



ARTICLE

Individual Maxim Tokens, not Abstract Maxim Types

Samuel Kahn

Indiana University, Indianapolis, IN, USA
Email: kahnsa@iupui.edu

Abstract

I argue that Kant’s Categorical Imperative should be applied to individual maxim tokens rather than abstract maxim types. The article is divided into five sections. In the first, I explain my thesis. In the second, I show that my thesis disagrees with Rawls. In the third, I argue for my thesis on the basis of the wording of the Categorical Imperative and on the basis of considerations about autonomy. In the fourth, I argue for my thesis on the basis of considerations about the ‘ought implies can’ principle. In the fifth, I provide a summary of the main argumentative moves and also explain some of the philosophical advantages of my thesis.

Keywords: Kant’s ethics; Categorical Imperative; maxims; general duties; ought implies can; Rawls; Formula of Universal Law; Formula of a Law of Nature; universalization formulations; autonomy

In this article I argue for the following thesis:

TOKEN

Kant’s Categorical Imperative should be applied to individual maxim tokens.

In so doing I argue against the following thesis:

TYPE

Kant’s Categorical Imperative should be applied to abstract maxim types.

The article has five sections. In section 1, I explain TOKEN and TYPE in more detail, and I elucidate the difference between them by appeal to two other philosophical debates: (a) the debate between deontologists and utilitarians and (b) the debate between act utilitarians and rule utilitarians. In section 2, I show that Rawls subscribes to TYPE, thus showing that I am not committing a straw person fallacy. In section 3, I appeal to the wording of the Categorical Imperative (CI) and to considerations about autonomy to motivate TOKEN. In section 4, I argue for TOKEN on the basis of Kant’s commitment to ‘ought implies can’. In section 5, I conclude with a summary and I gesture toward two advantages of TOKEN: (i) it allows for a middle ground between moral relativism and moral absolutism; and (ii) it gives a plausible account of general duties.

I. Explaining my thesis in more detail

There are three aspects of TOKEN and TYPE that require explanation: (I) Kant's CI; (II) the difference between individual maxim tokens and abstract maxim types; and (III) what it means to say that the CI 'should be applied' to one of these rather than the other. It is in explaining (III) that I appeal to the other philosophical debates enumerated in the introduction.

Kant's CI is his attempt to articulate the Supreme Law of Morality as it applies to imperfectly rational beings. What distinguishes the CI from the Supreme Law of Morality is that the former, as an imperative, includes the concept of constraint (G, 4: 412.26-413.15).¹ Constraint is needed because imperfectly rational beings do not do that which the Supreme Law of Morality requires merely because it is required (4: 413.15-18; see also Johnson and Cureton 2022: section 2).

Kant formulates the CI in different ways in different places and for different purposes. For example, the following four formulations of the CI, all from part II of the *Groundwork to the Metaphysics of Morals*, are often referred to as the universal law formulation (FUL), the law of nature formulation (FLN), the humanity formulation (FH), and the realm of ends formulation (FRE), respectively:

1. **FUL** Act only according to that maxim through which you at the same time can will that it become a universal law (G, 4: 421.7-8, emphasis omitted).
2. **FLN** Act thusly, as if the maxim of your action should become a universal law of nature through your will (G, 4: 421.18-20, emphasis omitted).
3. **FH** Act thusly, that you use the humanity in your person, as well as in the person of anyone else, always at the same time as end, never merely as means (G, 4: 429.10-12, emphasis omitted).
4. **FRE** All maxims of one's own lawgiving should agree with a possible realm of ends as a realm of nature (G, 4: 436.24-26).

There is debate about how to interpret these formulations (Korsgaard 1996: chapters 3 and 4); about how many (other) formulations there are (Baker 1988); about whether any or all of these formulations are equivalent (O'Neill 1989: chapter 7; Wood 1999: chapters 3 and 4); and about whether Kant regarded them as such (Pogge 1998; Wood 2006). However, all of these debates are downstream from the debate in this article. I am not saying that these other debates are unimportant. Rather, I am saying that I do not need to take a stand on or discuss them for current purposes.

A maxim is a subjective principle of volition (G, 4: 420.36-421.30). In other words, a maxim is a willed principle by which an agent governs her actions. Here are three maxims that frequently show up in discussions of Kant's and Kantian ethics:

LPM I will make a lying promise in order to get some ready money.

NBM I will never help anyone.

EM I will privilege self-interest over the moral law.

The LPM (lying promise maxim) and the NBM (non-beneficence maxim) are taken from Kant's examples in part II of the *Groundwork to the Metaphysics of Morals*. The EM (evil maxim) is taken from part I of the *Religion within the Boundaries of Mere Reason*.

There is debate about the general properties of maxims (Kitcher 2003); about the generality/specificity required for a principle to count as a maxim (as opposed to a specific intention) (Bittner 1974; Gressis 2010: sections 3, 4, and 5; O'Neill 2013: chapter 1); about the difference between maxims and laws (Gressis 2010: section 2); about whether agents always act on maxims (Nyholm 2017); and about the proper form of a maxim (Allison 1990: 85-94). Some might not accept all (or any) of the three examples just given as genuine maxims. However, these debates, as with the debates mentioned above about the CI, are downstream from the debate in this article. And, as before, in saying that, I do not mean to say that these debates are unimportant. Rather, I mean to say only that I do not need to take a stand on or discuss them for current purposes.

What I do need is a distinction between maxim types and maxim tokens. As with the type/token distinction more broadly, this is a distinction between universals and particulars: a maxim token is an individual maxim as acted on, or at least adopted, by an agent, whereas a maxim type involves a level of abstraction above the token. To put the point somewhat colourfully, a maxim token cannot be separated from the lived experience of an agent whereas a maxim type can. An example will hopefully make this clear. Suppose that the LPM is a genuine maxim type. Then two different agents can act on the LPM, making lying promises in order to get some ready money. Their individual instances of self-governance in accordance with the LPM are tokens of the type. Along the same lines, if the NBM is a genuine maxim type, then two distinct agents are able to instantiate it at different times and in different places, not helping anybody on principle. Their distinct instantiations are distinct tokens of the type. Finally, if the EM is a genuine maxim type, then two distinct agents at different times and places can adopt the EM as a principle and act accordingly. That means two tokens of the type. Just as two individuals can utter the same proposition or perform act tokens of the same type, two individuals can adopt the same maxim (type). When that happens, their willings are qualitatively identical, but numerically distinct, tokens of the same type.

