



Sticker, M. (2017). Kant's Criticism of Common Moral Rational Cognition. *European Journal of Philosophy*, 25(1), 85-108.
<https://doi.org/10.1111/ejop.12197>

Peer reviewed version

License (if available):
Other

Link to published version (if available):
[10.1111/ejop.12197](https://doi.org/10.1111/ejop.12197)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the accepted author manuscript (AAM). The final published version (version of record) is available online via Wiley at <https://doi.org/10.1111/ejop.12197> . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Kant's Criticism of Common Moral Rational Cognition¹

There is a consensus that Kant's aim in the *Groundwork* is to clarify, systematize and vindicate the common conception of morality. Philosophical theory hence serves a restorative function. It can strengthen agents' motivation, protect against self-deception and correct misunderstandings produced by uncritical moral theory. In this paper, I argue that Kant also *corrects* the common perspective and that Kant's *Groundwork* shows in which senses the common perspective, even considered apart from its propensity to self-deception and without being influenced by misleading theory, is deficient. Critical practical philosophy needs to *set right* agents about the stringency of some of their duties, and agents need to be *made aware* that they *have* certain other duties. I discuss how Kant corrects the common agent's notion of the stringency of the duty to not make false promises, and how Kant corrects the common agent's notion of duties to self. I finally discuss how his critical practical philosophy can become popular and achieve the correction of the common perspective. I stress the role of education informed by philosophical theory for this and contrast it with so called 'popular philosophy'.

It is uncontroversial in the Kant literature that the perspective of philosophically untutored agents is of great importance for *Groundwork* I, which, after all, describes the transition from 'common moral rational cognition' to philosophical cognition (IV:393.3-5). Recently, there has also been interest in the role of common cognition of morality for understanding the different formulations of the Categorical Imperative (Geiger 2015, Sticker, 2015a), for the deduction in *Groundwork* III (Sticker 2014), for Kant's conception of moral education (Sticker 2015b), for the methodology of his practical philosophy (Sticker forthcoming a, forthcoming b), for the Fact of Reason (Ware 2014, Sticker forthcoming c), for Kant's engagement with sceptics (Guyer 2003) and naturalism (Uleman 2010: ch.5), as well as for Kant's entire critical philosophy (see many works by Ameriks, for instance 2000, 2006, 2012). The most extensive discussion of the common perspective in Kant's practical philosophy to date is Grenberg's (2013) recent *Kant's Defense of Common Moral Experience*,

which argues that it is the ‘tragedy of Kant scholarship in the past 200 years’ to have lost sight of Kant’s ‘first intention of his project – the defence of a common approach to ethics’, which articulates ‘more clearly, what was always already in the practical consciousness of everyday people’ (ibid.1).

There is a consensus that Kant’s aim in the *Groundwork* is to clarify, systematize and vindicate the common conception of morality and that his ethics is not revisionist with regard to the common conception of morality.ⁱⁱ The clearest expression of this view we can find in a reply Grenberg (2015: 327-8) makes to Dean Moyar in a critical exchange about her book on the common perspective:

‘Moyar wonders at the end of his comments about my understanding of the relationship of the common to the philosophical, and asks whether we are licensed to “correct common morality with tools of philosophy” [...]. This one is easy: No!! To expand: the whole premise of my book has been that our common experience of morality is the lifeblood for practical philosophy. Unless we had these common experiences, philosophers would not be at all entitled to reflect upon things moral. And they are now entitled to reflect on things moral only to the extent that attentive reflection on our felt experience allows. The work of philosophy is not to correct common moral experience, but only to unpack, clarify and reveal it’.

Grenberg intends this as an exegetical statement about Kant’s practical philosophy, but she also endorses this view as a systematically appealing approach to ethics. Whilst the latter is certainly contentious, her idea that Kant strongly opposes philosophical revisions of the common understanding of morality I take to be the standard view in Kant scholarship.ⁱⁱⁱ

This standard view limits functions critical philosophy can serve for non-philosophers to three. Firstly, critical philosophy can strengthen agents’ motivation or their readiness to obey duty for the sake of duty.^{iv} This can be done by presenting an agent with fictional cases of outstanding moral conduct that reveal duty in its full purity and authority.^v Secondly, philosophy can protect against self-deception, or ‘rationalizing’ (IV:405.14) as Kant calls it,

and against the resulting ‘corruption’ (IV:405.17) of one’s initial rational conception of morality.^{vi} Thirdly, Philosophy can correct misunderstandings produced by uncritical, in particular, popular eudaemonist philosophical theory.^{vii} All of these functions are *restorative*, in the sense that they either strengthen an agent’s already existing moral commitments (the first), or protect against and refute misconceptions that are not part of the agent’s original grasp of morality or the grasp these agents have simply qua rational agents (the second and third).

According to this restorative conception, the best philosophy could do is to restore the original rational capacities and insights of the agent, as they were before the agent entangled herself in rationalizations and became acquainted with misleading theory. Whilst both the restorative and the corrective view, the latter of which is the one I shall be advocating, hold that philosophy should affect certain corrections in agents, the former view wants to correct agents insofar as they deviate from their original understanding of morality, whereas the latter view sees this original conception itself as needing certain corrections. The question at stake is whether the common agent prior to all self-deception and bad theory is infallible or even then stands in need of some correction.

