing potential duties. While the text is not conclusive, one downside of this stronger reading is that the third criticism would then make little sense — if there is no criterion whatsoever, then there is no test and, hence, nothing would pass the categorical imperative, not even immoral acts.

⁴ The formula of universal law states: ‘act only in accordance with that maxim through which you can at the same time will that it become a universal law’ (G 4: 421; see also 402, 437; KpV 5: 30).

⁵ See, for example, Stern (2009: §III.1), with further references.

⁶ Kant mentions at one point that the categorical imperative has such a counteracting function (see G 4: 405). I take this up below (in II.2).

⁷ See, for example, Knowles (2002: 215), with reference to PR §140R; see also Wood (1990) and Stern (2007).

⁸ In fact, some commentators think of the *Metaphysics of Morals* as something more like Kant’s practical anthropology and not really a treatise in the metaphysics of morals, which would deal mainly or even exclusively with *a priori* concepts and justification (see, for example, Timmermann 2007: Appendix F).

⁹ See, for example, Herman (1993: 122). One possible objection to this strategy, which I will not be able to discuss here, is to doubt that the distinction drawn between descriptive and normative content is defensible, especially when anthropological and sociological knowledge is counted among the former.

¹⁰ The formula of humanity states: ‘Act in such a way that you always treat humanity, whether in your own person or in the person of any other, never simply as a means, but always at the same time as an end’ (G 4: 429).

¹¹ See, for example, G 4: 439, where Kant directly equates ‘humanity [*Menschheit*]’ with ‘rational nature [*vernünftige Natur*]’. Thus, by ‘humanity’ Kant means something besides the human biological species. For him, humanity consists in the capacity to reason and to be

moved to action by pure practical reason alone. It is a normative ideal, not a mere descriptive category.

¹² For example, Kant writes that, '[a]utonomy is therefore the ground of the dignity of human nature and of every rational nature' (G 4: 436).

¹³ Still, the argument could also be run by way of the formula of humanity — presumably, to treat others as ends in themselves includes respecting and promoting their (permissible) ends.

¹⁴ A similar argument might be at the back of Hegel's mind when he claims that the welfare of all is a value for us (see PR §125; see also Knowles 2002: 187-189; Knowles does not mention this parallel with Kant). On the sense of rational willing at play here, see note 26 below.

¹⁵ Although we do not have a direct duty to promote our own happiness, Kant admits that we have an indirect duty to promote it. If we are unhappy or lack means (say, when we are extremely poor), then this may make it difficult or impossible to act morally. Hence, we have an indirect duty to avoid pain, adversity, and want, since we have a moral duty to preserve our moral agency and integrity (MS, 6: 388; see also G 4: 399). This is only indirectly a duty to our own happiness, since directly it is a duty to preserve our 'moral integrity' (MS, 6: 388).

¹⁶ It is here not my intention to suggest that this argument is successful or plausible, but merely to highlight that the obligatory end of one's own perfection plays a recognisable role in Kant's and Kantian reasoning.

¹⁷ See, in particular, Herman (1993 and 2007).

¹⁸ Hegel is often read to attack the central Kantian idea that morally worthy action requires that we do not act from inclination (see, for example, Knowles 2002: 178-184, with reference to PR §§123-124; see also Rawls 2000: 333, 335). Herman offers a reply that emphasises the fact that Kantian autonomy might require only that my desires and inclinations are shaped by moral considerations, not that I act without or against them (2007: Ch. 1).

¹⁹ See MacIntyre (1998: 96f).

²⁰ The issue of (institutional) context-dependence is at least part of the point of Hegel's discussion of 'Love thy neighbour as thyself' in the *Phenomenology of Spirit*, ¶425. See also MacIntyre's discussion of how circumstances change moral concepts (1998: Ch. 1, especially 1f) and alter what character traits are counted as virtues or vices (*ibid.*: 206).

²¹ See, for example, Herman (2007: especially Ch. 2); see also Herman (1993: Ch. 4).

²² On middle theory, see Herman (1993: 233-240). It is 'the translation of a formal conception of value into terms suitable for the particular contexts of human action and deliberation' (*ibid.*: 240).

²³ See MacIntyre (1998: 191, 197, 207f).

²⁴ See, for example, O'Neill (1989); for the textual basis underpinning this distinction, see G 4: 424. One common way of thinking about the categorical imperative is to ask whether or not we could conceive and rationally will a world in which one's maxim is a natural law governing this world and determining human nature (see, for example, Rawls 2000: 168f). This approach takes its lead from a variant of the universal law formula, namely, the law of nature formula which states that we should 'act as if the maxim of your action were to become by your will a universal law of nature' (G 4: 421). I will adopt this practice in the discussion which follows.