Claims about what the CI should be applied to are claims about the proper object of judgement. I think that the easiest way to make sense of this is by analogy with other philosophical debates. I shall mention two.

Consider the debate between deontologists and consequentialists. On one way of understanding this debate, the central question is about the proper object of judgement when it comes to value (Kahn 2014a: chapter 11). Deontologists assert that claims about the rightness or wrongness of rules are logically basic and, thus, that claims about the goodness or badness of the consequences of an action must be understood in terms of the rightness or wrongness of the rules followed by the intentional agents who brought about those consequences. On this way of understanding the debate, consequentialists, by way of contrast, assert that claims about the goodness or badness of states of affairs are logically basic and, thus, that claims about the rightness or wrongness of rules must be understood in terms of the goodness or badness of the states of affairs that would result from following these rules.

Alternatively, consider the debate between rule utilitarians and act utilitarians. On one way of understanding this debate, the central question is about the proper object of judgement when it comes to the principle of utility (Smart 1956: 344). Rule

utilitarians assert that act types are the proper object of judgement and, thus, that judgements about the rightness or wrongness of act tokens must be grounded in judgements about whether the act types instantiated in those tokens promote aggregate utility. On this way of understanding the debate, act utilitarians, by way of contrast, assert that claims about the rightness or wrongness of act types must be grounded in judgements about whether tokens of that type tend to promote aggregate utility.

I do not intend to take a stand in either of these debates here, nor do I assert that these are the only ways of characterising them. The point is simply to draw an analogy, one that hopefully will be illustrative. Similar to these ways of characterising debates about deontology/consequentialism and rule/act utilitarianism, the question at issue in this article is about whether the proper object of judgement for the CI is maxim tokens, or whether the proper object of judgement for the CI is maxim types. If the proper object of judgement for the CI is maxim tokens, then the deontic status of maxim tokens is logically basic and the deontic status of maxim types must be understood in terms of the deontic status of tokens of the type. If, by way of contrast, the proper object of judgement for the CI is maxim types, then the deontic status of maxim types is logically basic and the deontic status of maxim tokens can be inferred from the deontic status of the type. I am going to argue that the proper object of judgement for the CI is maxim tokens, not maxim types: I am going to argue for TOKEN and against TYPE.

It will be noticed that the options I am considering here are not, technically speaking, mutually exclusive: the CI could be applied to *both* maxim tokens *and* maxim types. Call this conjunctive thesis CONJUNCTION. As far as I am aware, nobody subscribes to CONJUNCTION. However, my arguments against TYPE also militate against CONJUNCTION and, indeed, in section 5 I suggest that this conjunctive thesis would be problematic.

It will also be noticed that the options I am considering here are not, technically speaking, exhaustive: the CI could be applied to general maxim types (e.g., false-promising maxims in general), a position that can be called GENERAL. However, not only do my arguments in favour of TOKEN militate against the thesis that the CI should apply *only* to general maxim types, but, more (and as I indicate in the conclusion of this article), I think GENERAL is both exegetically and philosophically unsustainable.

To make this more perspicuous, let me distinguish between GENERAL, TYPE, and TOKEN with an example. Consider the following maxim: 'I will make a lying promise in order to get a car', and consider two agents, X and Y, who adopt it. If GENERAL is true, then, in order to determine whether X and Y are acting permissibly, we do not apply the CI directly to the maxim type, much less to any agent's token adoption of it. Rather, we first classify this maxim together with the LPM as a lying promise maxim, and we then, second, determine that this maxim (along with the LPM) is impermissible because (we may suppose) the CI rules out lying promise maxims in general. From this, we infer that X and Y are acting impermissibly. Now, by way of contrast, suppose that GENERAL is false but that TYPE is nonetheless true. If TYPE is true, then the CI is applied to maxims in the abstract. Thus, on the supposition that the CI rules out the maxim 'I will make a lying promise maxim in order to get a car', we may again infer that X and Y are acting impermissibly. Finally, suppose that TYPE

is false and that TOKEN is true. Then we assess X's and Y's token adoptions directly and distinctly, and if we want to know about the maxim type or about lying promise maxims in general, then we have to generalise up.

To summarise: my goal in the following is to show that Kant's Categorical Imperative should be applied to individual maxim tokens and, in so doing, to show that it is not the case that Kant's Categorical Imperative should be applied to abstract maxim types, much less general maxim kinds. I begin, in the next section, by showing that I am not committing a straw person fallacy: I show that Rawls, unlike me, is committed to TYPE rather than to TOKEN.

2. Who thinks otherwise?

Rawls sets out his famous CI procedure, a procedure for determining whether a maxim is permissible, in four steps.

In the first step, the maxim that is to be tested using the CI is articulated. According to Rawls, maxims must have the form 'I am to do X in circumstances C in order to bring about Y unless Z' (Rawls 2000: 168). Thus, Rawls would not accept the LPM, the NBM, or the EM as complete. However (and as noted above), this can be ignored for current purposes.

In the second step, the maxim from the first step is universalised. Thus, 'I am to do X...' from the first step becomes 'Everyone is to do X...' in the second step. For example, the LPM would become something like 'Everyone will make a lying promise in order to get some ready money' (with the caveat noted in the first step, that this maxim is, on Rawls' account, incomplete).

In the third step, the universal precept from step 2 is considered to be a law of nature. This is done by appending to the end of it 'as if by a law of nature (as if such a law was implanted in us by natural instinct)' (Rawls 2000: 168). So, sticking with the LPM example, we have: 'Everyone will make a lying promise in order to get some ready money as if by a law of nature (as if such a law was implanted in us by natural instinct)'.

