

Pauline Kleingeld*

Open Access

Contradiction and Kant's Formula of Universal Law

DOI 10.1515/kant-2017-0006

Abstract: Kant's most prominent formulation of the Categorical Imperative, known as the Formula of Universal Law (FUL), is generally thought to demand that one act only on maxims that one can will as universal laws without this generating a contradiction. Kant's view is standardly summarized as requiring the 'universalizability' of one's maxims and described in terms of the distinction between 'contradictions in conception' and 'contradictions in the will'. Focusing on the underappreciated significance of the simultaneity condition included in the FUL, I argue, by contrast, that the principle is better read as requiring that one be able to will *two things* simultaneously without *self-contradiction*, namely, that a maxim be one's own and that it be a universal law. This amounts to a new interpretation of the FUL with significant interpretive and philosophical advantages.

Keywords: Formula of Universal Law, Categorical Imperative, contradiction, maxim, volitional self-contradiction.

Introduction

Immanuel Kant's most prominent formulation of the Categorical Imperative, known as the Formula of Universal Law (henceforth FUL), is generally thought to demand that one act only on the basis of maxims that one can will as universal laws. On existing interpretations, the question one should ask is whether one can will a candidate maxim as a universal law without this generating a contradiction. There are different accounts of the nature of the contradiction involved, but each interpretation seems to face problems. Kant insists that impermissible maxims 'contradict themselves' and involve a 'self-contradiction in the will', but on the standard interpretations it is impossible to interpret this in an ordinary sense of 'self-contradiction'. In this essay, I propose an alternative reading of the contradiction at issue and argue that it solves these problems.

*Kontakt: Prof. Dr. Pauline Kleingeld, University of Groningen, Faculty of Philosophy, Oude Boteringestraat 52, 9712 GL Groningen, Netherlands; pauline.kleingeld@rug.nl

 © 2017 Pauline Kleingeld, publiziert von De Gruyter. Dieses Werk ist lizenziert unter der Creative Commons Attribution-NonCommercial-NoDerivatives 3.0 Lizenz.

Brought to you by | University of Groningen
Authenticated

Download Date | 3/20/17 9:52 AM

The FUL requires that one act only on maxims that one can “*simultaneously* will as a universal law” (GMS, AA 04: 421, emphasis added). On leading interpretations, this simultaneity condition is routinely omitted from the description of the FUL. On the interpretation I defend in this essay, by contrast, it is essential to articulating the moral criterion expressed by the Formula. I argue that the FUL requires that one ask whether one can will a maxim as one’s own and simultaneously will this maxim as a universal law.¹ Maxims are impermissible if willing both simultaneously constitutes a self-contradiction of the will – which I term a volitional self-contradiction.² Consider, for example, the maxim to borrow money while promising falsely to repay it, when one believes oneself to be in need of money (cp. GMS, AA 04: 422). Willing this maxim as one’s own contradicts willing the maxim as a universal law, because – as Kant himself insists – if this maxim were a universal law no one could make false promises and get money in this way, and willing *that one act on the maxim oneself* contradicts willing *that no one does*. How exactly this volitional self-contradiction is *generated* will be discussed below.

I argue that this account of the nature and location of the relevant contradiction makes better sense of Kant’s repeated claims that maxims that fail the test ‘contradict *themselves*’ or that they lead to ‘*self-contradictions*’ of the will. Moreover, the account employs an ordinary and straightforward sense of ‘*self-contradiction*’ of the will, as meaning that one wills *A* and simultaneously wills *not-A*. If this argument is convincing, it provides not only a plausible interpretation of the texts but also a solution to important philosophical problems that have long been associated with the FUL.

I first examine the leading interpretations of the nature of the contradiction generated in the case of impermissible maxims, and I outline the main problems raised by these accounts of the FUL (section 1). I then develop the Volitional Self-Contradiction account and highlight its interpretive and philosophical

1 I use ‘willing a maxim’ as shorthand for ‘willing that a maxim serve as one’s action principle’. Willing something, as distinct from wishing and from feeling inclined towards something, involves a commitment to bringing something about to the extent to which this is in one’s power (GMS, AA 04: 394). In this case, it involves the commitment to using a maxim as one’s own individual action principle or policy. Note that the FUL assesses the moral permissibility of *maxims*; it concerns the moral permissibility of *actions* only indirectly. Ethics, on Kant’s conception, “does not give laws for *actions* [...] but only for *maxims* of actions” [Die Ethik giebt nicht Gesetze für die Handlungen [...], sondern nur für die Maximen der Handlungen] (MS, AA 06: 388).

2 The primary purpose of this choice of terminology is to avoid confusion, given that the term ‘contradiction in the will’ is currently used in an importantly different sense, as I discuss below. Another advantage is that it highlights Kant’s point that the relevant contradiction is a *self-contradiction*.

advantages, using two of Kant's own examples as illustrations (section 2). In the third section, I discuss the infelicity of using 'universalizability' as shorthand for the moral criterion articulated by the FUL, and I consider the disadvantages of describing Kant's view in terms of the distinction between 'contradictions in conception' and 'contradictions in the will' (section 3).

My focus in this essay is on the nature and location of the contradiction generated in the case of maxims that fail the criterion articulated in the FUL. I do not discuss Kant's derivation of the Formula, his argument for its validity, or its relation to the other formulations of the Categorical Imperative. Although a re-interpretation of the FUL of the sort offered here can be expected to have significant implications for the interpretation and assessment of Kant's arguments in support of it, as well as for the understanding of its relation to the other formulations of the Categorical Imperative, I defer the discussion of these issues to another occasion.

1 Where is the contradiction?

The FUL reads as follows:

*Act only in accordance with that maxim through which you can simultaneously [zugleich] will that it become a universal law.*³ (GMS, AA 04: 421, emphasis in original)

In virtually all discussions – from introductory textbooks to scholarly literature – this requirement is interpreted as follows: one ought to act only on maxims that one can will as universal laws. 'Simultaneously' is routinely omitted from the description of the moral criterion articulated by the FUL. The following representative statements from prominent Kant scholars and Kantian moral theorists are typical in this regard:

FUL [tells us] that morality requires that we act only on a maxim that could in fact be a universal law. (Guyer)⁴

³ "Handle nur nach derjenigen Maxime, durch die du zugleich wollen kannst, daß sie ein allgemeines Gesetz werde." One may wonder what to make of the locution 'maxim *through* which', because Kant's sentence construction is somewhat awkward. In many other passages, however, Kant mentions the idea of 'giving universal law through one's maxim' (GMS, AA 04: 432, 433, 434, 438, 439, 440; and the Formula of the Law of Nature, GMS, AA 04: 421). It seems that this idea is meant here as well.

⁴ Guyer 2007, 83.

The categorical imperative is the law of acting only on maxims that you can will to be universal laws. (Korsgaard)⁵

Kant's question is whether *everyone could will a maxim*. (O'Neill)⁶

[T]he criterion it [=FUL] embodies is whether a maxim can be willed as universal law for rational agents with autonomy. (Reath)⁷

[The FUL] tells us that it is permissible to act only on those maxims we could will to be universal laws. (Wood)⁸

These formulations differ in a number of important respects, and I do not mean to suggest that there is one standard reading of the FUL as a whole. My interest here is specifically in the fact that these authors all regard the requirement to be that one act only on maxims that can be willed as universal laws and that they all omit 'simultaneously' from their descriptions of what the Formula requires. The reason they omit it is not that they fail to notice it, and below I discuss the role they attribute to the simultaneity condition. Yet they clearly assume that this condition can be omitted from a description of the moral criterion articulated in the FUL without this affecting its meaning in any important way.⁹ I shall argue below, by contrast, that this omission has serious consequences.

1.1 Three leading accounts of the nature of the relevant contradiction

It is generally acknowledged that a crucial element of the moral criterion articulated through the FUL is the idea that some maxims, when 'universalized', would involve a *contradiction* of some kind, and that it would be impermissible to act on

⁵ Korsgaard 2009, 80; similarly on 81, 153, 209.