²⁵ A comment on ‘rational willing’ might be in order here. What Kant must mean by our inability to will the universalisation of a maxim is not something that one simply cannot actually will, since there are things that one can actually will that are morally problematic (think of Hare’s example of the imaginary Nazi who actually wills himself to be killed in the case in which he turns out to be a Jew, or of someone who is in such dire straits that he or she would agree to anything in order to be fed and housed). Rather, the contradiction in willing test is about what a *fully rational being* would be able to consent to. For this reason and in this sense, I speak of ‘rational willing’ here. This should also bring out that rational here is more than taking the required means to further one’s long-term self-interest — one can be rational in this (weaker) sense and not accept a moral duty to help others in need that goes beyond strict mutual advantage. A fully rational being, unlike us finite ones, would not be motivated by self-interest, but only by what pure practical reason required. To mark the difference from the weaker sense of rationality, some Kantians speak of ‘reasonable’ to denote the stronger one (see Rawls 2000: 164f).

²⁶ The exact nature of the contradiction is debated among Kantians, with the two leading interpretations being that a contradiction in conception is about (a) logical contradiction or (b) practical contradiction (acting in such a way as to undermine the very possibility of achieving the end one sets out to achieve by so acting). Fortunately, we need not enter into this debate here, since the response sketched in the main text is open on both interpretations. For further discussion, see Korsgaard (1996: Ch. 3); Herman (1993: 136–143).

²⁷ There is another objection here, although it is not clear whether it is Hegel’s or simply Hegelian (see also Henrich 1994: especially 102f). It is arguably implausible to think that the most important reason why physical abuse or torture is wrong is that we cannot universalise a maxim that included abuse and torture as an end or means. In fact, it is also implausible to think that the only or even the most important reason why physical abuse and torture are wrong is that they undermine our moral agency. If this is true, then one could argue that contradictions in conception or willing are not an adequate test for the (im)permissibility of maxims, since they do not reveal the right reasons why acting on the maxim is wrong. Especially on a Kantian picture, it is also those reasons which we want to know about and then act on. To Hegel’s credit, he seems to acknowledge that inflicting bodily harm is bad in itself (not just in virtue of resting on non-universalisable maxims or undermining our moral agency; see PR §48R). Admittedly, if those who ascribe moral realism to Kant are correct, then he could accept this too (although running the categorical imperative test would still not provide us with knowledge of the right reasons why inflicting bodily harm is wrong).

²⁸ See also MacIntyre (1998: 197f); and Knowles (2002: 205f).

²⁹ Could Kant fall back here on the objective end of humanity and the two obligatory ends to provide guidance that is substantive but nonetheless compatible with his formalist ethics? As we see in the main text, Kantians tend to appeal, at least in the first instance, to other considerations. Still, as we also see in the main text later on in this section, these ends do play a role in a fuller Kantian reply — specifically, Herman’s rules of moral salience embody them.

³⁰ As they say at the end of film trailers, ‘any resemblance to persons living or dead is purely coincidental’.

³¹ Admittedly, we might not test our maxims at all, but this would be a moral failing in at

least some instances and one which has little to do directly with the Act-Description-Problem. Also, we might test the false maxims (for example, we might be deceived about our own motives). Kant cannot exclude this. For him, the best we can do is to describe the actions and motives as sincerely and accurately as possible. The point in this context is that this requirement of sincerity already weeds out maxims of the sort that could yield false negatives or positives, and, hence, is an important resource in responding to the Act-Description-Problem — or so the Kantians argue.

³² To repeat, the maxims in question would be ruled out even before testing them with the help of the categorical imperative, not because they conflict with substantive moral norms imported from the outside, but because they are incompatible with the constitutive requirements of acting and wanting to test the maxim one intends to act on (as well as the first personal perspective involved in this). See also Herman (1993: 75).

³³ While I have concentrated on false positives, Kantians would argue that the three listed requirements of maxim-construction would also rule out the possibility that the categorical imperative generates false negatives (such as that it is impermissible for Rosa Parks to engage in civil disobedience since her maxim — always sit in the front on bus journeys — generates a contradiction in conception).

³⁴ (Herman 1993: Ch. 4). See especially the following concession: ‘An agent who came to the CI procedure with no knowledge of the moral characteristics of actions would be very unlikely to describe his action in a morally appropriate way. Kant’s moral agents are not morally naive. In the examples Kant gives of the employment of the CI procedure (G[, 4:] 422-423), the agents know the features of their proposed actions that raise moral questions *before* they use the CI to determine their permissibility. It is because they already realise that the actions they want to carry out are morally questionable that they test their permissibility. It is hard to see how any system of moral judgement which assessed maxims of action could work with morally naive or ignorant agents’ (Herman 1993: 75).