In the fourth step, the law of nature from step 3 is conjoined to existing laws of nature in order to determine whether the result can be willed coherently. If so, the original maxim is permissible; if not, the original maxim is impermissible.

Rawls makes various idealising assumptions about the way in which the CI procedure is to be applied in the fourth step: 'the thought is that an *ideal* reasonable agent considering whether to act from the maxim at step (1) implicitly accepts the requirements of pure practical reason represented in the steps leading up to and including step (4)' (Rawls 2000: 169, my emphasis). The idealisations associated with the CI include consideration of true human needs (pp. 173-5), limits on information (pp. 175-6), and others reminiscent of the original position from Rawls' theory of justice as fairness (Rawls 2001: part I, section 6).

The specifics of Rawls' idealisations are unimportant. What is important, however, is that, as I am going to argue now, Rawls' procedure evinces his commitment to TYPE as opposed to TOKEN.

As may be seen from the foregoing, the question, for Rawls, is not whether an actual concrete individual agent can will her individual maxim token as a law of nature. Rather, the question is whether an abstract, ideal agent can will the

corresponding maxim as a law of nature. The crucial move that Rawls makes (crucial for present purposes) is in inferring the deontic status of any tokens of the corresponding type from the answer to this question.

That is, instead of looking at whether an actual agent can universalise their maxim token as a law of nature and then generalising up to the deontic status of the maxim type, Rawls moves in exactly the opposite direction: he looks at whether an ideal agent can will the maxim as a law of nature and then infers that any instantiation of this maxim type will have the corresponding deontic status. In other words, on Rawls' account, we may infer that my token adoption of the LPM is impermissible from the fact that the CI procedure shows the corresponding type to be impermissible, and this presupposes TYPE rather than TOKEN.

Moreover, this kind of top-down procedure, inferring *from* the deontic status of the maxim type to the deontic status of its tokens, is standard practice in discussions of Kant's and Kantian ethics. We can see this from two things: (1) descriptions of the CI standardly evoke Rawls' ideas about true human needs – that is, Kantians are interested in whether maxims can be universalised in the abstract, not *in concreto*, on the basis of individual agents' actual needs – and, more especially, (2) applications of the CI standardly infer from the claim that (a) a maxim (type) can/not be universalised to the claim that (b) this maxim is im/permissible in all of its instances (tokens). That is, TYPE is presupposed by the standard operating procedure in discussions of the CI, especially in its universalisation formulations. Thus, for example, according to Willaschek, 'if my maxim not to borrow money I know I cannot repay can hold as a universal law, then so does your maxim with the same content' (Willaschek 2009: 64). To put this another way: instead of testing out maxim tokens individually, as required by TOKEN, Kantians test out a maxim independently of the messy details about any particular agent who might adopt that maxim, and they infer the deontic status of any such token from the results of this test, as is required by TYPE.

Now, it might be objected that what I am calling the standard practice in discussions of Kant's and Kantian ethics, the move from the permissibility of a maxim in the abstract to the permissibility of any tokens of that maxim, need not rely on TYPE. It could rely, instead, on TOKEN conjoined with the following uniformity thesis:

UNIFORMITY

All agents can universalise the same maxims.

However, there are at least two problems with this objection.

First, UNIFORMITY is false. What an agent is able to will depends to some extent on her beliefs. More particularly, what an agent is able to will at the same time as a law of nature depends to some extent on her beliefs. It follows that differences in beliefs can lead to differences in the ability to universalise a maxim. For example, I believe that non-Euclidean geometry is logically consistent, so I (unlike Hobbes) would not be able to will the maxim to prove otherwise (*a fortiori* I, unlike Hobbes, would not be able to universalise that maxim). Similarly, I believe that perpetual motion machines are physically impossible, so I (unlike Wilkins) would not be able to will the maxim to build one (*a fortiori* I, unlike Wilkins, would not be able to universalise that maxim). Other examples might appeal to beliefs about teleology in nature, such as the natural purpose of the motive of self-love. Thus, Kant, unlike those who eschew his

teleological view of nature, could not universalise the maxim to commit suicide from self-love. Indeed, Rawls was well aware of this (i.e., of the fact that what an agent is able to will depends to some extent on her beliefs). Not only would Rawls concede it; he would expand it. The whole point of the idealisations in the CI procedure is that what an agent is able to will as a law of nature depends not only on her beliefs but also on her needs, her psychological proclivities, etc.

Second, the fact that discussions of the CI do not discuss individual agents with their potentially false beliefs, unpalatable desires, idiosyncratic psychological tendencies, and strange embodied needs militates in my favour. This omission suggests, as argued above, that Kantians standardly presuppose TYPE, not TOKEN conjoined with UNIFORMITY. However, this second point requires some clarification.

An ideal agent can be imagined as adopting token maxims, in the same way that, say, a fictional character can be imagined as making token utterances. So, it is not the mere fact that Rawls appeals to an ideal agent that evinces his commitment to TYPE. Similarly, an abstract procedure can be used to assess a concrete agent's token maxim, in the same way that, for example, an abstract multiplication algorithm can be used to assess a concrete agent's utterance of an equation. So, it is not the mere fact that the CI procedure is abstract that evinces Rawls' commitment to TYPE.

What evinces Rawls' (and others') commitment to TYPE is, once again, the move from the im/possibility of a maxim, according to the CI procedure (or their favoured interpretation of the CI), to the im/possibility of all instances of that maxim.² It is the top-down approach, moving from type to token in inferring deontic status, that evinces the commitment to TYPE. This approach, although standard, is not always made explicit. But sometimes it is. Thus, according to Herman, '[a]n action [token] is said to be impermissible when it's an instance of a forbidden or impermissible *act type*: e.g., killing the innocent, torture' (Herman 2021: 135).

I am now going to argue that this gets things the wrong way around.

3. The second-personal Categorical Imperative

I think that there are at least two ways to see that the CI is about what concrete individual agents are able to will rather than about the im/possibility of abstract maxim types. The first is to contrast the FUL with the following principle:

L Conform your will to universal laws.³

Some maintain that understanding the difference between L and the FUL reveals a crucial lacuna in Kant's derivation of the latter (Aune 1980: 29; Wood 1999: 78-9). The current discussion might be relevant to that, but I maintain that understanding this difference reveals that TOKEN is correct whereas TYPE is not. Let me explain.