I agree that philosophy can and should have the functions that the restorative view ascribes to it. My problem with the restorative view is that it holds that this is *all* philosophy can and should do. In the present paper, I want to challenge the restorative view, according to which Kant *only* clarifies, systematizes and vindicates the common perspective. I argue that he also aims to *correct* this perspective. In some situations restoring an original conception of morality might be counter-productive, since this conception itself needs correction. I will show that there are two more functions for practical philosophy: Agents need to be *set right* about the stringency of some of their duties, and they need to be *made aware* of some of their duties. In a first section of my paper, I will discuss how Kant corrects the common agent’s

notion of the stringency of the duty to not make false promises. In a second section, I will discuss how Kant corrects the common agent's notion of duties to self. In a third section, I will discuss why Kant does not make his departures from the common perspective explicit, how his practical philosophy is supposed to achieve the correction of the common perspective and what might be the underlying reason for the insufficiencies of the common perspective.

I. Correction of Stringency

I begin with two remarks: Firstly, an important assumption for my investigation is that *Groundwork* I describes the common conception of morality, albeit with philosophical commentary, whereas *Groundwork* II offers a philosophically systematized conception of the themes and insights of *Groundwork* I. Differences between Kant's developed conception and the common perspective can be made apparent when we contrast specific elements of *Groundwork* I and II with each other. Kant does not present these differences *as* differences. They have to be uncovered in Kant's descriptions of common and philosophical cognition. In sec.III we will see that Kant does not make these differences explicit for a reason.

Secondly, it should be noted that not even the restorative view denies that Kant's ethics is innovative and revisionary with regard to some of the most *abstract* elements of duty. Kant emphasizes that common agents lack '*insight*' into the foundation or source of respect (IV:403.26-7). It takes a critical philosopher to discover that being motivated by a sense of duty is an exercise of autonomy. Kant claims that all previous academic attempts to determine the supreme principle of morality failed, because they conceived of the source of morality as a material good and did not discover that an agent bound by duty is 'subject *only to his own* and yet *universal legislation*' (IV:432.29-30, V:64.6-65.4). Only autonomy, Kant believes, can be a foundation or source of the common notion of duty, a source the common agent is unaware of but already implicitly acknowledges when she judges that actions for the sake of self-

interest lack moral worth.^{viii} These general revisions of agents' self-understanding as moral creatures are not the sort of corrections that I will be focusing on, since they do not affect agents' conception of what their duties are in particular situations and how stringent these duties are.

Groundwork I opens with Kant's famous assertion that '[i]t is impossible to think of anything at all [...] that could be taken to be good without limitation, except a **good will**'.^{ix} Kant admits, however, that this proposition involves something 'strange' that can give rise to 'suspicion' (IV:394.34-5). He believes that if common agents think fully through the competing candidates for unlimited goodness presented at the beginning of *Groundwork* I, they will discover that the good will is the only viable candidate. If not every single agent, then at least their 'common reason' is in 'agreement' with his claim about the unconditional goodness of the good will (ibid.32-6). Kant's opening proposition intends to capture on a highly abstract level the common agents' implicit awareness that morality does not stand in need of any ulterior incentives to be rationally appealing.

Kant argues that for finite human beings the good will is only present as necessitation or duty, since 'subjective limitations and hindrances' (IV:397.8), desires and inclinations, can detract these beings from the right course. In *Groundwork* I, he presents and analyses five paradigmatic cases for commonly acknowledged duties. The first four (IV:397.11-399.26) are supposed to reveal how a sense of the authority of duty affects the common agent's moral assessments. The common agent, for instance, does not consider a shopkeeper's honesty as truly moral as long as 'self-serving purpose' (ibid.397.18-9), rather than a sense of the authority of duty, motivates the shopkeeper's actions.^x The fifth example (IV:402.16-403.17) is presented towards the end of *Groundwork* I as an example for how universality functions as a criterion in ordinary reasoning about duty. What matters for our purpose is that the first

three examples, the honest merchant, impermissibility of suicide and other's happiness as an obligatory end, match duties that Kant later uses to illustrate his philosophical formulae of the Categorical Imperative.^{xi} Example four and five, however, show important differences to exemplary cases discussed in *Groundwork* II. In the present section, I focus on the case of false promises.

In an important footnote in the Introduction to the *Critique of Practical Reason*, Kant claims that his practical philosophy does not aim at advancing a 'new principle of morality', but merely a '*new formula*' for a pre-philosophical principle (V:8.28-33). In *Groundwork* I, Kant states that this principle is already present in ordinary reasoning as a 'standard of its judging' (IV:403.37), or as the notion: 'never to proceed except in such a way *that I could also will that my maxim should become a universal law*' (IV:402.7-9). The way Kant describes the pre-philosophical standard of judging in IV:402.7-9 is in wording almost identical to the Universal Law Formulation (ULF) of the Categorical Imperative, explicitly stated in *Groundwork* II.^{xii} The *Groundwork* I formulation, however, contains a first-personal address ('*I could also will...*'), which emphasizes that the principle is part of *my* moral reasoning about *my own actions*. The common agent does not think of the supreme moral principle 'as separated [...] in a universal form' (IV:403.36, see V:155.16) or as an abstract formula. The ULF in *Groundwork* II, by contrast, is phrased as an *imperative* and makes explicit that the supreme principle of morality is an imperative and a universal and abstract law, which applies to all finite rational beings.

In *Groundwork* I, Kant takes himself to be describing how common agents *actually* reason ('*I always consider*'), and to clarify a principle that common agents 'always actually have before their eyes' (IV:403.36-7, see also *ibid.*402.15).^{xiii} Kant's *Groundwork* I example for common moral reasoning is introduced as a response to the question of whether I may

‘when I am in trouble, make a promise with the intention not to keep it?’ (ibid.16-7).^{xiv} To answer this question

‘in the very quickest and undeceptive way [...] I ask myself: would I actually be content [zufrieden] that my maxim (to extricate myself from a predicament by means of an untruthful promise) should hold as a universal law (for myself as well as for others) [...] Then I soon become aware that I could indeed will [wollen] the lie, but by no means a universal law to lie; for according to such a law there would actually be no promise at all, since it would be futile to pretend my will to others with regard to my future actions, who would not believe this pretence; or, if they rashly did so, would pay me back in like coin, and hence my maxim, as soon as it were made a universal law, would have to destroy itself’ (IV:403.4-17).