⁶ O'Neill 2004, 98.

⁷ Reath 2006, 204 f.; similarly on 206, 219.

⁸ Wood 2006, 350; see also Wood 1999, ch.3, and Wood 2008, 70.

⁹ Because it is widely regarded as adequate to formulate the FUL requirement without including 'simultaneously', authors omit it not merely from their own explanations of what the FUL means but also from their descriptions of the FUL itself. Here are just a few examples: "Consider whether your maxim can be asserted as a universal principle" (Hegel 1991 [1821], § 135, *Zusatz*); "So act, that the rule on which thou actest would admit of being adopted as a law by all rational beings" (Mill 1985 [1861], 207); "Kant's *Formula of Universal Law*: It is wrong to act on any maxim that we could not will to be a universal law" (Parfit 2011, vol. 1, 182).

such maxims. There are different accounts of the contradiction involved,¹⁰ but each of the leading interpretations is in some respect dissatisfying.

On one influential line of interpretation, the contradiction involved in willing a maxim as a universal law is a contradiction between willing the universalized maxim and endorsing a certain value. Following Hegel, critics have asserted that contradictions emerge only on the basis of substantive evaluative presuppositions. As Hegel puts it: "A contradiction must be a contradiction with something, that is, with a content that is already presupposed as a fixed principle."¹¹ As an example, he mentions theft. The absence of property, in a world in which the maxim of stealing is a universal law is not contradictory, he argues; only if one presupposes that there should be property, or that property should be respected, does it become contradictory to will the maxim of stealing as a universal law.¹² The relevant contradiction, according to Hegel, is a contradiction between endorsing a presupposed value, on the one hand, and willing a given maxim as a universal law, on the other. On the basis of this account of the contradiction, the FUL is understandably criticized as being 'empty', since the Formula by itself is insufficient to identify any maxim as impermissible.¹³

Over the past few decades, several Kantian moral theorists have developed new responses to this objection. Onora O'Neill and Christine Korsgaard, among others, have argued that it is possible to give an account of the relevant contradiction *without* reference to substantive values (such as the value of property, in Hegel's example). Instead, they claim, the contradiction generated by trying to 'universalize' an impermissible maxim can and should be understood in terms of the presuppositions of finite rational agency itself. They claim that willing impermissible maxims as universal laws involves a 'practical contradiction' (on Korsgaard's account) or 'inconsistency in acting' (on O'Neill's account).

On Korsgaard's account, the sense in which an immoral maxim involves a contradiction is that willing it as a universal law means willing the thwarting of one's own purpose. In her seminal article, "Kant's Formula of Universal Law",

10 In this essay, I discuss only those interpretations of the FUL that are most widely embraced in the current literature. It will become clear to readers familiar with other interpretations that the Volitional Self-Contradiction interpretation differs from these, too. For helpful reviews and critical discussions of additional interpretations of the FUL, see Korsgaard 1996, Timmons 2006, and Galvin 2009. See also note 17 below.

11 Hegel 1991 [1821], § 135 (translation modified).

12 *Ibid.*

13 *Ibid.* Hegel's version of this 'empty formalism' objection has been the most influential, but he was not the first to formulate it. As Jens Timmermann notes, G. A. Tittel had already articulated a similar complaint in 1786 (Timmermann 2007, 82, n.70).

she argues that there is a “*specifically practical sense of ‘contradiction’*” involved in willing a maxim that, when universalized, would thwart the purpose stated in the maxim (in the case of a ‘contradiction in conception’), or thwart a purpose that is essential to the will (in the case of a ‘contradiction in the will’).¹⁴ In a world in which such a maxim is a universal law, the agent is unable to act on it so as to achieve the purpose stated in the maxim (in the case of contradictions in conception), or unable to achieve a purpose that is essential to the will (in the case of contradictions in the will). An example from the first set of cases is the maxim to promise falsely to repay a loan, in order to get money easily: If this maxim were a universal law, then promises to repay, made by those requesting loans, would not be believed, and one could not get easy money by promising falsely to repay. An example from the second set of cases is the maxim of egoism. Suggesting that general effectiveness in the pursuit of one’s ends is an ‘essential purpose’ of the will as such, Korsgaard explains the practical contradiction that emerges in the case of the maxim of egoism as follows: In a world in which the maxim of egoism is a universal law, we will not receive the help we might need in order to be effective in the pursuit of our own ends; therefore, willing this maxim as a universal law involves thwarting the will’s own purpose.¹⁵

The simultaneity condition, on Korsgaard’s account, comes down to the requirement that it be possible to will a maxim as a universal law and, *in the world of the universalized maxim*, simultaneously to act on this maxim. She argues that, both in the case of the contradiction in conception test and in the case of the contradiction in the will test, the agent’s purpose is thwarted “*in the world where maxims that fail these tests are universal law*”.¹⁶ In the world of the universalized maxim, an agent cannot will to act on the maxim without at the same time willing the frustration of his own end.¹⁷

This explains why, on Korsgaard’s rendering of the FUL, ‘simultaneously’ can be omitted. The moral requirement, on her view, is that one should be able (rationally, without practical contradiction) to *will one’s maxim as a universal law*. The simultaneity condition plays a role only in the *test* for finding out whether this requirement is met. This test is whether the maxim’s universalization involves a practical contradiction for agents acting in the world of the universalized maxim.

¹⁴ Korsgaard 1996, 93, emphasis added; 92, 102.

¹⁵ *Ibid.*, 96 f.

¹⁶ *Ibid.*, 96 (emphasis added).

¹⁷ *Ibid.*, 93. On the ‘Logical Contradiction’ interpretation, too, the contradiction is located in the world of the universalized maxim. The contradiction is explained as follows: “The agent could literally no longer act on the original maxim in a world containing the U[niversal]-C[ounterpart] maxim” (Galvin 2009, 69).

If it does, then the maxim cannot be willed to be a universal law and is morally impermissible. Because the role of the simultaneity condition is restricted to this test, on Korsgaard's account, 'simultaneously' can be omitted from the description of the FUL itself.

O'Neill's interpretation of what it means to be able to will one's maxim as a universal law differs from Korsgaard's, because on her view it means that the maxim *can be willed by all*. She argues that rational agents as such are committed to certain rational principles, including the principle of willing the necessary and some sufficient means to any end to which one is committed. Furthermore, she presupposes that moral deliberators have basic knowledge of their own vulnerability and of the normal effects of acting on the maxims under consideration.¹⁸ In light of these principles and conditions, she suggests, we can determine whether it is possible for all to will a given maxim. On O'Neill's view, a maxim can fail this requirement in two ways. First of all, we cannot will a maxim as a universal law if the project stated in the maxim is *impossible* in a world in which the maxim is a universal law (such as the project of making false promises in a world of universal deception); universalizing such maxims yields a contradiction in conception.¹⁹ Second, we cannot will a maxim as a universal law if it would be *inconsistent*, for agents who are aware of the background conditions of their own agency, *to will* the universalized version of a maxim; this would be a contradiction in the will. For example, it would be inconsistent, for agents who are aware of their own vulnerability and finitude, to will a world of egoism, because they know (as a matter of empirical fact) that they often need help for the survival of their own agency.²⁰ Willing this maxim as a universal law would contradict willing what they will *qua* human finite rational beings.

The role O'Neill attributes to the simultaneity condition is rather different from that envisioned by Korsgaard, but again it plays a role only in the *test* to see whether a maxim can coherently be willed as universal law. Thus, she writes:

To universalize maxims agents must satisfy themselves that they can both adopt the maxim and simultaneously will that others do so. In determining whether they can do so they may find that they are defeated by either of the two types of contradiction [viz., contradictions in conception or contradictions in the will].²¹

¹⁸ O'Neill 1989, 81–104.

¹⁹ *Ibid.*, 132.

