

“On a supposed right to lie from philanthropy”
Penultimate draft. Forthcoming in *Kant Lexicon*.

Since the format of the *Kant Lexicon* doesn't permit acknowledgement footnotes, I'd just like to seize the opportunity here to thank Krupa Patel, Arthur Ripstein, Sergio Tenenbaum, Shelley Weinberg, Allen Wood, and Julian Wuerth for invaluable feedback on earlier drafts of this entry.

“On a supposed right to lie from philanthropy” (“Über ein vermeintes Recht, aus Menschenliebe zu lügen” (1797) (SRL, 8:425-30/CEPP:611-615) Few writings in the history of philosophy have provoked more controversy than Kant's short essay “On a supposed right to lie from philanthropy”; it is here that we find Kant's discussion of the infamous example of lying to the murderer at the door. Kant's essay responds to a challenge that Benjamin Constant poses to a fundamental principle of Kant's practical philosophy: that lying is always morally wrong. Constant claims that this principle commits Kant to the bizarre position “that it would be a crime to lie to a murderer who asked us whether a friend of ours whom he is pursuing has taken refuge in our house.... It is a duty to tell the truth” (SRL, 8:425/CEPP:611). Such a principle, Constant continues, renders society impossible, because it would give the murderers the right to information that helps them harm innocent others (*ibid.*). Kant's response to Constant raises deep interpretive and philosophical puzzles resulting in the controversial history surrounding this essay.

Kant starts by clarifying that all we can be morally responsible for is being truthful and not for telling others the truth as such; after all, truth is objective and beyond what we can control and subject to our will (SRL, 8:426/CEPP:611). Moreover, Kant explains, the aim is to defend the view that even under conditions wherein one is “compelled by an unjust constraint ... [and lies] in order to prevent a threatened misdeed to himself or to another,” one still does not have “the *authorization* (the right) to be untruthful” (SRL 8:426). Then, after responding that someone who is truthful in response to the murderer's question cannot be held legally accountable for what ensues, Kant argues that “one *who tells a lie*, however well disposed he may be, must be responsible for its consequences even before a civil court and must pay the penalty for them, however unforeseen they may have been... To be *truthful* (honest) in all declarations is ... a sacred command of reason prescribing unconditionally, one not to be restricted by any conveniences” (SRL, 8:427/CEPP:612). Many have viewed this as Kant stubbornly defending the bizarre line of reasoning that Constant is criticizing: One has a duty to be truthful also to murderers who are pursuing innocent victims, and, to make matters worse, if one lies and the lie happens to help the murderers in their terrible pursuits, one is legally responsible for the murderers' subsequent violence.

To see why Kant is not, as is commonly claimed, tripped up in this essay, but rather is expressing a position consistent with his moral writings generally, we need to appreciate how the essay affirms three core ideas of Kant's. First, from the point of view of virtue, lying is always impermissible. Second, Kant maintains both that although no one has the right (authorization) to lie, not all lies are legal wrongs and that from the point of view of private right, lying is wrong only when it deprives others of something to which they already have a right. Finally, Kant believes there is an important distinction to be drawn between “formal” and “material” wrongdoing, which is revealed when we find ourselves in situations like this one where morally good ways are impossible to choose. Hence, if in these situations we choose to fight ‘evil with

evil,' we do not wrong anyone in particular (commit a "material" wrong), but we still do wrong in general (commit a "formal" wrong).

That Kant views lying as always impermissible from the point of view of virtue (first-personal ethics) is indisputable. It's not just, as we learn in the *Groundwork for the Metaphysics of Morals* and *The Critique of Practical Reason*, that we have a perfect duty not to lie, but also, as he argues in *The Metaphysics of Morals*, that lying—understood as "the contrary of truthfulness"—is "The greatest violation of a human being's duty to himself regarded merely as a moral being (the humanity in his own person)" (MM, 6:429 [1797]/CEPP:552). Lying is contrary to a "*formal*" duty to oneself "as a moral being" and "to [one's] inner freedom, the innate dignity of a human being" and it involves making "one's basic principle to have no basic principle and hence no character" (MM, 6:420/CEPP:545). Similarly, in the *Religion within the Boundaries of mere Reason*, Kant views lying as the root of all evil and intimately links self-deception (lying to oneself) to the worst ways in which we can morally lose our way in life (Rel, 6:37/CEPP:83-4).