Both L and the FUL are addressed to individual agents (second-personally): they are imperatives. But the compliance conditions for L and the FUL are not the same. Whereas L tells an agent to will only maxims that *can be* universal laws, the FUL tells an agent to will only maxims that *he can will at the same time* as universal laws.

Conceptually, L and the FUL come apart in the following way. On the one side, L presupposes that there are some abstract universal laws that have that status entirely independently of the agent's will and, thus, that an individual maxim token is permissible if but only if it is in conformity with these laws, an instance of a

permissible abstract maxim type. On the other side, the FUL, by way of contrast, determines whether an individual maxim token is permissible directly, by virtue of whether an individual agent is able to will that individual maxim token at the same time as a universal law.

Concretely, L and the FUL come apart precisely because UNIFORMITY, the uniformity thesis discussed in the previous section, is false. If (as argued above) agents with different beliefs are able to universalise different maxims, then L and the FUL can yield different results regarding the permissibility of a maxim. To see why, suppose that agent A is able to universalise maxim M but agent B is not. Suppose, further, that M conforms to a universal law. Then L entails that M is permissible for either A or B whereas the FUL entails that M is permissible for A but impermissible for B. In other words, given these assumptions, the FUL entails that A's token adoption of the abstract maxim type M is permissible whereas B's is not. More, the FUL has no way of assessing the abstract maxim type except through the individual wills of agents and, thus, through its individual tokens.

Now, the argument in the previous paragraph is predicated on the falsity of UNIFORMITY. But note that, even if UNIFORMITY were true (it is not, but even if it were), L and the FUL would generate different results if either (i) there are maxims that conform to universal laws but that agents are unable to universalise or (ii) there are maxims that do not conform to universal laws but that agents are able to universalise. I argued that UNIFORMITY is false on the basis of (a) the connection between willing and believing and (b) the variation in beliefs among agents. But, even if someone rejects this argument, she nonetheless might accept (i) or (ii) on the basis of ideas about (uniform) imperfections in human volitional capacities. And accepting (i) or (ii) would suffice for my purposes here. However, I do not want to pursue this line of reasoning. Instead, I want to turn to the second way, hinted at above, of seeing that TOKEN is true.

The CI is supposed to be a principle of autonomy. Kant explains what this means when he contrasts his approach to other approaches in ethics. He argues that other philosophers have failed to articulate plausible ethical systems because they have failed to realise that a human 'is subjugated only by his own and nevertheless universal lawgiving will, and that he is only obligated to act . . . according to his own will' (G, 4: 432.29-32, emphasis omitted). Kant then uses this idea to motivate calling the CI 'the principle of the autonomy of the will', contrasting his principle to 'every other, which I therefore count as [a principle of] heteronomy' (4: 433.10-11, emphasis omitted).

What is important for present purposes is *not* that Kant thinks that every other ethicist before him has failed to articulate an adequate account of the Supreme Principle of Morality, much less that Kant thinks that all of these ethicists have failed for the same reason. Rather, what is important is why Kant sees fit to call his principle one of autonomy: because it is a principle that makes perspicuous that an agent's obligations are generated from her *own* will, from whether *she* can will a maxim at the same time as a universal law.

The reason this is important is that it entails that the CI is intended to apply to an individual agent's token willings rather than to abstract maxim types. Autonomy considerations reveal that a maxim is not, on Kant's account, supposed to be im/permissible in the abstract. Rather, individual token willings are im/permissible, and what makes them so is not whether they are tokens of an im/permissible maxim type,

but rather whether the individual agent under consideration is able to universalise them.

On this basis (that is, the contrast between L and the FUL and the foregoing argument about autonomy), I assert that the CI is about what individual agents are able to will, about individual maxim tokens rather than abstract maxim types. Thus, I conclude that TOKEN is true and that TYPE is false.

4. Ought implies can

In this section, I argue that the version of 'ought implies can' (OIC) to which Kant is committed coheres better with TOKEN than with TYPE. The argument has three parts. First, I distinguish two different versions of OIC. Then, I explain why one version coheres better with TOKEN. Finally, I argue that Kant is committed to that version.

Here are the two different versions of OIC:

OIC-TYPE

If an agent ought to A, then she has the ability to perform acts of type A in general.

OIC-TOKEN

If an agent ought to A, then she has the ability to perform the act token A.

The difference between OIC-TYPE and OIC-TOKEN can be illustrated by means of an example. Suppose that there is a car bearing down on a child and that P is nearby. Suppose, further, that it is a very hot day, that P went for a run that morning, and that P did not hydrate well after the run – so, if P tries to rush out to save that child, P will suffer a cramp that will prevent her from succeeding. From the fact that P went for a run that morning, it may be inferred that, in general, P has the ability to perform acts of this kind, even though, in this particular situation, P is destined to fail. So, in this example, OIC-TYPE does not nullify P's duty of rescue, but OIC-TOKEN does.

Now, I do not think that there are entailment relations between OIC-TYPE and TYPE or between OIC-TOKEN and TOKEN. Nonetheless, I do think that OIC-TYPE and TYPE cohere better than OIC-TYPE and TOKEN, and, similarly, I think that OIC-TOKEN and TOKEN cohere better than OIC-TOKEN and TYPE. To see why, note that, if TYPE is true, then the deontic status of a maxim can be determined in the abstract, independently of the capacities of an individual agent at a particular place and time. So, TYPE coheres better with a version of OIC that does not require a connection between obligation and these kinds of particularised capacities. Similar reasoning can be used to make plausible the connection between OIC-TOKEN and TOKEN. Let me try to make this clearer with an example.