²⁰ *Ibid.*, 81–104, 126–135, esp. 132; see also Bojanowski 2014, 217.

²¹ O'Neill 1989, 95. In earlier work, O'Neill gave a somewhat different account of the contradiction, but there too she conceived of the contradiction as being contained in the conception or in the willing of the universal counterpart of a maxim that fails the test. Moreover, she also explicated the problem in terms of intending to do x and simultaneously intending that "every-

O'Neill here interprets the simultaneity condition as a principle of addition: If I can adopt a maxim and simultaneously will that *all others* also adopt it, then this *adds up* to the result that I can will the maxim as a universal law. In other words, on her view, too, the FUL requires that one act on maxims that one can will as universal laws; again it seems that 'simultaneously' can be omitted from the formulation of the moral criterion.²²

1.2 The problem of interpreting 'self-contradiction'

The interpretations provided by Korsgaard and O'Neill have greatly strengthened the case for Kantian ethics. They provide clear strategies to explicate the way the FUL provides normative guidance for assessing maxims, using only a conception of human rational agency – without, that is, presupposing substantive values.

Yet, even on these interpretations of the FUL, there is something dissatisfying about the account of the relevant 'contradiction' involved in willing a maxim as a universal law. For it does not fully match Kant's own articulation of the problem with maxims that fail the test. Kant discusses this problem in terms of such a *maxim*, or the *will* that adopts it, *contradicting itself*. He writes that a maxim that fails the test cannot "harmonize with itself" [mit sich selbst zusammenstimmen] but must "necessarily contradict itself" [sich nothwendig widersprechen] or be "in conflict with itself" [sich selbst widerstreiten] (GMS, AA 04: 422, 437). Because the will is internally conflicted in such cases, Kant also writes that "such a will would contradict itself" [ein solcher Wille [würde] sich selbst widersprechen] (GMS, AA 04: 424), there would be a "contradiction in our own will" [Widerspruch in unserm eigenen Willen] (*ibid.*), and the will would "be in conflict with itself" [sich selbst widerstreiten] (GMS, AA 04: 437, 423).

On Korsgaard's and O'Neill's accounts, by contrast, at least in cases of 'contradictions in the will', the contradiction is actually a contradiction between the universalized maxim and something entirely different, namely, an essential purpose

one else" do x, and she introduced empirical premises to generate the contradiction (cp. O'Neill [Neill] 1975, 59–93).

²² Similarly, Paul Guyer writes that the "FUL tells us to act on a maxim that could also be a universal law" (Guyer 2007, 83), and that "this requirement is to act only on particular principles that could also be universal laws, that is, to act only on principles that everyone else could also accept and act upon" (*ibid.* 42). Here, 'also' is used in the sense of 'additionally', and the FUL is read as requiring us to act on maxims that we could will to be principles not just for ourselves but also, in addition, for all others, such that, when we add it all up, the maxim could be a law for all, that is, a universal law.

of the will as such or the presuppositions and conditions of human rational agency more broadly. In other words, their accounts do not yield a *self-contradiction of the maxim*. On Korsgaard's account, moreover, the contradiction lies in the maxim's being *self-defeating* when universalized, or in the will's *thwarting* its own purpose, but this does not seem fully equivalent to a 'self-contradiction'. It is precisely the seeming impossibility of understanding the contradiction in the ordinary sense of the term that led her to introduce the notion of a 'specifically practical sense' of 'contradiction'.

In what follows, I argue that there is a way to take Kant's language of 'self-contradiction' of the maxim literally, and that the simultaneity condition provides the key to a significantly different reading of the FUL with important interpretive and philosophical advantages. On this reading, the simultaneity condition, spelled out in full, is a condition of being simultaneously willable: the possibility of willing a maxim as one's own action principle (as willed independently of or prior to its universalization) and simultaneously willing this maxim as a universal law, without this generating a volitional self-contradiction. On this alternative account of the FUL, 'simultaneously' is a crucial element in the formulation of *the moral criterion itself*, and leaving the word out means losing the essence of the principle.

2 The Volitional Self-Contradiction account

2.1 'Simultaneously' in the FUL and related formulations

Kant's most prominent statements of the principle of morality, in each of his main works in moral theory, all include a simultaneity condition. They all include the word '*zugleich*', which means 'simultaneously' or, synonymously, 'at the same time'.²³ Here again is the FUL, which Kant presents, with emphasis, as 'the' Categorical Imperative:

Act only in accordance with that maxim through which you can simultaneously will that it become a universal law. (GMS, AA 04: 421, emphasis in original)

²³ Kant uses the term not only in the temporal but also in the logical sense, for example when he speaks of "ends" [Zwecken] that "are simultaneously (that is, by their concept) duties" [die zugleich (d.i. ihrem Begriffe nach) Pflichten sind] (MS, AA 06: 385).

Later in the second section of the *Groundwork*, Kant offers what he calls the ‘general formula’ of the Categorical Imperative:

Act in accordance with that maxim which can simultaneously make itself into a universal law.²⁴ (GMS, AA 04:436 f.)

In the *Critique of Practical Reason*, Kant formulates the ‘fundamental law of pure practical reason’ as follows:

So act that the maxim of your will can always simultaneously hold as the principle of a universal legislation.²⁵ (KpV, AA 05: 30)

Finally, here is the formulation of the Categorical Imperative in the Introduction to the *Metaphysics of Morals*:

Act on the basis of a maxim that can simultaneously hold as a universal law!²⁶ (MS, AA 06:225; almost identical on 06: 226)

These are the formulations highlighted by Kant as the most important statements of the principle of morality, in his three most important works on the subject, and they all include ‘simultaneously’. So do most of the related formulations in other passages, such as the formulations he calls the ‘Principle of Autonomy’ (GMS, AA 04:440) and the ‘Formula of an Unconditionally Good Will’ (GMS, AA 04:437).²⁷

There are a few formulations that leave it out, most notably the statement known as the Formula of the Law of Nature (GMS, AA 04: 421).²⁸ Furthermore, there are several formulations in which Kant uses ‘also’ [auch] (GMS, AA 04: 402,

24 “Handle nach der Maxime, die sich selbst zugleich zum allgemeinen Gesetze machen kann.”

25 “Handle so, daß die Maxime deines Willens jederzeit zugleich als Princip einer allgemeinen Gesetzgebung gelten könne.”

26 “Handle nach einer Maxime, die zugleich als allgemeines Gesetz gelten kann.”

27 The ‘Principle of Autonomy’ is the command “not to choose otherwise than so that the maxims of one’s choice are simultaneously included as universal law within the same volition” [nicht anders zu wählen als so, daß die Maximen seiner Wahl in demselben Wollen zugleich als allgemeines Gesetz mit begriffen seien] (GMS, AA 04: 440). The ‘Formula of an Unconditionally Good Will’ reads: “Act in accordance with maxims that can simultaneously have themselves as universal laws of nature for their objects” [Handle nach Maximen, die sich selbst zugleich als allgemeine Naturgesetze zum Gegenstande haben können] (GMS, AA 04: 437).

28 See, however, note 48 below.

403, and 447).²⁹ 'Simultaneously' is used in most versions, however, including all of the most prominent ones.

Given that so many of the formulations of the Categorical Imperative contain a simultaneity condition, there is good reason to take a closer look at the contribution it makes to the meaning of the statements. Perhaps omitting it is not as harmless as it seems.

I do not mean to imply that Kant was always consistent in his description of the moral criterion or in its application to specific examples. Rather than focusing on the differences between Kant's formulations,³⁰ however, I shall focus on the question of just *what* one should be able to *will simultaneously*, according to the FUL. This turns out to provide the key to a new³¹ understanding of the contradiction generated in the case of maxims that fail to meet the FUL criterion.