In this passage Kant describes how an agent who wants to do the right thing and who lacks philosophical sophistication would reason about the question as to whether it is ever OK to make a lying promise. The agent asks whether she can be ‘content’ (IV:403.5) with the maxim of lying promises as a universal law and whether she can ‘will’ (IV:403.11) this. She realizes that she cannot. In a world in which lying promises were universal ‘there would actually be no promise at all’, since no one would ever believe a promise (IV:403.13-5). The maxim of lying promises ‘as soon as it were made a universal law, would have to destroy itself’ (IV:403.16-7).

It is important for us that Kant here does not only describe the common agent’s verdict, but also the reasoning behind it. He walks the reader through the common agent’s way of reaching her verdict. According to how Kant describes the case, the destruction of the maxim of lying promises is not what carries the normative weight for the common agent. The common agent does not simply judge that lying promises are impermissible, because the corresponding maxim turns out to contain a contradiction revealed in universalization. There is a crucial step in the common agent’s reasoning between the verdict and the contradiction. The common agent rejects lying promises because of what she can be *content with* and *will* if

universalized, not because a universal maxim of lying promises would be self-undermining. Kant introduces considerations pertaining to self-underminingness or logical inconsistencies only after he stated that the common agent cannot will a world of universal false promising and as an explanation for this. According to the way Kant describes it here, the common agent thinks lying promises are impermissible because a world in which no one honoured one's promises is detrimental to her own intention of obtaining money by making false promises (it would make her attempts 'futile' IV:403.13), or, if for some reason people still fell for these schemes, the agent herself would be constantly defrauded by others (they would pay her 'back in like coin' IV:403.15-6). The agent cannot be content to live in such a world.

From Kant's discussion of lying promises after he introduced the ULF and the LNF in *Groundwork* II^{xv} an important difference to the common conception emerges. When universalizing the maxim to make promises without intention of honouring them

'I then see at once that it could never hold as a universal law of nature and harmonize with itself, but must necessarily contradict itself. For the universality of a law that everyone, once he believes himself in need, could promise whatever he fancies with the intention not to keep it, would make the promise and the end one may pursue with it itself impossible' (IV:422.29-35).

Here the step between the verdict and the self-destruction of the maxim is omitted. The philosopher who applies the Law of Nature Formula (LNF) can 'see at once' that the maxim of lying promises must 'necessarily contradict itself'. The justification of the duty to not make false promises does not take a detour via what agents can be content with or will. Those who properly understand the LNF see that the maxim of lying promises, when universalized, is *self-contradictory or self-undermining* and this is what disqualifies the maxim^{xvi}, even if, for some reason, an agent would be content with a world without the practice of promises or a world where she would be constantly defrauded.

The important difference between the way the common agent and the philosopher think of the duty to not make lying promises becomes apparent shortly after Kant revisits the examples in *Groundwork* II. Kant claims that

‘Some actions are such that their maxim cannot even be *thought* without contradiction as a universal law of nature; let alone that one could *will* that it *should* become such. In the case of others that inner impossibility is indeed not to be found, but it is still impossible to *will* that their maxim be elevated to the universality of a law of nature, because such a will would contradict itself. It is easy to see that the first conflicts with strict or narrow (unrelenting) duty, the second only with wider (meritorious) duty’ (IV:424.3-12).

The different kinds of contradictions that justify a duty (contradiction in willing and in thought) also distinguish duties into two different classes of stringency. There is a class of unrelenting or ‘necessary’ (IV:429.15, 29) duties, those duties that derive from a contradiction in thought; and a class of meritorious duties, those duties that derive from the fact that a maxim can merely not be willed as universal.

Kant distinguishes classes of duties first in a provisional way in the *Groundwork* (IV:421.fn.), and later in a definite way in his *Metaphysics of Morals*.^{xvii} In the *Metaphysics of Morals*, Kant provides more details about his distinction between strict and wide duties. Strict duties by themselves specify concrete actions we ought to refrain from no matter what. Wide duties instruct agents to incorporate certain obligatory ends into their maxims. They allow for ‘leeway’ or ‘latitude’ (VI:390.6-7). ‘Latitude’ means that it is up to the agent to discover the concrete means to further an obligatory end. Furthermore, wide duties are *conditional* in the sense that they are only to be exercised if their exercise does not violate a strict duty. In other words, strict duties always trump wide duties.^{xviii} It is controversial whether latitude also extends to the question of *how much* an obligatory end is to be furthered within the constraints of strict duty.^{xix}

The reasoning of the common agent as it is presented in IV:403 strongly suggests that common agents think of the duty to not make false promises along the lines of *wide* duties. They believe the duty is grounded in what they cannot *will* or be content with when universalized. Kant's philosophical treatment in *Groundwork* II, by contrast, shows: The duty is grounded directly in a contradiction in *thought*. Kant believes that a duty's grounding also determines its *stringency*. Being unaware of the true stringency of a strict duty, because an agent does not understand that the duty is not grounded in what she can be content with, might lead an agent to believe that this duty is less important to fulfil than other duties, and leaves an agent unable to account for the special moral standing of this duty compared to other duties.

Before we move on, I will address two objections to my reading.