Returning to his essay on the right to lie, Kant clarifies that his argument is consistent with the analysis of lying from the point of view of virtue (first-personal ethics), according to which "Untruthfulness is a violation of duty to oneself" (SRL, 8:426n/CEPP:612). But, he continues, the essay focuses only on "a duty of right" (ibid.), which is why he avoids using formulations concerning our duties of virtue in the body of the text. For Kant, although virtue (first-personal ethics) and right (enforceable justice) are compatible and complementary, they are not co-extensive (MM, 6:230/CEPP:387). Virtue involves internal exercises of freedom (proper usage of our self-reflective reasoning capacities and actions motivated by these reflections as necessary). As we learn in the *Groundwork* and the second *Critique*, for Kant, to act virtuously is to act on universalizable maxims from the motivation of duty. In contrast, right tracks external exercises of freedom and interaction with others. We exercise external freedom rightfully when we interact in the world (in space and time) in ways reciprocally respectful of one another's innate right to freedom (one's "right of humanity in one's own person") and the corresponding duty of rightful honor (MM, 6:236f/CEPP:392 and MM, 6:240/CEPP:395, cf. MM, 6:213f/CEPP:374-376; MM, 6:218-221/CEPP:383-385). Our innate right to freedom is our right to "independence from being constrained by another's choice... insofar as it [our exercise of freedom] can coexist with the freedom of every other in accordance with a universal law" (MM, 6:237/CEPP: 397), whereas our rightful honor "consists in asserting one's worth as a human being in relation to others" (MM, 6:236/CEPP:392). To understand what Kant is saying regarding right and lying in his short essay on the right to lying, we must first appreciate why Kant thinks that not all lies are legally wrong and why lying is a private legal wrong only when it deprives others of something to which they already have a right.

Kant's main analysis of right and lies is found in *The Metaphysics of Morals*. Here, Kant argues that a human being is "authorized to do to others anything that does not in itself diminish what is theirs, so long as they do not want to accept it – such things as merely communicating his thoughts to them, telling or promising them something, whether what he says is true and sincere or untrue and insincere ... for it is entirely up to them whether they want to believe him or not" (MM, 6:238/CEPP:394). Words do not have coercive (physical causal) powers and hence cannot deprive others of what is rightfully theirs. Rather, the general principle is that if I declare something as truthful and you take me up on my invitation to trust what I have said, I become a co-author for what happens next. Because of my lie, you don't quite know what you're actually doing any longer, which is why I become responsible for bad consequences of my lie. Moreover,

one way successfully to use words so as to deprive others of what is rightfully theirs is contractual fraud, which involves “the false allegation that a contract has been concluded with someone, made in order to deprive him of what is his” (MM, 6:238n/CEPP:394). Contractual fraud is a way of stealing by means of words. Another way to deprive others of what is theirs by means of words is to deny them of their rightful honor, of public recognition of the lives they have lived. Not only are these kinds of speech always ethically wrong, but insofar as they involve “slander” (false defamation) legal suits can be brought to court (MM, 6:466/CEPP:582). In sum, private lies are *legally* punishable when by lying one “violates another’s rights” by (intentionally or unintentionally) taking something that rightfully belongs to that person. In contrast, as we saw above, from the point of view of first-personal ethics (virtue), there are no such limitations; it is never virtuous to deceive oneself or others (MM, 6:429/CEPP:552).

Kant makes it clear that “I indeed do no wrong to him who compels me to make the [lying] statement” (SRL, 8:436/CEPP:612). After all, it’s not just that the murderer neither has the right to my information nor the right to threaten me, but he/she also does not have a legal claim on me requiring that I, as a private person, do not lie in response to a question. Rather, as we have seen, the general rule is that if I make a statement (declaration) to another that I know to be untruthful, then I become legally responsible for its bad consequences. This is why Kant argues in the essay that even “a well-meant lie *can* ... also become by *accident* (*casus*) punishable in accordance with civil laws” (SRL, 8:426/CEPP:612). In contrast, someone who is truthful in situations involving threatening murderers thereby chooses to abstain from the interaction, saving the full responsibility of what follows for the murderer. Of course, there are many ways to change the example such that the reasoning changes too, such as by invoking scenarios where a public court would simply find related lying unpunishable. Nevertheless, the main point is that being truthful in response to threats is not to commit a private legal wrong; it cannot bring legal liability for what happens next.