2.2 Volitional Self-Contradictions: A new reading of the FUL

What is it that we are morally required to be able to 'will simultaneously'? I would like to suggest the following answer: that the maxim be our own individual action principle and that it be a universal law. A maxim, on Kant's account, is an agent's action principle [Princip, Grundsatz], that is, the principle on which an agent acts (GMS, AA 04: 421n.; KpV, AA 05: 19). *Acting on the basis of* a maxim implies willing that it serve as one's own action principle. Textually speaking, then, it is possible to read the simultaneity condition in the FUL as referring to the simultaneous compossibility³² of willing that a maxim be one's own maxim and willing

29 "I ought never to proceed except so that I could also [*auch*] will that my maxim should become a universal law" [ich soll niemals anders verfahren als so, daß ich auch wollen könne, meine Maxime solle ein allgemeines Gesetz werden] (GMS, AA 04: 402), and "I ask myself only: Can you also [*auch*] will that your maxim become a universal law?" [Kannst du auch wollen, daß deine Maxime ein allgemeines Gesetz werde?] (GMS, AA 04: 403). For reasons clarified below, I propose that we read 'also' in the sense of 'simultaneously', which is one of its core meanings.

30 For example, there are important differences between the requirements that one's maxim simultaneously be able to 'hold as', 'make itself into' or 'be willed as' universal law. In this essay I focus on the FUL, which is why I focus on the formulation in terms of 'willing'.

31 'New', that is, relative to the interpretations currently considered in leading scholarly debates. I cannot rule out the possibility that this reading of the FUL has been presented somewhere in the vast Kant literature of the past two centuries. I am unaware of any discussion of this reading elsewhere, however, and it is not included in overviews such as those by Korsgaard 1996 or Galvin 2009, nor in the detailed taxonomy in Timmons 2006.

32 'Simultaneous compossibility' might seem to be a pleonasm, but it is not. As Kant explains in the *Critique of Pure Reason*, it is not contradictory to say of a person that he is 'young' and

that it be a universal law. Or, put more simply, the FUL can be read as requiring us to *act only on maxims that we can will as our own maxim and simultaneously will as a universal law*.³³

Thus, to return to the example mentioned in the introduction and elaborated in more detail below: consider Kant's example of the maxim to borrow money while promising falsely to repay it, when one believes oneself to be in need of money. Willing to act on this maxim and simultaneously willing that it be a universal law constitutes a volitional self-contradiction. The reason for this is that if the maxim is a universal law, then it is impossible to get money by making false promises, and willing that *I act on the maxim* contradicts willing that *no one acts on the maxim*. A full understanding of how this contradiction is generated requires further discussion, however. In the remainder of this section I develop the main features of the account I propose. Subsequently, in section 2.3, I discuss this example, as well as Kant's example of the maxim of egoism, in full detail.

The first point to note concerns the *location* of the relevant contradiction that emerges in the case of impermissible maxims. On existing interpretations, the relevant contradiction is located *in*, or seen as 'involved in', the conception or the willing of a world in which the maxim is a universal law. By contrast, if we read the FUL along the lines I propose, the relevant contradiction should be located *between* willing that the maxim be one's own action principle (as willed independently of or prior to its universalization) *and* willing that it be a universal law.

Importantly, Kant's early discussions of self-contradictions of the will, in his remarks on the *Observations on the Feeling of the Beautiful and Sublime* (probably from 1764–65), are striking evidence of exactly this understanding of the relevant contradiction. They suggest a close connection between Kant's account of this contradiction and Jean-Jacques Rousseau's account, in the *Social Contract* and *Emile*, of the "contradiction" between the "particular will" and the "general will".³⁴ Here is how Kant describes the good will:

that he is 'old' (not young), because he can be 'young' first and 'old' later. But he cannot be 'young' and 'old' *simultaneously* [zugleich] (KrV, B 192). Similarly, one could will a certain maxim as one's subjective action principle on one occasion, and as an objective universal law in other circumstances. In such cases, the agent could be said to will the maxim and to will the maxim as a universal law without willing both *simultaneously*.

33 This formulation may seem a bit wordy, but it has the crucial advantage of eliminating the ambiguity in Kant's condensed formulation; shortening it would reintroduce the very ambiguities I am aiming to dispel.

34 Rousseau 1997 [1762a], 1.7; see also 1979 [1762b], 462.

This will contains both the merely private and also the universal/general³⁵ will, or the human being regards himself as simultaneously in agreement with the universal/general will.³⁶ (HN, AA 20: 145)

Furthermore, Kant writes that morally impermissible willing involves a 'self-contradiction' of the will:

The will of human beings would contradict itself if they willed something which on the basis of the common will they would abhor.³⁷ (HN, AA 20: 161).

Using the example of theft, he describes this self-contradiction as follows:

From a private perspective I will something and from a public perspective I reject the same thing.³⁸ (HN, AA 20: 161)

In these passages, Kant does not articulate the contradiction as residing in the conception of the general will, as consisting in a contradiction between the general will and some substantive value, as willing the thwarting of one's private purpose, or as the impossibility that all others also will what one wills privately. Rather, Kant clearly locates the self-contradiction *between the particular will*, or the 'merely private' perspective, on the one hand, *and the universal/general will*, or the 'public' perspective, on the other. Of course, there are important differences between these early notes and Kant's later moral theory, if only because he does not yet formulate his point in terms of maxims and laws. Kant still uses similar language some twenty years later, however, in the Mrongovius Lectures on Moral Philosophy, reportedly saying that "[m]y private will often does not harmonize with my will taken as a universal/general rule"³⁹ (V-Mo/Mron II, AA 29: 627).

35 Note that the German word '*allgemein*', which in the context of Kant's moral theory is usually translated into English as 'universal', also means 'general'. I here include both translations, so as to make visible the connection between the Rousseauian conception and Kant's notions of 'allgemeiner Wille' and 'allgemeines Gesetz'. This connection is obscured by the usual translations as 'universal will' and 'universal law'. These translations are otherwise entirely appropriate, however, given Kant's own distinction between 'general' and 'universal' rules (cp. KpV, AA 05: 36).

36 "Diese Willkühr enthält nun so wohl den bloß eigenen als auch den allgemeinen Willen oder es betrachtet sich der Mensch zugleich *in consensu* mit dem allgemeinen Willen."

37 "Contradiceret hominum voluntas sibimet ipsi si vellent quod ex voluntate communi abhorrent."

38 "[...] idem secundum privatum volo et secundum publicum aversor".

39 "Mein Privat Wille stimmt mit meinem Willen als allgemeine Regel genommen oft nicht überein."

The fact that Kant locates the relevant self-contradiction of the will between the ‘private’ and the ‘general’ or ‘universal’ perspective is important. For this suggests that the self-contradiction of which Kant speaks in connection with the FUL should be understood in a parallel fashion, as a contradiction between willing a maxim as one’s own private maxim and willing it as a universal law.

If we understand the relevant contradiction in this way, however, then omitting the simultaneity condition from the description of the FUL – thereby reducing the latter to the universalizability requirement – means losing its very point. For on the reading I propose, the point of the FUL is that one ought to be able *both* to will a certain maxim *and*, simultaneously, to will this maxim as a universal law, without volitional self-contradiction. If one reads the FUL as requiring merely the universalizability of one’s maxims – as the leading interpretations have it – then one loses one side of this self-contradiction. Unsurprisingly, it then becomes hard to provide a satisfactory account of the contradiction without introducing auxiliary assumptions, such as substantive values, an account of ‘essential purposes’ of the will, or an account of the nature of human rational agency more broadly.