(i) My claims about Kant's conception of common moral cognition so far hinge on the single example Kant provides in IV:403. One might worry that this example is not supposed to represent how *all* common agents think about lying, but merely how some think.^{xx} I do concede (below in this section) that common agents might differ to some degree in how much they understand about the stringency of their duties. We should bear in mind, however, that in *Groundwork* I Kant intends to present the common cognition of morality and to transition to philosophical cognition. This presumably means that the common cognition he describes here is paradigmatic and that Kant provides examples that he thinks are particularly well-suited demonstrations of this cognition. It is unlikely that Kant would appeal to cases that he considers idiosyncratic and unrepresentative of common cognition. Furthermore, even if it were only some common agents who are afflicted by the problem I outlined, this would still mean that there is a non-restorative function for philosophy, albeit one that is only important for some common agents.

(ii) It might be argued that the common agent is not as confused as I make her out to be. After all, even in *Groundwork* I it is stated that the maxim of lying promises 'as soon as it

were made a universal law, would have to destroy itself' (IV:403.16-7). There is at least an implicit acknowledgement that the maxim is contradictory or self-defeating when universalized. There are three things we should keep in mind, though.

Firstly, *Groundwork* I presents the common understanding of morality alongside a fair bit of philosophical commentary. It is clear that the common agent who applies a universalization test becomes aware that she cannot *will* lying promises as a universal maxim. It is less clear whether the underlying reason for this, that universalizing the maxim leads to a contradiction in thought, is apparent to her, or whether this is Kant speaking and a foreshadowing of the philosophical treatment of the example in *Groundwork* II. In addition, it is possible that some common agents have a better understanding than others of why false promising is forbidden. Some might stop at the realization that they cannot will universal deception, because it would make it more difficult to achieve their goals if everyone made false promises; others, those who have thought more about the issue, will also realize that there is a more fundamental problem with lying promises, namely, that universalization reveals it to be a contradictory maxim.^{xxi}

Secondly, even if the common agent is aware of the contradiction in thought, the immediate reason *why* she rejects a maxim upon universalization is that she cannot will it. It is only insofar as a contradiction in thought would lead to a state of the world that she acknowledges as unsuited to her purposes that she rejects the maxim ('it would be futile to pretend my will to others with regard to my future actions' IV:403.13-4). At the very least, the common agent does not understand that in the case of strict duties what matters is the contradiction in thought, not what she can be content with. This is also reflected in the *Groundwork* I discussion of suicide in which the notion of a contradiction is not even mentioned (IV:397.33-398.7). The common agent acknowledges her strict duties either without any awareness of the role of contradiction in thought or at least unaware of the true normative significance of this notion.^{xxii}

Thirdly, whilst the common agent's understanding of the duty to omit lying promises might be deficient or incomplete, we have to give her some credit: The idea that contradictions in thought can inform us about our duties is not obvious. After all, there is nothing *per se* morally bad about a contradiction (though contradictions of course make for bad arguments). Kant's assumption that a contradiction in conception carries moral significance has been the basis for much opposition to Kant's practical philosophy.^{xxiii} Kant's grounding of *wide* duties, by contrast, expresses the much more intuitive notion that if there is something that we cannot will others (or everyone) to do, this should make us doubt whether it is ok for us to do. The common agent is certainly no fool for sticking to the more commonsensical notion of what she can will when universalized, and the burden of proof is on Kant to show that there is more to morality than what we can and cannot *will*.

Summary: The common agent grounds the strict duty not to make false promises in much the same way as she grounds wide duties: on what she cannot *will* as universalized. She does not understand that there is a fundamental difference between the two kinds of duty and hence she does not understand the differences in stringency. It is only as the result of the philosophical Categorical Imperative formula that clarity about the grounding of the duty not to make lying promises can be obtained. Kant, in *Groundwork* II, revises the common agent's understanding of the stringency of this duty.

II. Correction of wide duties to self

The penultimate example for a duty in *Groundwork* I is of a duty to self. At this point Kant has already made clear that common agents are aware that morality requires certain treatment of oneself (or rather: to omit certain actions, namely suicide) (IV:397.33-398.7). There is something else morality requires with regard to oneself:

‘To secure one’s own happiness is one’s duty (at least indirectly); for lack of contentment with one’s condition, in the trouble of many worries and amidst unsatisfied needs, could easily become a great *temptation to transgress one’s duty*’ (IV:399.3-7).

Human beings naturally ‘have the most powerful and intimate inclination to happiness’ (IV:399.8). However, there are also *moral reasons* to not damage, in pursuit of quick pleasure, capacities agents need for moral actions as well as for pursuit of their happiness (ibid.8-26).

The duty to secure happiness here is an example for a non-strict duty to self,. Outside *Groundwork I*, Kant sometimes seemingly accords securing one’s happiness the status of a *wide* duty to self.^{xxiv} In the *Metaphysics of Morals*, however, he clarifies:

‘To seek prosperity for its own sake is not directly a duty, but indirectly it can well be a duty, that of warding off poverty insofar as this is a great temptation to vice. But then it is not my happiness but the preservation of my moral integrity that is my end and also my duty’ (VI:388.26-30).

One’s own prosperity and happiness are only *indirect* duties, or they are only morally significant insofar as they preserve moral integrity. Indirect duties occupy a peculiar position in Kant’s system of duties. This status becomes most apparent in the *Metaphysics of Morals*’ discussion of the moral standing of animals. Kant emphasizes that humans cannot have duties to non-humans (God or animals) (VI:442.1-25). Not treating animals in cruel ways, however, is ‘indirectly’ part of the rational agent’s duty to herself (VI:443.23-5). ‘Indirectly’ here cannot function as a specification of a distinct class of genuine duties. After all, as Kant said, we do not have duties to animals. According to Timmermann (2006: 298), ‘indirect duty is not even a lesser kind of duty: it is not a species of duty at all’. Indirect duties are prescriptions to create conditions that are auspicious to the fulfillment of direct duty. They

concern ‘the use of *means*’ to duties (ibid.297), and they are not grounded in the Categorical Imperative, but in the need to be able to fulfil our proper or direct duties.