Consider again the case of P, for whom the general ability to rescue and the specific ability to rescue diverge. As noted above, OIC-TOKEN, but not OIC-TYPE, nullifies P's duty of rescue. But, if TYPE is true, then P's duties are determined with regard to P's general capacities, not P's capacities in the moment. This makes it hard to see how OIC-TOKEN and TYPE would sync up: they seem to generate inconsistent results (in this case, that P does not have a duty of rescue and P does have a duty of rescue, respectively). Conversely, if TOKEN is true, then P's duties are determined with regard to P's capacities in the moment, not P's general capacities. This makes it hard to see how OIC-TYPE and TOKEN would sync up: they too seem to generate inconsistent

results (in this case, that P does have a duty of rescue and P does not have a duty of rescue, respectively).

In order better to explain the difference between OIC-TYPE and OIC-TOKEN, and also to begin to explain why I think Kant is committed to the latter, I want to examine the recent work of Herman and Kohl, who argue (independently) that Kant is committed to OIC-TYPE rather than OIC-TOKEN.

Herman's argument is based on the claim that OIC-TOKEN has absurd results:

If I am obligated to repay a debt on Wednesday but I squander the money on Tuesday, it seems reasonable to say that my inability to repay my debt does not remove my obligation. Surely it is not the case on Wednesday that I have no obligation to repay my debt. (Herman 1993: 163)

Herman's idea here is that, if someone has a debt and renders herself unable to repay it by squandering her money the day before it is due, then OIC-TOKEN, but not OIC-TYPE, entails the absurd result that the person's obligation to repay the debt is no longer binding.⁴ On the basis of this, Herman concludes that it is not the case that agents can be obligated to perform action *tokens* only if it is possible to perform them; nonetheless 'we can be obliged to do actions only of a *kind* that it is possible for us to do' (Herman 1993: 164, emphasis in original). The conclusion that Kant is committed to OIC-TYPE rather than OIC-TOKEN seems to be based primarily on considerations of charity.

The problem with Herman's objection is that OIC-TOKEN does not have the absurd result that she attributes to it. Squandering one's money the day before a debt is due is not a way of rendering the obligation non-binding; it is a way of infringing on the obligation. In other words, that squandering is itself constitutive of a blameworthy failure to fulfil the obligation to pay. Thus, if an agent squanders her money the day before a debt is due, she is not thereby released from her obligations and free of blame: Herman's motivation for subscribing to OIC-TYPE rather than OIC-TOKEN does not withstand critical scrutiny.⁵

Kohl also contends that OIC-TOKEN has absurd results. His two main arguments concern rationality and epistemic considerations, respectively:

Since the rationality of our choices cannot be affected by the (for us) unforeseeable vagaries of the empirical world, and since oughts (for Kant) are practical rules that provide conclusive standards for rational choice, these vagaries cannot determine what effects we ought to aim at accomplishing either: oughts must be tailored to the perspective of agents who deliberate from a position of inevitable uncertainty concerning their ultimate success in accomplishing intended effects. (Kohl 2015: 692)

... on Kant's view every person can know what she morally ought to do (5: 36; 6: 375); empirical circumstances that determine whether our capacities suffice for accomplishing the effects that moral oughts tell us to aim at are typically unknowable for us; thus, the validity of moral oughts cannot depend on those circumstances. (Kohl 2015: 701)

Kohl's arguments in both of these passages build on the fact that OIC-TOKEN nullifies obligations based on an agent's abilities 'in the moment'. In the first argument, Kohl uses this, along with the Kantian thesis that immoral action is (*eo ipso*) irrational, to show that, if OIC-TOKEN rather than OIC-TYPE were true, agents' rationality would depend on the unforeseeable vagaries of the world. In the second argument, Kohl argues that OIC-TOKEN, unlike OIC-TYPE, would make it impossible for an agent to know what she morally ought to do. Kohl contends that both of these results are unacceptable. However, Kohl's arguments do not work.

Kohl's first argument does not work because he misunderstands the implications of both OIC-TOKEN and OIC-TYPE. This may be seen from the example he uses to illustrate his point. Kohl considers an agent who is positioned to rescue a child from oncoming traffic but who, because of the particular situation, is doomed to fail because she will suffer a cramp (I used this example at the outset of this section to explain the difference between OIC-TOKEN and OIC-TYPE). Kohl argues that, in this case, OIC-TOKEN nullifies the duty of rescue and, thus, renders the attempt irrational, whereas OIC-TYPE does not. From this, Kohl concludes that OIC-TOKEN is absurd whereas OIC-TYPE agrees with intuition.

To see why this does not work, let us follow Kohl in distinguishing between the attempt to rescue, and the actual rescuing. The agent in Kohl's example has the general ability to rescue, but, because of her cramp, she does not have the ability to rescue in that instance. The agent also has the general ability to attempt to rescue, and this ability is not compromised by her cramp. From all of this, it follows that (a) OIC-TYPE is consistent with the rationality of the agent's attempt to rescue, but (b) so is OIC-TOKEN, because neither OIC-TYPE nor OIC-TOKEN nullifies a duty to attempt to rescue. However, it also follows that (c) OIC-TYPE, because it does not nullify the duty of rescue (because the agent retains the general ability to rescue), is consistent with the agent's failure to rescue being a blameworthy duty abrogation and, thus, irrational on Kant's account (because, as Kohl notes and as noted above, failure to fulfil a duty is, on Kant's account, irrational). But, this means that OIC-TYPE makes the agent's irrationality determined by 'unforeseeable vagaries of the empirical world', exactly the result Kohl wants to avoid. Finally, it follows that (d) OIC-TOKEN, by way of contrast with OIC-TYPE, does nullify the agent's duty of rescue. So, OIC-TOKEN does *not* make failure irrational.

From this, it may be seen that, in precisely the kind of case Kohl envisions, one in which the agent knows that her capacities *generally* suffice for accomplishing something, but does *not* know (or have any reason to suspect) that these capacities will fail *in this particular case*, an agent can have a duty to attempt to X but not a duty to X. Thus, Kohl's appeal to ideas about rationality in his first argument actually ends up supporting OIC-TOKEN rather than OIC-TYPE (*pace* Kohl), exactly the opposite of what he wants.