Consequently, the first important advantage of the Volitional Self-Contradiction reading of the FUL is that it gives a straightforward reading of Kant’s repeated claims that impermissible maxims yield a *self*-contradiction of the maxim and a *self*-contradiction of the will. For on the Volitional Self-Contradiction interpretation, in the case of a failing maxim, willing the maxim as one’s own maxim contradicts willing the maxim as a universal law. In this sense the maxim (*qua* maxim of the agent) *contradicts itself* (*qua* maxim of all). In the case of maxims that fail the FUL requirement, one would will a certain maxim as a universal law but reject it as one’s own maxim (at any rate in the case at hand); or, conversely, one would will it as one’s own maxim but reject it as a universal law. In fact, this is exactly what Kant says. He writes that “in every case” of a violation of a duty, if we looked at ourselves from a rational perspective,

[...] we would encounter a contradiction in our own will, namely, that a certain principle objectively should be necessary as universal law and yet that subjectively it should not hold universally but allow exceptions.⁴⁰ (GMS, AA 04: 424)

We should not misunderstand Kant as saying here that immoral *actions* are contradictory. He explicitly denies this (GMS, AA 04: 424). The self-contradiction emerges in moral reflection on the permissibility of our *maxims*, and it is a

40 “[...] so würden wir einen Widerspruch in unserm eigenen Willen antreffen, nämlich daß ein gewisses Princip objectiv als allgemeines Gesetz nothwendig sei und doch subjectiv nicht allgemein gelten, sondern Ausnahmen verstaten sollte.”

self-contradiction of the *will*. The men in Kant's examples pause, before acting, to ask themselves whether their maxims are against duty (GMS, AA 04: 421–423). Kant's claim is that they learn, in the course of their moral reflection, that willing the maxim as their own principle is incompatible with (simultaneously) willing it as a universal law.

A second important advantage of the Volitional Self-Contradiction interpretation is that it makes it possible to read the contradiction at issue in an ordinary sense of 'contradiction', namely, as willing A and simultaneously willing not-A. To explain, let me start with a simple example of a volitional self-contradiction concerning an action: willing that I eat chocolate while simultaneously willing that nobody eats chocolate. Note that the content of the second volition in this example is that *nobody* eats chocolate, not that *nobody else* eats chocolate. It is contradictory to will that *I myself* eat chocolate (= willing A) and simultaneously to will that *nobody* eats chocolate (= willing not-A). Similarly, volitional self-contradictions can concern *maxims*. For example, if I will that *I myself* act on a certain maxim (= willing A) while simultaneously willing that *nobody* acts on this maxim (= willing not-A), this similarly constitutes a volitional self-contradiction. For this too amounts to willing A and willing not-A.

We should now examine more closely how volitional self-contradictions are *generated* in moral reflection. Kant's account of their emergence includes several elements that I first present here, before providing textual evidence of their role in Kant's argument concerning two examples. First, what Kant seems to suggest is that one should spell out the *necessary and immediate implications* of the maxim's becoming a universal law. He does not describe the task as one of imagining the empirically probable (but contingent) consequences over time, on the basis of one's past experience. Rather, one is to construct an idea of a world in which the maxim is a universal law, purely on the basis of the content of the maxim. Given this idea of what it would mean for the maxim to be a universal law, one can then determine whether willing this is compatible with willing the maxim as one's own individual action principle. Kant makes two further assumptions that are worth mentioning. He assumes that in the world in which the maxim is a universal law, this law is *known* to be such. Furthermore, he uses the premise that if one knows that x necessarily and immediately implies y, then one's willing x entails one's willing y. On this premise, if the universalization of one's maxim entails the impossibility of acting on the maxim, then willing the maxim as a universal law equals willing the impossibility of acting on the maxim.⁴¹

41 Kant uses a related premise in his argument for the thesis that it is a duty to promote the high-

On the basis of these three elements, we can see how a volitional self-contradiction emerges in cases in which universalizing a maxim makes acting on it impossible. If the universalization of a given maxim necessarily and immediately makes acting on this maxim impossible, then willing that the maxim be my individual principle and simultaneously willing that this maxim be a universal law constitutes a volitional self-contradiction. For then it amounts to willing *that I act* on the maxim and simultaneously willing *that nobody acts* on the maxim (or, alternatively, to willing *that I can act* on the maxim while willing *that nobody can act* on it). In other words, it comes down to willing A and simultaneously willing not-A. Kant's account of the volitional self-contradiction generated in the case of the maxim of false promising exemplifies this argumentative pattern, as I show below.

Volitional self-contradictions may also emerge in cases in which universalizing the maxim does not make acting on it impossible, however. For in the case of some maxims, it is contradictory to will the maxim as one's own and simultaneously to will that it be a universal law, simply by virtue of its content. For example, one cannot coherently will to give limitless priority to one's own interests over those of others, as a matter of principle, and simultaneously *will* a world in which one's interests are completely subordinated to those of others acting on the same principle. Note that in this world others are not merely *trying* to subordinate one's interests to theirs; they are *actually* subordinating them. One can coherently will to act on this principle oneself, and one can coherently conceive of a world in which all agents act on the principle, but one cannot coherently will to act on this principle oneself and simultaneously will that all do. Absolutely subordinating the interests of all others to one's own (on one's own maxim) and simultaneously *willing* that one's own interests be absolutely subordinated to those of all others (in willing that one's maxim be a universal law) is contradictory. Kant's account of the problem encountered in the case of the maxim of egoism, discussed below, follows this pattern.

Accordingly, we can distinguish two different ways in which volitional self-contradictions are generated when maxims are subjected to the FUL test. In a first set of cases, the maxim cannot coherently be a universal law, which makes it self-contradictory to will the maxim as one's own and simultaneously will it as a universal law, because this means willing that one act on a certain maxim and simultaneously willing that nobody does. In a second set of cases, by contrast, it is possible to act on the maxim in a world in which it is a universal law, but given

est good; see Kleingeld 2016, 44–47. Whether the premise commits Kant to the view that willing x entails willing all that is logically implied by x is a different matter that I do not pursue here.

what is contained in the idea of a world in which the maxim is a universal law, it is nevertheless self-contradictory to will the maxim and simultaneously will it as a universal law. In other words, although it is possible to conceive of such a maxim as a universal law, willing the maxim as a universal law nevertheless contradicts (simultaneously) willing it as one's own maxim.

This distinction between two sets of maxims matches Kant's distinction, in the *Groundwork*, between the first and second pairs of the four famous examples. The first pair – Kant's suicide and false promising examples – belongs to the first set. The second pair of examples – Kant's examples of the maxim of not developing one's talents and the maxim of egoism – belongs to the second set.

2.3 Two illustrations

To illustrate that it is both possible and fruitful to analyse Kant's texts along the lines sketched above, let me discuss Kant's examples of the maxim of false promising and the maxim of egoism in more detail.⁴²

The maxim of false promising

In the first set of cases, the action specified in the maxim would be impossible if the maxim were a universal law, and Kant assumes that this is true of the maxim to borrow money and promise falsely to repay it, when one believes oneself to be in need of money (GMS, AA 04: 422). He spells out the necessary and immediate implications of turning the maxim into a universal law, arguing that, in a world in which this maxim were a universal law, making false promises in order to get money would be impossible. In such a world, he states, “no one would believe that he had been promised anything”,⁴³ and the alleged promises of people asking for loans would be regarded as “vain pretense” [eitles Vorgeben] (ibid.). In other words, if the maxim is a universal law, acting on the maxim is impossible. As a result, willing to act on the maxim oneself contradicts willing that the maxim

⁴² Not all of Kant's examples are equally felicitous illustrations. Regarding the suicide example Kant writes that *nature* would contradict itself if the suicide's maxim were a universal law, because a *feeling* that is meant to further life would destroy life (GMS, AA 04: 422). This description of the relevant contradiction seems to differ from the terminology of the FUL and from the terminology Kant uses in his discussion of the other examples, where he describes the contradiction in terms of the *maxim's* or the *will's* contradicting itself.

⁴³ “[... indem] niemand glauben würde, daß ihm was versprochen sei [...]”.

be a universal law. Thus, this maxim “can never hold as universal law of nature and harmonize with itself, but must necessarily contradict itself” (ibid.).⁴⁴ In the literature, this lack of ‘harmony of the maxim with itself’ is usually read as being located entirely within the world of the universalized maxim. This is not the only way in which the passage can be read, however. For it can also be read along the lines I have sketched above, namely, as referring to the lack of ‘harmony’ between the agent’s private maxim, as conceived *prior* to its universalization, and the maxim when conceived as universal law.