In the *Metaphysics of Morals*, Kant calls it an ‘*amphiboly* in [...] *concepts of reflection*’ (VI:442.21) or an ambiguity in one’s understanding of duty to accord an indirect duty the status of a direct duty. In the dense paragraph in which Kant discusses this amphiboly, he does not specify whether this is usually a mistake of philosophers or of laypeople. Kant simply speaks of it as a mistake of human beings. It seems to be a problem for the philosophically uneducated agent as well as for the philosopher.

The subtle distinction between direct and indirect duties is not part of the common perspective. In the example in *Groundwork* I, the qualification that securing happiness is only duty indirectly is inserted in brackets: ‘To secure one’s own happiness is one’s duty (at least indirectly)’ (IV:399.3). This strongly indicates that the qualification is philosophical commentary. The common agent is aware only that her own happiness matters morally. Kant agrees with the common agent that, firstly, agents owe more to themselves than simply not to kill themselves and, secondly, that the significance of one’s own happiness has to be accommodated by a moral theory. He splits the two aspects, though, and treats them in different works. *Duties to self* are further discussed in *Groundwork* II and in more detail in the *Metaphysics of Morals*. Kant argues that agents cannot will as a universal law of nature the maxim of letting one’s ‘talents rust’. Agents have the wide and direct duty to develop talents and to become ‘a useful human being in all sorts of respects’ (IV:422.37-423.16). It is developing talents not happiness that we morally owe to ourselves (see also VI:385.32-387.23, 391.26-393.10).

Post *Groundwork* I Kant does not seriously entertain the idea that one’s own happiness could be a direct duty, since he is not concerned anymore with explaining the common conception of what duties there are. In later parts of the *Groundwork*, Kant rather engages the

popular philosophers, and in the Second Critique he deals with his academic colleagues and important figures from the history of philosophy, such as the Ancient Schools, Hume, Spinoza and Moral Sense Theorists. Kant, however, makes a concession to the common perspective in the *Doctrine of Virtue's* discussion of stupefying oneself. It is, according to Kant, a direct duty to not deprive oneself of agency by means of drugs or excessive eating (VI:427.1-33). Even here Kant, however, emphasizes that this is no concession to happiness for its own sake, and that one's own happiness can never be a *direct* duty (ibid.5-11).

Kant accommodates the significance of *personal happiness* in his conception of the highest good. The complete good is not simply morality, but virtue united with proportionate happiness (V:110.31-111.5).^{xxv} It is telling that in his discussion of the highest good Kant discusses and criticises ancient philosophical schools not pre-theoretical conceptions of the highest good (V:111.18-112.26, 125.14-127.17). Kant sees the highest good as a particularly pressing problem for academic debate. For questions common agents might wonder about, such as what their duty is in particular situations, we can set aside the highest good (see also VIII:280.5-10).

In the previous section, we saw that common agents do not understand the stringency of some duties. Now we also learn that they are in some cases aware of the *wrong duty*. They do understand that there are wide duties to self and that personal happiness matters, both of which is correct, but they muddle the two, when they think of happiness as a proper duty to self. The only indication in *Groundwork* I that this is incorrect is Kant's remark in brackets that securing one's happiness is only an *indirect* duty. Strangely, Kant does here not clearly state that developing talents *instead of* securing happiness is a direct and wide duty to self. Kant simply glosses over the common agent's mistake and corrects her tacitly.^{xxvi}

It is instructive to see how Kant addresses this issue a few years later in the Second Critique. Kant stresses that his philosophical remarks 'may seem superfluous [...] but they

may serve at least to afford the judgement of common human reason somewhat greater distinctness' (V:36.6-8). The requirements of morality can be 'seen quite easily and without hesitation by the most common understanding' (ibid.28-9). This sounds as if there is no revision of the common cognition of morality taking place in the Second Critique, since what their duty is, is supposedly already apparent to the common agent. However, at the end of this work, in the Doctrine of Method, when he discusses moral education, Kant acknowledges two kinds of moral education. One which is concerned with fostering the right motivation and a genuine moral disposition (V:159.31-160.2). This kind of education receives the bulk of the attention in the Doctrine of Method. Another kind of education is supposed to contribute to a 'sharpening' of moral appraisal and to show how to 'distinguish different duties' (V:159.22-31). Kant is aware that moral education can and should do more than to strengthen motivation. It can correct moral appraisal, more specifically, it can correct our notion of what duties we have or sharpen our appraisal of concrete cases and help us distinguish different duties and their stringencies from each other.