The problem with Kohl's second argument is more straightforward: OIC-TOKEN does not make it impossible for an agent to know what her duties are. Assuming that knowledge is factive but that justified belief is not, an agent might have a justified belief that she ought to D even though she is unable to D, and, in such a case, OIC-TOKEN would entail that this justified belief is not knowledge. But knowledge does not require apodictic certainty, and so the fact that an agent cannot attain that level of certainty in regard to whether her capacities will suffice in a given instance

does not threaten her ability to attain knowledge regarding her duties if OIC-TOKEN is true. Returning again to agent P from the beginning of this section, who believes she is able to rescue the child, OIC-TOKEN entails: (a) if P is destined to fail on account of a cramp, this belief is mistaken and does not count as knowledge and (b) if P is not destined to fail on account of a cramp (or whatever), then this belief might count as knowledge.

Based on this, it seems to me that the arguments for ascribing OIC-TYPE to Kant are, at best, weak. Moreover, I think that there are strong independent arguments in favour of ascribing OIC-TOKEN to Kant. Here are three.

First, OIC-TOKEN seems to be what Kant has in mind when he talks about OIC. For example, in an unpublished reflection Kant says that the 'subjective grounds and calculations of imputation are: the intention. The knowledge. The capacity. The readiness. The opportunity' (Refl 7128, 19: 254.15-16). Similarly, in the *Critique of Pure Reason* Kant remarks that, 'since it is prescribed that such [actions] ought to happen, so they must also be able to happen' (CPR, A806/B834; 3: 524.22-23). Note that in the first passage imputation requires not merely the capacity but also the opportunity, exactly the kind of language that is sometimes used in modern discussions to distinguish between general capacity and ability in a specific instance (opportunity to exercise the general capacity). Note also that the second passage contains no caveat about these actions merely being able to happen *in general*.

Second, OIC-TOKEN makes better sense than OIC-TYPE of Kant's highest good argument. According to this argument, (1) agents have a duty to promote the highest good, a world in which all agents are supremely virtuous and happiness is distributed in proportion to virtue; (2) the highest good is a real possibility only if God exists and agents have immortal souls; therefore, (3) agents are warranted in believing God exists and that they have immortal souls, at least for the purposes of fulfilling the duty to promote the highest good (Wood 1970).

The highest good argument makes a tacit appeal to OIC in the move from (1) and (2) to (3). The problem is that, on at least one way of understanding what it means to have a general capacity, OIC-TYPE is not strong enough to make this move coherent: agents might not have the opportunity to promote the highest good in worlds in which God does not exist and there are no immortal souls, but that does not impugn the general capacity to promote the highest good in these worlds. That is, if we understand general capacities as appealing to what we can do in nearby possible worlds in which blocking conditions are lifted or enabling conditions are realised, and if the existence of God and immortal souls are understood as enabling conditions for the realisation of the highest good, then we have the general capacity to promote the highest good provided that, in a nearby possible world in which God exists and we have immortal souls, our actions would do so, even if, in this world, precisely because God does not exist and we do not have immortal souls, our actions do not have this effect. Thus, charity seems to require ascribing OIC-TOKEN to Kant for the purposes of making sense of the highest good argument. Because the highest good argument is a mainstay of Kant's philosophy (it appears in most of Kant's major works, including the *Critique of Pure Reason*, the *Critique of Practical Reason*, and the *Critique of the Power of Judgement*), this generates strong grounds for ascribing OIC-TOKEN to Kant in general.

Third, OIC-TOKEN coheres better than OIC-TYPE with Kant's reasons for subscribing to OIC. As noted at the beginning of section 1, the difference between

the CI and the Supreme Law of Morality is that the former involves constraint because imperfectly rational beings do not always follow the moral law. Kant's commitment to OIC follows from the fact that he takes the Supreme Law of Morality to describe what an imperfectly rational agent would do if she were governed by reason. But Kant does not take the ability to be governed by reason to be a merely general capacity: this is something that all agents *ultima facie* ought to do in all instances and, more, something that all agents are able to do in all instances.

On these grounds (i.e., the failure of the arguments for ascribing OIC-TYPE to Kant and the direct textual and indirect doctrinal evidence in favour of ascribing OIC-TOKEN to him) I argue that Kant was committed to OIC-TOKEN. But, as already argued, OIC-TOKEN coheres better with TOKEN than with TYPE, whence I maintain that the argument of this section supports my thesis.

5. Conclusion

In this article, I argued in favour of TOKEN and against TYPE: I argued that Kant's CI should be applied to individual maxim tokens and not to abstract maxim types. I began by explaining what this means. I did so directly and indirectly – by analysing the terms in TOKEN and TYPE, and also by analogy with two other philosophical debates. I explained that Rawls (and many others besides) subscribe to TYPE rather than TOKEN. And then I gave three arguments in favour of TOKEN. The first was based on the wording of the CI. The second was based on ideas about autonomy. The third was based on how to understand Kant's commitment to OIC. I want to wrap up by gesturing towards two advantages of TOKEN: (1) it opens up a middle ground between moral absolutism and moral relativism and (2) it gives a plausible account of general duties.

Moral absolutism is the view that all agents at all times have the same duties. Many object to this view on the basis of the many cultural differences across societies in different places and at different times. The idea behind this objection is that actions do not occur in a vacuum; it is necessary to take various background beliefs and practices into consideration when assessing what someone is doing. This is especially so when someone is in a place where prevailing beliefs and practices are not what she takes them to be. Moral absolutism does not seem to be able to capture these nuances and, thus, it seems to be untenable.

Moral relativism is at the opposite extreme: it says that there is no absolute standard that can be used to assess actions and that the assessment must be based *solely* on cultural context. Many object to this view on the basis of morally bankrupt socially accepted beliefs and practices, including those bound up with racism, antisemitism, sexism, homophobia, etc. Moral relativism does not seem to be able to capture the fact that some actions are wrong regardless of cultural context.