As is often noted in the literature, Kant clearly assumes in this example that the maxim would be *known* to be a universal law; that is, he assumes that in the world in which the maxim is a universal law everyone knows that repayment promises on the part of people asking for loans are false. Kant indeed makes this assumption explicit elsewhere with regard to the related maxim of not keeping one’s promises. According to the *Naturrecht Feyerabend* lectures on natural law, Kant argues that in a world in which this maxim is a universal law, one knows both that one does not keep one’s own promises and also that others know this. Under these conditions it is impossible to make genuine promises (V-NR/Feyera-bend, AA 27: 1326; cp. also GMS, AA 04: 403). This is a necessary and immediate implication, not a matter of the gradual erosion of the practice of promising as a result of a gradual decline in trust. One simply *cannot* make a real promise if both the promisor and the promisee know that the promise will not be kept. Similarly, in the case of the false promising example in *Groundwork* II, those asking for money would (try to) make false promises to repay their loans, but none of the people with money to lend would take them seriously. Indeed, as Kant puts it, no one would believe he had been promised anything at all, so in this world one cannot get a loan by making a false promise to repay it. Therefore, willing the maxim as one’s own action principle and simultaneously willing the maxim to be a universal law amounts to willing *that I* obtain easy money by making false promises (through acting on my maxim) while willing *that nobody* obtains easy money by making false promises (in a world in which this maxim is a universal law). This is a genuine volitional self-contradiction.

It is important to note what this volitional self-contradiction does *not* consist in, so let me contrast this account with the interpretations offered by Hegel, Korsgaard, and O’Neill. First, the contradiction is not a contradiction between my universalized maxim and a presupposed value, such as the value of truthfulness or the social utility of the practice of promising. The contradiction is a contradiction

44 “[Da sehe ich nun sogleich, daß sie] niemals als allgemeines Naturgesetz gelten und mit sich selbst zusammenstimmen könne, sondern sich nothwendig widersprechen müsse.”

'with something' – as Hegel puts it – but not with something entirely *different*. From the perspective of the account I propose, what Hegel fails to see is that the relevant contradiction is a contradiction between willing the maxim as one's own action principle and willing the very same maxim as a universal law. Generating this contradiction does not require the presupposition of substantive values.

Furthermore, on the reading proposed above, the *relevant* contradiction – the volitional self-contradiction – is not the inconsistency in the *conception* of the maxim as a universal law. On the Volitional Self-Contradiction account, the fact that one's purpose is thwarted or that one's project is impossible when the maxim is a universal law does not constitute the relevant contradiction but rather produces it. We cannot will the maxim and simultaneously will that it be a universal law *because* the maxim cannot be a universal law. For willing the maxim as one's own maxim and simultaneously willing the maxim as a universal law would mean willing to act on the maxim oneself and simultaneously willing that nobody – again, also excluding oneself – acts on the maxim. (Incidentally, if this interpretation of the relevant contradiction is correct, it implies that what is known in the literature as the 'contradiction in conception test' is actually a misnomer. I discuss this terminological implication in section 3.)

The maxim of egoism

In the second set of cases, acting on the maxim is possible in a world in which it is a universal law – that is, the maxim can be a universal law. Nevertheless, the content of the maxim makes it impossible, without volitional self-contradiction, to will to act on the maxim and simultaneously to will that the maxim be a universal law. One of Kant's examples here is the maxim of egoism. Kant does not actually formulate the maxim at issue explicitly. He describes only the egoist's "manner of thinking" [Denkungsart] (GMS, AA 04: 423). The egoist does not want to take anything away from others, but he wants to pursue merely his own ends and never help others who need and desire assistance. Kant also asserts that the egoist wishes [wünscht] to be helped when he needs [bedarf] it and that he hopes that others will assist him (GMS, AA 04: 423). It is clear, therefore, that this person is not someone who is committed to an ideal of individual self-sufficiency. Rather, Kant seems to have had a very ordinary kind of egoist in mind, namely, someone who refuses to help others but who wants others to help him when he is in need of assistance. Thus, although there are different ways of capturing this set of commitments in a maxim, the maxim involved must be something like 'I shall never help others in need of assistance, but I shall want others to help me when I need their help to achieve my own ends'. Kant assumes that a world of egoists acting

on this maxim is not impossible – and indeed, in such a world no one would ever help anyone else, but this would not prevent egoists in need from wanting that others help them. It is clear from the content of his maxim, however, that it is contradictory for the ordinary egoist to will his maxim and simultaneously to will a world in which no one ever helps others in need of assistance. After all, he wants others to help him when he needs it to achieve his ends. Therefore, it is self-contradictory for him to will this maxim as his own and simultaneously to will that it be a universal law.⁴⁵

A few further clarifications are in order. First, this volitional self-contradiction is generated regardless of whether one needs help at the moment of deliberation. In fact, Kant stipulates that the egoist in the example is not in need of anyone's help (GMS, AA 04: 423). This should not be puzzling, since what is at issue is not a particular egoistic *action* in specific circumstances but the adoption of the *maxim* of egoism. As I mentioned above, maxims are *principles* of action. They are general policies adopted for the long haul.

Second, the previous point may help to explain Kant's use of the term 'wish' in his discussion of the example. He describes a wish as a desire that does not currently lead to action ("inactive desire" [unthätige Begierde], V-Met-L1/Pölitz, AA 28: 253, cp. MS, AA 06: 452) and as a "striving to be the cause of the reality of the object while being conscious of the insufficiency of our powers"⁴⁶ (V-Met-K2/Heinze, AA 28: 743, cp. MS, AA 06: 213). In both senses, Kant's use of 'wish' is apt here. The egoist in the example does not actively desire the help of others because he is not in need of it at the moment of deliberation; and were he to need and seek their beneficence on some future occasion, this would mean that his own powers would be insufficient.

Third, a potential worry needs to be addressed, namely, that the impermissibility of the maxim of egoism implies that any form of competition is morally impermissible.⁴⁷ In trying to win a game, for example, I give priority to my own interests over those of others. Does it follow from Kant's rejection of the maxim of egoism that it is also morally wrong to take part in competitive sports and games?

Kant was not opposed to competition, and he repeatedly praises games for providing relaxation and cultivation of the mind through "a peaceful battle"

⁴⁵ Kant's argument, as reconstructed here, does not suffice to rule out the permissibility of all forms of principled refusal to help others. For example, the maxim of refusing to *give or accept* help would not run into the same difficulties, though it may well run into others. My reconstruction here concerns only Kant's own example.

⁴⁶ "[...] das Bestreben [...], Ursache von der Wirklichkeit der Gegenstände zu werden, aber mit dem Bewußtsein der Unzulänglichkeit unserer Kräfte [...]"

⁴⁷ I thank two anonymous referees for the *Kant-Studien* for raising this issue.

[einem friedlichen Kampfe] (Anth, AA 07: 152; cp. 232). When he describes someone who cheats in a game as feeling self-contempt, this self-contempt concerns the cheating, not the playing (KpV, AA 05: 37). But how does he reconcile his endorsement of competitive games with his rejection of the maxim of egoism?