The mistake of the common agent I uncovered in this section is potentially grave. She thinks that her *own* happiness has a higher moral status than it in fact has. It seems to be on the same level as her duty to further *other's* happiness. Both duties are treated by the common agent like wide and direct duties. There are many situations in which we can either contribute to others' happiness or pursue our own. According to the common perspective, it seems that these can be potentially difficult moral cases, in which morality (more exactly: duties of the same stringency) is on both sides. Not only does this make moral deliberation more difficult than Kant believes it is, it can also serve the egoist or eudaemonist to deem herself perfectly moral in her pursuit of personal happiness. She might think that she at least complies with her duty to thoroughly secure her own happiness.^{xxvii}

Matters are even worse if we add the confusion about stringency discussed in the previous section. The common agent might be unaware that the duty not to make false promises is more stringent than other duties, and she might even believe that omitting lying promises and securing one's own happiness are both duties of roughly equal stringency (wide, direct duties). Take the much discussed case of the murderer at the door who wants to know whether I hide my friend. I am aware that a (convincing) lie would drastically increase my friend's chances to evade the murderer – though he could still run into the murderer due to an unfortunate coincidence. Let us bracket what my action would mean for my friend and focus on what it would mean for me.^{xxviii} Depending on how close I am to my friend, it might be reasonable to think that letting the murderer get to my friend will undermine my happiness. A convincing lie would therefore serve to extract myself from a situation that would damage my happiness.^{xxix} For a common agent the situation presents itself as dilemma-like and her decision would come down to the balancing of different supposedly moral goods against each other^{xxx}: The awareness that we cannot will lying as universal vs. my duty to secure my own happiness. Whilst this might be a natural way for many philosophers to think about murderers at doors, it is clearly not how Kant believes we should think about it. He believes that morally there is nothing to be said for lying to the murderer.

Summary: The common agent does get morality right in a very general sense. She does know that lying is bad and that she cannot treat herself any way she likes (even beyond the prohibition of suicide). She does not need to be corrected about these very general issues. Yet, her insights might be insufficient to handle cases such as the murderer at the door. Cases can be difficult for the common agent when different morally salient aspects draw in different directions (the badness of lying vs. what I supposedly owe to myself).

III. Popular Philosophy and Education

Kant does not romanticize or idealize common agents. While he believes that they are not prone to certain mistakes to be found amongst his academic colleagues, he also acknowledges that the common agent is not an entirely reliable judge concerning the stringency of certain duties, the moral significance of personal happiness and of what she owes to herself. Yet, Kant never explicitly states that and in what sense his philosophical theory corrects and revises the common moral cognition. He even contrasts the correctness of common judgements regarding concrete moral cases with the potential of uncritical philosophical theory to mislead (IV:403.23-36, V:35.16-18, 155.12-18).

In this final section, I have three aims. Firstly, I will explain why the common agent and her moral capacities matter to Kant, and why Kant consequently downplays the revisionary aspects of his philosophy. Secondly, I will describe what role Kant's revisions accord to critical philosophy and education. Thirdly, I will close with some indications of what the underlying reason for the insufficiencies of the common agent might be.

In the *Groundwork*, Kant claims that it 'should actually have been possible to presume all along' that cognition of what ought to be done is the affair even of the most common agent (IV:404.8-11). As a justification Kant only states that practical insight is 'incumbent upon everyone' (ibid.9). That Kant sees no need to support this statement with additional arguments suggests that he took it to be the default position of his, enlightened, time.^{xxxix} It was chiefly his reception of Rousseau that led Kant to endorse the egalitarian notion that morality is not only for the educated but for everyone.^{xxxix} Rousseau changed Kant's attitude towards the 'rabble', and 'common workers', and taught him that on a fundamental level the common, philosophically untutored agent is no different from the most educated philosopher and deserves as much respect (XX:44.8-16, 176.1-4). Kant's Rousseauian discovery reflects an 'egalitarian ideal of enlightenment' (Ameriks 2000: 102), namely, the idea that 'what is

expected of, and most significant about, human beings must be in principle equally accessible to all and should not depend on the accident of particular external conditions' (ibid.228 see also A/B:831/860).

At least from the First Critique onwards, Kant explicitly conceives of morality as the most essential aspect of human existence.^{xxxiii} In the field of natural science there remain 'many questions insoluble' (A/B:477/505), whereas '[o]ne must be able to know what is *just* or *unjust* in all possible cases in accordance with a rule, because our obligations are at stake, and we cannot have any obligations to do what *we cannot know*' (A/B:476/504, see also V:37.28-39). Mistakes or ignorance about metaphysics, science or matters of fact are less threatening (if at all) for the dignity of a person than ignorance of morality. People's scientific and theoretical understanding of the world is contingent upon education, and the scientific state of the art (for instance, pre-Newton everyone was in ignorance about the fundamental laws of the world of appearances). This, however, does not undermine the dignity of ordinary rational humans. The common agent still has the capacity to get moral questions right, and living up to the standards of morality is an agent's '*highest vocation*' (V:86.10); a vocation that distinguishes humans from other animals and even 'infinitely elevates' their 'worth as an intelligence' (V:162.18-23, see also ibid.61.32-62.7).

If an agent's highest vocation, his dignity, worthiness to be happy and perhaps his conditions in an afterlife depend on whether an agent lives up to moral standards, the world would be intolerably unfair if some rational agents, due to no fault of their own, were not in principle able to fully live up to these standards, whereas others were gifted with this ability by nature or by a special kind of upbringing. It is one of Kant's innovations, albeit one deeply rooted in the enlightenment tradition, to devise a theory of morality, according to which all finite rational agents, regardless of their social status and level of education, are competent moral agents and, in principle, morally criticisable and able to criticize themselves. That it

could be *difficult* to find out in a concrete situation that a false promise (or a lie) is an impermissible means of preserving one's happiness is very much at odds with Kant's otherwise very optimistic statements that it is within the power of every rational human agent to determine what the moral option is, and that this does not require a philosophical background, 'wide-ranging acuteness' or experience 'with regard to the course of the world' (IV:403.19-20, see also V:27.22, 32.2-7, 36.4-6). Sometimes Kant even claims that 'what *duty* is, is plain of itself to everyone' (V:36.31), and that an agent is 'certain on the spot' (VIII:287.16-7), and 'without hesitation' (V:36.28-9) what her duty is.^{xxxiv}

Kant's reason for not making his correction of the common perspective explicit is a practical or pedagogical concern connected to his egalitarianism and high esteem for the common agent. Kant worries that if he were to declare that the common agent requires some form of philosophical correction or advice, this could be taken to support the idea that the common agent cannot be treated as fully responsible and an equal partner to Kant's well educated readership. This, however, would be the wrong lesson to draw, according to Kant. Instead of adopting a condescending view towards the less well educated, it is incumbent upon the educated and powerful to create a society or culture in which the right kind of instruction is readily accessible to everyone.