If TOKEN is true, then there is an absolute duty to follow the prescriptions of the CI, and there might be uniformity across many (even all) agents regarding many of these prescriptions. But this uniformity is contingent, something to be discovered on the basis of the assessment of individual maxim tokens. If, by way of contrast, TYPE (or CONJUNCTION, the thesis that the CI can be applied to maxim tokens *and* maxim types) is true, then the deontic status of agents' individual maxim adoptions is determined in the abstract by appeal to the deontic status of the corresponding

maxim types. Thus TOKEN, unlike TYPE (and CONJUNCTION), charts a middle course between the Scylla of moral absolutism on the one side and the Charybdis of moral relativism on the other.⁶

General duties include prescriptions like ‘do not lie’ and ‘do not kill’. What is distinctive about general duties is that they have exceptions. For example, although it is generally wrong to lie, it might be permissible to do so when throwing a surprise party for a friend; although it is generally wrong to kill, it might be permissible to do so when someone is terminally ill, in great pain, and requesting to be put out of her misery. Because general duties have exceptions, it is difficult to understand both (a) how one might argue for them and (b) how one might argue against them.

GENERAL, the thesis that the CI is applied to general maxim kinds, gives a simple answer to (a): general duties are established directly – we simply group the maxims and then apply the CI accordingly. But this simple answer comes with a cost: it rules out the possibility of exceptions, the distinguishing mark of general duties. For that reason, I think that TOKEN provides a more plausible way of understanding general duties. I would argue that general duties are based on generalisations about agents and the circumstances in which they find themselves; these generalisations, which might be based on induction or abduction, are then used to make generalisations about whether tokens of a general kind of maxim are usually adopted permissibly. There are then two levels of approximation: there is approximation about the tokens of a type, and there is approximation about the types of a kind. And exceptions can enter in at either level of approximation.

If the generalisations are plausible, they establish a presumption against maxims of a certain kind. The way to argue against that presumption is to evince a large number of exceptions that make plausible the idea that the generalisations on which the presumption is based are mistaken and need to be rethought.⁷

Thus, I think that, in addition to the textual, doctrinal, and philosophical arguments in favour of TOKEN that I have advanced in the main text of this article, TOKEN has additional advantages over alternative theses in this problem space, advantages that I have tried to explain in the last few paragraphs.

Notes

1 All citations are to the standard Academy pagination. All translations are my own. I use the following abbreviations: *G* for *Groundwork to the Metaphysics of Morals*; *CPR* for the *Critique of Pure Reason*; and *Refl* for Kant’s unpublished reflections.

2 Some might still try to resist the conclusion that Rawls (and others) are committed to TYPE. They might argue that, so long as agents have all the same beliefs, desires, needs, etc., they are capable of universalizing the same maxims. This might not be sufficient to ground UNIFORMITY. But it can ground a close cousin, **UNIFORMITY’**: All agents can universalize the same maxims, provided they have the same beliefs, desires, needs, and other psychological proclivities mentioned by Rawls.

I think there are three things that are worth mentioning here. First, as a matter of detail, UNIFORMITY’ is false. What agents are able to will depends not only on the psychological proclivities mentioned by Rawls but also on their other maxims. Rawls does not mention this, and he does not need to, because he abstracts away from it in discussing an ideal agent. But it suffices to falsify both UNIFORMITY and this variant of it. Second, even if UNIFORMITY’ can be patched, these provisions render it unfit for the purpose for which UNIFORMITY was introduced, namely: to ground the inference from the non/universalizability of a maxim to its im/permissibility in all of its instances, and this is

precisely because the im/missibility of a maxim, if we have adopted TOKEN, is based on whether a particular agent is able to universalize that maxim, not on whether that agent can universalize that maxim in counterfactual circumstances. Third and relatedly, we are constrained here by two things: (a) what Rawls and other Kantians actually say and (b) what Kant actually says. This is because (a) I am arguing about how correctly to interpret Rawls and other Kantians in (b) their expositions of Kant's ethics.

3 An analogous principle for FH can be found in Kahn (2014b: section 1.2).

4 In the literature on OIC, this kind of objection is called an appeal to culpable inability. The objection is that any time an agent has an obligation, OIC makes it possible for her to nullify that obligation simply by making it impossible to fulfil (Kahn 2019: chapter 3 section 4).

5 The problem with Herman's argument is considerably deeper than this paragraph indicates. As may be seen from the quotations reproduced above, Herman does not include the words 'in general' in her articulation of OIC-TYPE. This omission undermines any attempt to distinguish OIC-TOKEN from OIC-TYPE. To make this concrete: in Herman's thought experiment, the agent is unable to perform an action token of paying a debt, and precisely because of this, the agent is unable to perform a token of the type 'paying a debt'.

Moreover, from the fact that there is a time *t* at which an agent is able to perform an act token of type A, it does not follow that the agent is *in general* able to perform act tokens of type A. The reason this is important is that, if Herman's omission is corrected, then her thought experiment can easily be filled in with details to generate a problematic result for OIC-TYPE. For example, suppose that an agent acquires a debt to pay a large sum of money by a certain date. The agent acquires this debt knowing that she will receive money to pay it the day before it is due (maybe the agent is informed that she will receive a large inheritance or a bonus by a certain date, and so she borrows money from a friend and goes out and buys a car, promising to pay the debt the day after the inheritance or bonus is received). Suppose, further, that the agent never has had anywhere close to that amount of money before. Then the agent does not *in general* have the ability to pay out sums of that kind and so she can have no corresponding obligation if OIC-TYPE is true. Similar results can be obtained by thinking about an agent who, for whatever reason, only temporarily has a physical or mental (in)ability.

Now to be fair to proponents of OIC-TYPE, this result can be meliorated by virtue of the fact that there are other action types that agents in these situations in general will be able to perform that will 'cover things up' so to speak. Nonetheless, this result leads me to think that Herman's argument evinces a misunderstanding of *both* OIC-TOKEN *and* OIC-TYPE.

6 Some might object at this juncture on the following grounds. If the results of the CI are dependent on an agent's beliefs, then the middle ground that I am advocating between absolutism and relativism has counterintuitive results. For example, suppose that an agent believes that members of group X are not as rational as members of group Y. It seems that, on the view that I am advocating, this agent should hold members of group X and members of group Y to different moral standards. This seems problematic. I would like to say three things about this.