The answer is found in Kant's designation of games as *peaceful* battles. He conceives of world peace as allowing international competition within the constraints of a rightful global legal order (IaG, AA 08: 26, ZeF, AA 08: 367). In light of this conception of peace, there is a crucial difference between the maxim of egoism in Kant's example and the maxim of someone playing a game. As a player of a game, one pursues one's self-interest subject to the consent of one's opponents and within the parameters set by the rules of the game. Pursuing one's self-interest is restricted to trying to run faster than one's competitors, trying to design the smartest strategy to move one's pieces across a board, and so on. Moreover, doing so is limited by a set of rules. One is not permitted to take shortcuts, knock one's opponent unconscious, and so on, and violation of such rules warrants sanctions or disqualification. Thus, Kant's point in speaking of a 'peaceful' battle seems to be that *within proper limits* competition and the pursuit of self-interest are morally permissible. Whether the limits set by any particular game are proper depends on whether the rules of the game themselves are justifiable. Given Kant's argument against drunkenness, for example, one would expect him to be morally opposed to competitive drinking games (MS, AA 06: 427); also, one would expect him to oppose 'extreme' games in which contestants are allowed intentionally to maim or kill each other. But on the condition that the pursuit of self-interest is properly limited, competitive games are indeed permissible. There is no volitional self-contradiction involved in trying to win (morally permissible) games, while also willing that all players act on this principle. Indeed, many games cannot even be played without all players trying to win.

Let me end this discussion of the maxim of egoism by contrasting the proposed account of the contradiction again with the accounts provided by Hegel, Korsgaard, and O'Neill. Hegelians may read Kant's example as suggesting that willing a world of egoists contradicts one's commitment to something else, such as one's own long-term interests or the moral value of kindness, and they may complain that this contradiction emerges only on the presupposition of this other substantive commitment. On Korsgaard's interpretation, willing that the maxim of egoism be a universal law involves willing the thwarting of an 'essential purpose' of the will as such (*viz.*, the effectiveness in achieving one's ends). On O'Neill's reading, willing the maxim as a universal law is inconsistent with the background conditions of human agency, such as our vulnerability. This is to say that on each of these interpretations the contradiction is a contradiction between willing the maxim as a universal law, on the one hand, and some specific

volitional content that is different from the maxim itself, on the other. This turns the contradiction into something other than a self-contradiction of the maxim. On the account I propose, by contrast, generating the contradiction does not require any such auxiliary assumptions regarding substantive values, essential purposes of the will, or the background conditions of human agency. Rather, the reason why the maxim of egoism fails the criterion articulated in the FUL is that it would be self-contradictory to will to act on this maxim and simultaneously will that the maxim be a universal law.

In concluding this section, I believe it is safe to say that Kant did not specify the nature of the relevant contradiction clearly enough – otherwise his texts would not have given rise to the current range of interpretations. I have not argued that the standard interpretations have no textual support at all. Rather, I have argued that they fail to capture important elements of Kant’s description of the FUL, and that the Volitional Self-Contradiction reading has significant interpretive and philosophical advantages. Importantly, it makes it possible to take Kant’s language of ‘self-contradiction’ at face value, to do so in an ordinary sense of ‘contradiction’, and to do so without additional presuppositions such as a commitment to substantive values, a conception of essential purposes of the will, or an account of the background conditions of human rational agency.⁴⁸

48 Let me add a remark here concerning the few passages where Kant omits ‘simultaneously’, such as the Formula of the Law of Nature. Of course, these omissions may well stem from common human imperfection on Kant’s part, but there is another possibility as well. In several such passages Kant is referring to maxims that agents already have (as in ‘my maxim’, or ‘the maxim of your action’). If agents have already adopted a certain maxim, then when they ask themselves whether they can will their maxim to be a universal law (or whether they can will that it become a universal law ‘through their maxim’, as Kant puts it in the FLN), they naturally examine the simultaneous compossibility of willing their maxim and willing it as a universal law. If an action principle is turned into a universal law ‘through one’s maxim’, then it is no longer merely one’s maxim but simultaneously a universal law. If this is right, then the simultaneity condition may indeed be tacitly implied in cases where the maxim under consideration is a maxim an agent already has. By contrast, in cases where the criterion is not formulated in terms of maxims an agent has already adopted, as in the case of the FUL itself, the simultaneity condition needs to be mentioned explicitly.

3 Beyond 'universalizability' and 'contradictions in conception'

In light of the Volitional Self-Contradiction account of the FUL, some of the standard terminology used to describe Kant's position – though not used by Kant himself – starts to look less felicitous. In fact, this terminology may be part of the explanation of why the Volitional Self-Contradiction reading has escaped notice as a live interpretive option. The most obvious example here is the use of 'universalizability' as a description of the moral criterion that our maxims ought to satisfy. From the perspective of the account I am defending, this term is misleading at best. It captures merely one aspect and misses the essence of the moral criterion articulated by the FUL. What the Volitional Self-Contradiction account makes clear is that 'willing one's maxim as a universal law' is only one of two sides of a comparison. The FUL actually requires that one be able to *will two things simultaneously* regarding any maxim on which one acts, namely, that the maxim be one's own individual action principle and that the maxim be a universal law.

Another example is the well-known distinction between 'contradictions in conception' and 'contradictions in the will'.⁴⁹ The wording of this distinction – again, not Kant's own – suggests that in the first case there would be *merely* a contradiction in conception and no (self)-contradiction in the will. Yet Kant writes that "in every case" of a violation of a duty, if we examine the maxim at issue in light of the FUL, we find a "contradiction in our own will"⁵⁰ (GMS, AA 04: 424, emphasis added), and the Volitional Self-Contradiction account clarifies why. For on this account, *all* failing maxims generate contradictions between willing the maxim as one's own and willing it as a universal law. In other words, *all* failing maxims generate self-contradictions in the will, or, in the terminology adopted in this essay, volitional self-contradictions.

This is not to deny Kant's distinction between two sets of impermissible maxims. On the interpretation I am defending, this distinction – which Kant presents as tracking violations of perfect and imperfect duties, respectively – corresponds to the two ways in which volitional self-contradictions may be generated.

⁴⁹ Both the term 'universalizability' and the distinction between 'contradictions in conception' and 'contradictions in the will' seem to have been introduced in the twentieth century. Marcus Singer traces 'universalizability' back to G. C. Field's work of the 1920s (Singer 1985, 53). The distinction between 'contradictions in conception' and 'contradictions in the will' has gained wide currency since O'Neill (Nell) 1975; see Timmons 2006, 164.

⁵⁰ "[...] bei jeder Übertretung einer Pflicht [...]"; "[...] einen Widerspruch in unserm eigenen Willen [...]"

In the first set of cases, the volitional self-contradiction is generated by the fact that the maxim cannot be a universal law, because this impossibility makes it contradictory to will the maxim as one's own maxim and simultaneously to will it as a universal law. In the second set of cases, it is not impossible for the maxim to function as a universal law, but given this maxim's content, willing the maxim contradicts willing what is contained in the idea of a world in which the maxim is a universal law.

This interpretation finds further confirmation in the fact that, in his general reflections on the principle of morality, Kant mentions only *one* kind of contradiction. For example, he writes:

*That will is absolutely good [...] whose maxim, if made a universal law, can never conflict with itself. This principle is, accordingly, also its supreme law: act always on that maxim whose universality as a law you can will simultaneously; this is the only condition under which a will can never be in conflict with itself, and such an imperative is categorical.*⁵¹ (GMS, AA 04: 437)

Even the passage on which the distinction between 'contradictions in conception' and 'contradictions in the will' is based is fully compatible with the Volitional Self-Contradiction reading. This passage reads as follows:

Some actions are such that their maxim cannot even be *thought* without contradiction as a universal law of nature, let alone that one can also/still [*noch*] *will* that it *ought* to become one. In the case of other actions, however, that inner impossibility is not found, but it is still impossible to *will* that the maxim [of such actions] be raised to the universality of a law of nature, because such a will would contradict itself.⁵² (GMS, AA 04: 424, emphasis in original)

The passage is somewhat ambiguous, and there are different ways to translate '*noch*', which, in existing translations, is often left out. Yet the second half of the first sentence ("let alone that one can also/still will ...") certainly can – and

51 "Der Wille ist schlechterdings gut, [...] dessen Maxime, wenn sie zu einem allgemeinen Gesetze gemacht wird, sich selbst niemals widerstreiten kann. Dieses Princip ist also auch sein oberstes Gesetz: handle jederzeit nach derjenigen Maxime, deren Allgemeinheit als Gesetzes du zugleich wollen kannst; dieses ist die einzige Bedingung, unter der ein Wille niemals mit sich selbst im Widerstreite sein kann, und ein solcher Imperativ ist kategorisch."