In addition, Kant believes that, under current conditions, the educated are not necessarily better off than those who lack all higher or academic education. In fact, they might be worse off. Kant warns that uncritical moral theory, instead of setting agents right, 'can easily confuse' reasoning 'with a host of alien and irrelevant considerations and deflect it from the straight course' (IV:404.27-8). After his discussion of rationalizing and corruption at the end of *Groundwork* I, Kant elaborates in *Groundwork* II on *popular philosophy*, which promises help to agents who are confused about their duty. However, as an institutionalized and widespread form of collective rationalizing popular philosophy, according to Kant, cannot

help agents and it can even reinforce rationalizing. He therefore proposes a transition from popular moral philosophy to a proper philosophical system.

Kant is concerned that for many of his colleagues popularity functions as a means of justifying philosophical claims (IV:388.19-31, 409.15-9).^{xxxv} Founding a philosophy on what is popular results in a ‘mixed’ or eclectic doctrine (IV:388.27, 410.8, 22, 26, 411.3, see also IV:409.30-2, 426.18-9, IX:148.11-8). Popular philosophers piece together elements from virtue ethics, eudaimonism, perfectionism, moral sense, theology, etc., without an underlying and unifying rational principle (IV:410.3-18). Their teachings make an agent ‘waver’ (ibid.411.7) in her reasoning between various kinds of principles.^{xxxvi} Furthermore, an eclectic doctrine is composed of pure and empirical elements. It obscures what is special about moral commands, namely, that they are not conditioned on anything empirical and that the moral law demands unconditional obedience. The educated public has a ‘taste’ (ibid.388.25) for these mixtures, since they give impure principles a veneer of rational justifiability without truly committing an agent to genuinely moral principles. Popular philosophers do not treat duties, such as the duty not to make false promises, as unconditional, since for them everything is conditioned on happiness. Popular doctrines of morality offer a systematization of the insufficiencies exhibited in *Groundwork* I, not a correction. In fact, they even make matters worse, since they treat personal happiness as the most important concern, and strict duties, such as omitting false promises, as conditioned on the promise of happiness.^{xxxvii}

Kant also has a more theoretical objection to popular philosophy. It lacks philosophical depth or content. A philosophy starting from what is popular can never go ‘further than it can by groping with the help of examples’ (IV:412.18, see also A:ix), and hence cannot discover fundamental philosophical principles. Kant’s take on exemplary assessments of moral cases as presented in *Groundwork* I and II differs in two important ways from popular philosophy. Firstly, Kant believes that the philosopher should be critical towards examples (see esp.IV:408.28-33), and not take anything agents say as beyond criticism. Kant for instance

admits that the common agent might not immediately agree with his *Groundwork* I assertion about the sole unconditional worth of the good will (see my section I). He believes, however, that she can be brought to agree when someone illustrates this thesis and provides additional argumentative support as he does when he explains why other candidates for unconditional goodness fail. Secondly, Kant believes that philosophers have to move beyond examples and find underlying principles and a metaphysical framework that shows how these principles can determine actions. Once a principle of ordinary reasoning is made sufficiently clear and is metaphysically vindicated, this principle can be used to critically investigate the common conception of morality and to uncover and correct elements that are not in line with this principle.

It is important that Kant believes that popularity can become ‘true *philosophical popularity*’ (IV:409.27). True philosophical popularity is not a criterion for philosophical correctness but a manner of *presenting ‘in concreto’* (IX:19.17-29) via discussion and analysis of specific cases a theory, which has a justification in a source other than popular assent or majority opinion, (see also VI:376.28-31, IX:100.13-16). Such a presentation can be ‘very commendable’ (IV:409.20) in order to provide a doctrine with ‘*access*’ (ibid.24) to the minds of agents without philosophical training. Kant claims that he intends to present the abstract principle and foundation of morality in the *Groundwork*, and the same material in a popular manner in a future metaphysics of morals (IV:391.34-392.2).^{xxxviii}

Kant is well aware that it will help common agents little if philosophers come to their aid with abstract principles ‘from the arsenal, of metaphysics’ (VI:376.5, see also B:xxxii). In the conclusion of the Second Critique, Kant states that philosophy is supposed to establish a ‘*doctrine of wisdom*’ (V:163.28), which can ‘serve *teachers* as a guideline to prepare well and clearly the path to wisdom which everyone should travel, and to secure others against taking the wrong way’ (V:163.29-31). Only ‘subtle investigations’ in which ‘the public need take no

interest' can clarify and ground a philosophy sufficiently for it to become popular whilst avoiding the pitfalls of a *merely* popular philosophy (ibid.33-5). It is not the task of the critical practical philosopher to directly improve people's dispositions. However, education has to rest on principles; and a moral education must rest on moral principles.^{xxxix} Critical practical philosophy has to find these principles and thus to determine the guidelines for the right moral education.

Implementing these guidelines and thereby popularizing critical philosophy and improving the general public based on its doctrines is the role of correctly informed popular philosophers, as well as teachers, 'popular instruction for children' (VI:14.8-10) and of religious authorities, religious instruction and preaching (VIII:403.5-10). The critical philosopher is supposed to prepare the foundations for the enterprise of moral correction and moral education, and is neither engaged in a purely academic enterprise, nor directly engaged in the application of his research. Kant might not always clearly distinguish between the role of the educator and of the philosopher, since he himself was an educator before he became a full-time academic.^{xl} He believes that philosophy and education can be (and perhaps: should be) done by the same person, albeit in different roles.