³⁵ Herman is aware that her proposal ‘does not preserve the idea of a simple technique for maxim assessment’, but suggests that this ‘loss (if it is a loss) will be more than made up if the CI procedure [that is, testing maxims by way of the categorical imperative] can be made to work’ (*ibid.*: 131).

³⁶ See Halter (1998).

³⁷ See Arendt (1994: 135f).

³⁸ To claim that the purported advantage disappears is not to say that there are no other contrasts between the categorical imperative and counsels of prudence that might remain.

³⁹ Herman suggests that the choice-worthy aspects of maxims to which the categorical imperative points us — such as the objective end of humanity — provide critical tools with which we can evaluate rules of moral salience from the outside (from a practice-transcendent vantage point) (1993: Ch. 4, especially 92). However, this is an unconvincing suggestion: if we need sensibility and rules of moral salience to make use of the abstract content purportedly generated by the idea of the moral law, then this abstract content cannot provide an independent critical check on these sensibilities and rules.

Bibliography

- Ameriks, K. (2000), 'The Hegelian Critique of Kantian Morality', in *Kant and the Fate of Autonomy*. Cambridge: Cambridge University Press, 309-337. Reprinted from Moen, M. & den Ouden, B. (eds) (1987), *New Essays on Kant*. New York: Peter Lang.
- Arendt, H. (1994), *Eichmann in Jerusalem: A Report on the Banality of Evil*. New York: Penguin [first published 1964].
- Halter, M. (1998), *Stories of Deliverance: Speaking with Men and Women who Rescued Jews from the Holocaust*, trans. M. Bernard. Chicago: Open Court.
- Hegel, G. W. F. [1807], *Phänomenologie des Geistes* [abbreviated as 'PS']. Reprinted in Hegel, G. W. F. (1986), *Werke* (Volume 3). Frankfurt am Main: Suhrkamp. Translated as (1977) *Phenomenology of Spirit* by A. V. Miller. Oxford: Oxford University Press.
- Hegel, G. W. F. [1821], *Grundlinien der Philosophie des Rechts* [abbreviated as 'PR']. Reprinted in Hegel, G. W. F. (1986), *Werke* (Volume 7). Frankfurt am Main: Suhrkamp. Translated as (1991) *Elements of the Philosophy of Rights* by H. B. Nisbet. Cambridge: Cambridge University Press.
- Henrich, D. (1994), *The Unity of Reason: Essays on Kant's Philosophy*. Cambridge, MA: Harvard University Press.
- Herman, B. (1993), *The Practice of Moral Judgement*. Cambridge, MA: Harvard University Press.
- Herman, B. (2007), *Moral Literacy*. Cambridge, MA: Harvard University Press.
- Kant, I. [1785], *Grundlegung der Metaphysik der Sitten* [abbreviated as 'G']. Akademie edition 4: 395-463. Translated as (1996), 'Groundwork of the Metaphysics of Morals' in *Practical Philosophy*, trans. and ed. M. J. Gregor. Cambridge: Cambridge University Press, pp. 41-108.
- Kant, I. [1787], *Kritik der praktischen Vernunft* [abbreviated as 'KpV'], 5: 1-163. Translated as (1996), 'Critique of Practical Reason' in *Practical Philosophy*, trans. and ed. M. J. Gregor. Cambridge: Cambridge University Press, pp. 137-271.
- Kant, I. [1793], *Metaphysik der Sitten* [abbreviated as 'MS'], 6: 203-493. Translated as (1996), 'The Metaphysics of Morals' in *Practical Philosophy*, trans. and ed. M. J. Gregor. Cambridge: Cambridge University Press, pp. 362-603.
- Korsgaard, C. M. (1996), *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.
- Knowles, D. (2002), *Hegel and the Philosophy of Right*. London: Routledge.
- MacIntyre, A. (1998), *A Short History of Ethics: A History of Moral Philosophy from the Homeric Age to the Twentieth Century* (second edition). London: Routledge.
- O'Neill, O. (1989), *Constructions of Reason*. Cambridge: Cambridge University Press.
- Rawls, J. (2000), *Lectures on the History of Moral Philosophy*, ed. B. Herman. Cambridge, MA: Harvard University Press.
- Stern, R. (2007), 'Freedom, Self-Legislation and Morality in Kant and Hegel:

Constructivist vs. Realist Accounts' in E. Hammer (ed.), *German Idealism: Contemporary Perspectives*. London: Routledge.

Stern, R. (2009), 'The Autonomy of Morality and the Morality of Autonomy' in *Journal of Moral Philosophy* 6.3: 395-415.

Timmermann, J. (2007), *Kant's Groundwork of the Metaphysics of Morals. A Commentary*. Cambridge: Cambridge University Press.

Wood, A. (1990), *Hegel's Ethical Thought*. Cambridge: Cambridge University Press, pp. 127-173.