Let us finally turn to Kant’s main concern in his response to Constant, namely that although lying to the murderer involves no “material” wrongdoing, it does involve “formal” wrongdoing. After clarifying that the murderer is not wronged by the lie (no material wrongdoing occurs), Kant continues by arguing, “I nevertheless do wrong in the most essential part of duty *in general* by such falsification, which can therefore be called a lie (though not in a jurist’s sense)” (SRL, 8: 426/CEPP 612). Similarly, after agreeing with Constant’s general claim that one must never abandon a true proposition, he continues: “But here one must understand not the danger of *harming* (contingently) but of *doing wrong* generally ... and, though by a certain lie I in fact wrong no one, I nevertheless violate the principle of right with respect to all unavoidable necessary statements *in general* (I do wrong formally though not materially)” (SRL, 8:429/CEPP:614f). Consider also this: “Thus a lie, defined merely as an intentionally untrue declaration to another, does not require what jurists insist upon adding for their definition, that it must harm another.... For it always harms another, even if not another individual, nevertheless humanity generally, inasmuch as it makes the source of right unusable” (MM, 6: 426/CEPP:612). Lying is not always to do a legal, material wrong to another, particular human being; a murderer doesn’t have a right to anyone’s helpful, truthful declarations. Nevertheless, when one lies, one commits a formal wrong; lying wrongs everyone (“humanity”) because the action is necessarily inconsistent with a moral, rightful (and, of course, virtuous) world.

One way to make this argument clearer is to notice that Kant makes the same argument at the end of his account of private right in the Doctrine of Right in *The Metaphysics of Morals* – a work published the same year as the on the right to lie essay (1797). There he argues that if a

group of people who are interacting have “the intention to be and to remain in this state of externally lawless freedom [the state of nature],” then they “do *one another* no wrong at all when they feud among themselves... But in general they do wrong in the highest degree by willing to be and to remain in a condition that is not rightful” (MM, 6:307f/CEPP:452). By fighting it out, the interacting persons don’t (materially) wrong one another, but they do wrong in general (formally) by choosing to stay in a condition where justice is impossible. In a footnote here, Kant adds that “This distinction between what is merely formally wrong and what is also materially wrong has many applications,” (MM, 6:307f/CEPP:307), and my suggestion here is that in the essay on the right to lie, Kant aims to show one more situation (lying to the murderer at the door), in which the same distinction (between material and formal wrongdoing) is central to a good philosophical analysis of it.

In *The Metaphysics of Morals*, Kant argues that no one has the right to refuse to enter or leave civil society: one can be rightfully forced to enter civil society, and no one can destroy civil society (return to the state of nature) (MM, 6:307-314/CEPP:452). Correspondingly, the argument in the essay on the right to lie can explain why no one can have a right to lie *to* a public authority or *as* a representative of the public authority. These can be punishable public crimes because public authority is not yet another private person, but a public person—a general will that represents everyone, and yet no one in particular. That is to say, when we undertake an action that attacks public institutions, we don’t commit private wrongs (wrong another private person), but we do public wrongs (wrong everyone and no one in particular). Hence, lying becomes a punishable wrong when done to or by a public authority exactly because the public authority represents everyone and no one in particular; formal wrongs become punishable. For example, lying under oath is perjury (a “crime against justice”) and if one lies as a public official, one can be charged with public wrongs such as corruption of or misconduct in public office.

Does this last point mean that when people act in the name of a state, they can oblige us to tell him/her the truth? To use the most historically prominent response to Kant’s essay on lying to the murderer at the door: If the person at the door is a Nazi soldier searching for Jewish people hiding in my home, must I tell the truth to this soldier? On the above analysis, the answer to this question is a firm “no.” Representatives of horrible regimes, such as Nazi Germany, do not have the right to not be lied to because they do not represent a public authority. Nazi-Germany was a violent effort to deprive some people of their right to freedom, making it what Kant in *Anthropology from a Pragmatic Point of View* calls a “barbaric” regime (A, 7:331/CEA:426 [1798]). Such a regime cannot be seen as exercising public authority since it was not grounded on a legal commitment to representing the basic rights of each. Accordingly, once a lawful regime was re-established in Germany, those taking part were held legally responsible for their violent actions in the name of the barbaric regime, and so guilty of private crimes as well as public crimes (“treason”); to what extent their actions were punishable would depend on the circumstances (of actual choices available and the presence of threats). Finally, we can also account for Kant’s view that lying to murderers—as war heroes did to Nazi soldiers—is never experienced as something morally uncomplicated (rightful or virtuous) to do, but as something profoundly regrettable, as something that comes at an irreducible moral cost: it always involves committing a formal wrong. Being a hero is therefore not something anyone wants to be, it is rather something one subsequently must learn to live with. One must learn to live with how life put one into a situation where there were no morally unproblematic ways out, where with the hope of a better future and as motivated by affectionate love of humankind (“Menschenliebe”),

one chooses to commit actions (lying, killing) one does not have a right – and so cannot feel morally authorized – to do.

Helga Varden
University of Illinois at Urbana-Champaign