52 "Einige Handlungen sind so beschaffen, daß ihre Maxime ohne Widerspruch nicht einmal als allgemeines Naturgesetz gedacht werden kann; weit gefehlt, daß man noch wollen könne, es sollte ein solches werden. Bei andern ist zwar jene innere Unmöglichkeit nicht anzutreffen, aber es ist doch unmöglich, zu wollen, daß ihre Maxime zur Allgemeinheit eines Naturgesetzes erhoben werde, weil ein solcher Wille sich selbst widersprechen würde."

probably should – be read as referring to the volitional self-contradiction that is generated in the case of maxims that cannot be conceived as universal laws. The wording of the sentence is consistent with the idea developed above, namely, that if the universalization of a maxim necessarily and immediately makes acting on the maxim impossible, then willing the maxim as one's own and simultaneously willing it as a universal law yields a volitional self-contradiction. In short, volitional self-contradictions arise in *both* sets of cases, although they emerge in two different ways.

Concluding remarks

I have argued that there are good reasons to favour the Volitional Self-Contradiction account over the standard interpretations of the FUL. The Formula is usually thought to demand that one act only on maxims that one can will as universal laws. On the interpretation I propose, by contrast, the Formula demands that one act only on maxims that one can will as one's own and simultaneously will as universal laws. This interpretation yields a different account of the nature and locus of the contradiction generated in the case of impermissible maxims. The relevant contradiction is not a contradiction in (or involved in) conceiving or willing a maxim as a universal law but a contradiction between willing a maxim as one's individual action principle and willing it as a universal law.

Compared to existing accounts, the Volitional Self-Contradiction interpretation has significant advantages. It explains the sense in which the relevant contradiction is a '*self-contradiction*'; it makes it possible to describe this self-contradiction in an ordinary sense of 'contradiction'; and it explains how this self-contradiction can be generated without presupposing substantive values, introducing a conception of essential purposes of the will, or developing a broader account of the background conditions of human agency.

If the argument of this essay is convincing, this has important implications for our understanding of Kant's moral theory. Spelling out these further implications lies beyond the scope of this paper, however. I will not, therefore, engage in a discussion of the FUL's application to a wider range of maxims, nor will I discuss the so-called 'false positives' and 'false negatives' the Formula is often alleged to yield.⁵³ Furthermore, it remains to be seen whether the Volitional Self-Contradic-

⁵³ A full account of the practical application of the FUL first requires an in-depth discussion of the notion of a *maxim*. Many allegedly wrong results depend on one's particular understanding of the notion of a maxim, and not just on one's understanding of the FUL. For perceptive discus-

tion interpretation sheds new light on Kant's derivation of the FUL and whether it helps to clarify the relation among the different formulas of the Categorical Imperative. Given the advantages of this interpretation for our understanding of the FUL itself, however, these issues seem well worth exploring.⁵⁴

References

- Translations are my own, but I have benefited from the translations available in the *Cambridge Edition of the Works of Immanuel Kant* (Cambridge, 1992–).
- Allison, Henry E. (2011), *Kant's Groundwork for the Metaphysics of Morals: A Commentary*. Oxford.
- Bojanowski, Jochen (2014), "Rationales Wollen – Über das Verhältnis von Kategorischem Imperativ und Goldener Regel". In: *Philosophie nach Kant: Neue Wege zum Verständnis von Kants Transzendental- und Moralphilosophie*, ed. Mario Egger. Berlin, 211–221.
- Galvin, Richard (2009), "The Universal Law Formulas". In: *The Blackwell Guide to Kant's Ethics*, ed. Thomas E. Hill, Jr. Oxford, 52–82.
- Guyer, Paul (2007), *Kant's Groundwork for the Metaphysics of Morals: A Reader's Guide*. London.
- Hegel, Georg Wilhelm Friedrich (1991 [1821]), *Elements of the Philosophy of Right*, ed. Allen W. Wood, trans. H. B. Nisbet. Cambridge.
- Kleingeld, Pauline (2016), "Kant on 'Good', the Good, and the Duty to Promote the Highest Good". In: *The Highest Good in Kant's Philosophy*, ed. Thomas Höwing. Berlin, 33–49.
- Korsgaard, Christine M. (1996), *Creating the Kingdom of Ends*. Cambridge.
- Korsgaard, Christine M. (2009), *Self-Constitution: Agency, Identity, and Integrity*. Oxford.
- Mill, John Stuart (1985 [1861]), *Utilitarianism*. In: *The Collected Works of John Stuart Mill*, ed. John M. Robson, vol. 10. Toronto.
- Nell (O'Neill), Onora (1975), *Acting on Principle: An Essay on Kantian Ethics*. New York.
- Nyholm, Sven (2015), "Kant's Universal Law Formula Revisited." In: *Metaphilosophy* 46, 280–299.

sions of a range of objections that certain results are 'false', see Allison 2011, 196–203; Nyholm 2015; O'Neill 1989, 83–104; Sensen 2014.

⁵⁴ For helpful comments, I am grateful to Joel Anderson, Katharina Bauer, Jochen Bojanowski, Martin van Hees, Patricia Kitcher, Sem de Maagt, Sven Nyholm, Herman Philipse, Stefano Lo Re, Andrews Reath, Joe Saunders, Irina Schumski, Oliver Sensen, Martin Sticker, Jens Timmermann, Eric Watkins, Marcus Willaschek, and two anonymous referees for the *Kant-Studien*. I also benefited from comments I received when I presented this paper at the University of Groningen, the 2015 Leuven Kant Conference (with comments by Kate Moran), Rainer Forst and Darrel Moellendorf's Colloquium at the Goethe University in Frankfurt am Main, the American Philosophical Association 2016 Eastern Division meeting (with comments by Robert Loudon and Julian Wuerth), and the conference on 'Agency, Persons, and Kant', at the University of Notre Dame. I would like to thank the Netherlands Organization for Scientific Research (NWO) for financial support.

- O'Neill, Onora (1989), *Constructions of Reason: Explorations of Kant's Practical Philosophy*. Cambridge.
- O'Neill, Onora (2004), "Rationality as Practical Reason". In: *The Oxford Handbook of Rationality*, eds. Alfred R. Mele and Piers Rawling. Oxford, 93–109.
- Parfit, Derek (2011), *On What Matters*. Oxford.
- Reath, Andrews (2006), *Agency and Autonomy in Kant's Moral Theory: Selected Essays*. Oxford.
- Rousseau, Jean-Jacques. (1997 [1762a]), *The Social Contract*. In: *'The Social Contract' and Other Later Political Writings*, ed. Victor Gourevitch. Cambridge.
- Rousseau, Jean-Jacques (1979 [1762b]), *Emile: Or on Education*. Trans. Allan Bloom. New York.
- Sensen, Oliver (2014), "Universalizing as a Moral Demand". In: *Estudos Kantianos* 2, 169–184.
- Singer, Marcus G. (1985), "Universalizability and the Generalization Principle". In: *Morality and Universality*, eds. Nelson Potter and Mark Timmons. Dordrecht, 47–73.
- Timmermann, Jens (2007), *Kant's Groundwork of the Metaphysics of Morals: A Commentary*. Cambridge.
- Timmons, Mark (2006), "The Categorical Imperative and Universalizability (GMS, 421–424)". In: *Groundwork for the Metaphysics of Morals*, eds. Christoph Horn and Dieter Schönecker. Berlin, 158–199.
- Wood, Allen W. (1999), *Kant's Ethical Thought*. Cambridge.
- Wood, Allen W. (2006), "The Supreme Principle of Morality". In: *The Cambridge Companion to Kant and Modern Philosophy*, ed. Paul Guyer. Cambridge, 342–380.
- Wood, Allen W. (2008), *Kantian Ethics*. Cambridge.