To begin, Kant's ethics is more about an agent's own duties and responsibilities than about blaming and praising other agents. As such, Kant's ethics is more about the standards that an agent holds herself to than about the standards she holds other agents to. However, I will be told that this is hardly a solution to the problem, for, if an agent believes that members of group X are less rational than members of group Y, then, on the view that I am advocating, it seems that this agent also will have different duties and responsibilities to members of group X and members of group Y.

This leads me to the second thing I would like to say. I am not convinced that this is as counterintuitive as it might seem at first blush. Infants and small children arguably are less rational than adults, and not only does it not seem like a problem that our duties and responsibilities to them are different from our duties and responsibilities to adults, but, more, it seems like it would be a problem if it were otherwise. However, I will be told that the problem arises when an agent's beliefs about these groups are *false* – when an agent believes that members of group X are not as rational as members of group Y but the agent is in error about this. The view that I am advocating entails that, even in such a case, the agent's duties track their (mistaken) beliefs, and *that* is the problem.

This leads me to the third thing I would like to say. I am still not convinced that this is as counterintuitive as it might seem at first blush. For one thing, such an agent might be to blame for

her mistaken belief if, for example, it was formed negligently or as the result of self-deception. For another thing, imagine how we would assess an agent who took members of X and Y (not) to differ in this way but nonetheless took them to be owed the same (or different) duties – this, I think, would be a far greater moral failing.

Of course, there is more to be said here about each of these points, and there are other objections that need to be addressed. But this is not the place to put forth a fully fledged articulation or defence of the view.

7 In saying this I am disagreeing with Korsgaard. Korsgaard maintains that general duties have the status of provisional universalities, meaning that they are treated as exceptionless until an exception is encountered, whereupon the duty is modified accordingly (Korsgaard 1999: 24). I think that this is problematic for at least three reasons. First, as a reading of Kant, this makes incoherent Kant's inclusion of what he calls 'casuistical questions' in the *Metaphysics of Morals*. That is, after arguing for a general duty, Kant poses various questions that seem to be genuinely open-ended and up for discussion for the reader, something that is difficult to reconcile with the idea that these general duties should be treated as exceptionless. Second, it is difficult to see how such a duty could be argued for: if it is established on the basis of premises that have exceptions (as I have argued above), then it should be accepted that the duty will have exceptions; if the duty is taken *not* to have exceptions, then it seems that the premises must be similar, and so an exception would require a new argument (not merely a slight modification to the duty in the form of a caveat). Third and finally, Korsgaard's proposal does not agree with general practice: as kids become adults, they learn about ever more exceptions to the general duties they learned as children, but they carry around in their heads and teach their children the same general duties they originally learned; fine print is not added on to that cultural heritage.

References

- Allison, Henry E. (1990) *Kant's Theory of Freedom*. Cambridge: Cambridge University Press.
- Aune, Bruce (1980) *Kant's Theory of Morals*. Princeton: Princeton University Press.
- Baker, Judith (1988) 'Counting categorical imperatives'. *Kant-Studien*, 79(4), 389–406.
- Bittner, Rüdiger (1974) 'Maximen'. In Gerhard Funke (ed.), *Akten des 4. Internationalen Kant Kongresses* (Berlin: Walter de Gruyter), 485–98.
- Gressis, Rob (2010) 'Recent work on Kantian Maxims I'. *Philosophy Compass*, 5(30), 216–27.
- Herman, Barbara (1993) *The Practice of Moral Judgment*. Cambridge, MA: Harvard University Press.
- (2021) *The Moral Habitat*. Oxford: Oxford University Press.
- Johnson, Robert and Cureton Adam (2022) 'Kant's Moral Philosophy'. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Stanford, CA: Stanford University Press). <https://plato.stanford.edu/entries/kant-moral/>
- Kahn, Samuel (2014a) *A Problem Based Introduction to Philosophy*. Dubuque, IA: Kendall Hunt.
- (2014b) 'The interconnection of willing and believing in Kant's and Kantian ethics'. *International Philosophical Quarterly*, 54(2), 143–57.
- (2019) *Kant, Ought Implies Can, the Principle of Alternate Possibilities, and Happiness*. Lanham, MD: Lexington Press.
- Kohl, Markus (2015) 'Kant and "Ought Implies Can"'. *The Philosophical Quarterly*, 65(261), 690–710.
- Kitcher, Patricia (2003) 'What is a maxim?' *Philosophical Topics*, 32(1/2), 215–43.
- Korsgaard, Christine (1996) *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.
- (1999) 'Self-constitution in the ethics of Plato and Kant'. *The Journal of Ethics*, 3(1), 1–29.
- O'Neill, Onora (1989) *Constructions of Reason*. Cambridge: Cambridge University Press.
- (2013) *Acting on Principle*. Cambridge: Cambridge University Press.
- Nyholm, Sven (2017) 'Do we always act on maxims?' *Kantian Review*, 22(2), 233–55.
- Pogge, Thomas (1998) 'The categorical imperative'. In Paul Guyer (ed.), *Kant's Groundwork of the Metaphysics of Morals* (Lanham, MD: Rowman and Littlefield), 89–215.
- Rawls, John (2000) *Lectures on the History of Moral Philosophy*. Cambridge, MA: Harvard University Press.
- (2001) *Justice as Fairness*. Cambridge, MA: Harvard University Press.
- Smart, J. J. C. (1956) 'Extreme and restricted utilitarianism'. *The Philosophical Quarterly*, 6(25), 344–54.

- Willaschek, Marcus (2009) 'Right and coercion'. *International Journal of Philosophical Studies*, 17(1), 49–70.
- Wood, Allen (1970) *Kant's Moral Religion*. Ithaca, NY: Cornell University Press.
- (1999) *Kant's Ethical Thought*. Cambridge: Cambridge University Press.
- (2006) 'The supreme principle of morality'. In Paul Guyer (ed.), *The Cambridge Companion to Kant and Modern Philosophy* (Cambridge: Cambridge University Press), 342–80.