Kant believes that it is not the right reaction to the insufficiencies of the common perspective to treat her as less than fully rational and less than fully responsible for her actions. After all, the common agent's mistakes are also the mistakes of uncritical popular philosophers and of the educated public who listens to these philosophers, and who even functions as popular philosophers' standard of philosophical correctness. Kant believes that it is incumbent upon critical philosophers, such as he himself, to improve public education by informing the educators. The common agent never needs to engage with philosophers directly or engage in philosophy herself, in this sense it is true that, as Kant says, the common agent has 'no need of science and philosophy to know what one has to do' (IV:404.5-6). However,

as I hoped to have shown, the common agent does need critical philosophy indirectly to correct public moral education. Kant believes that based on the common agent's acknowledgment that morality is unconditional and that duty is determined by the criterion of universality the common agent can be lead by the right kind of moral education to revise her misunderstandings about the role of happiness and the strictness of her duties.^{xli}

Critical practical philosophy does not merely serve a *restorative* function, as I called it in the introduction of my paper. It is not merely supposed to protect the original pre-philosophical understanding of the common agent and to strengthen motivation, but also supposed to facilitate an education that sets common agents right about the stringency of some of their duties in the first place, and that makes them aware of what they owe to themselves. Without this education there is a danger that agents will not merely fail to be motivated but in some situations even have the wrong idea of what morality requires of them. This danger, as I hope to have shown, is a danger independent of the pervasive threats of self-deception and bad theory. Even intellectually upright agents who could withstand the propensity to self-deception, if these agents exist, can find themselves in a situation in which they make the wrong call regarding what they ought to do morally.

The upshot of my paper should cast doubt on the view that Kant's moral theory aims to justify a *conventional* conception of morality via an unconventional theoretical and metaphysical framework (see Timmermann 2004: xiv for this view). This view needs qualification. Kant is not conventional at least according to how he himself conceives of the *de facto* conventions (among the common as well as the academically educated people) regarding lying and duties to self. Furthermore, my investigation also gives us reasons to re-think the view that Kant's take on morality is less 'revolutionary' than his take on theoretical knowledge (Schneewind 1998: 533, see also Timmermann 2004: xv). I take it, however, that

Kant's correction of the common perspective regarding morality is much narrower in scope than the changes his Copernican Revolution aims to bring about.

Finally, let me briefly address an important question that I cannot go into in any depth here: What is the underlying structure responsible for making common rational cognition require correction by philosophy? Let me quickly address two possible answers that I deem unsatisfactory and then close with a positive suggestion.

Firstly, if my reading is correct, I have uncovered insufficiencies prior to and independent of the kind of corruption that is due to *self-deception and bad moral theory*. Kant describes the two insufficiencies I uncovered in this paper as part of the common cognition of morality and in abstraction from specific views people might have held at specific times. The insufficiencies seem to be part of the reasoning agents engage in qua rational agents and without being familiar with specific theories, doctrines, ideologies, etc. Furthermore, Kant describes them before he introduces the natural dialectic and the threat of corruption in IV:405 as well as popular philosophy in *Groundwork II*, and in his description of the insufficiencies he makes no mention of self-deception or bad theory. These insufficiencies are part of the conception of morality before the natural dialectic necessitates a transition to *Groundwork II*. Self-deception and misleading theory are not responsible for the mistakes I outlined in this paper – though they of course can amplify these already existing insufficiencies, and they produce their very own confusions with which philosophers and educators have to deal.

Secondly, I do not think that a *dialectic* inherent in reason itself is responsible for the insufficiencies I uncovered. These insufficiencies, whilst they require philosophical correction, are not the kind of systematic and dialectical tensions that infest our entire world view and endanger our notion of freedom and of agency, as we find them diagnosed in the

First Critique. Dialectic in Kant's practical philosophy comes in the form of a natural dialectical tension between our sensuous and rational nature and the propensity to resolve this tension via self-deception or rationalizing (*Groundwork* IV:405) as well as in the form of reason's quest for the 'unconditioned totality of the object of pure practical reason, under the name of the highest good' (Second Critique V:108.11-2). I argued above that self-deception, whilst very important for Kant, is not the reason for the insufficiencies of common cognition described in *Groundwork* I.^{xlii} I argued in sec.II that issues pertaining to the highest good are for Kant highly academic matters and do not affect an agent's awareness of her concrete duties. An inherent dialectic as we find it in Kant's practical philosophy is not responsible for the two kinds of insufficiencies I discovered.^{xliii}

Maybe it helps to turn the question on its head: Why should the common agent without training or philosophical education be correct about what her duty is in all cases? This question seems far more difficult to answer than the question why there can be mistakes in ordinary reasoning. The most natural explanation for the mistakes I have uncovered is that morality can at times be complicated even for an agent who wants to do the right thing. This is so, in particular, when a number of morally salient (or seemingly morally salient) aspects are at play and pull in different directions. The insufficiencies in the common understanding of morality are problems of application, and quite possibly the same problems that will lead Kant to later present a *Casuistry* in which he acknowledges that there can be difficult cases that require rational capacities trained 'in how to seek truth' (VI:411.20-1). These cases aim to fine-tune agents' moral capacities when faced with situations in which it seems that different and mutually exclusive options are each supported by morality: for instance, loyalty to one's master and the duty to tell the truth or at least not to lie (VI:431.25-34), or the good of one's country and the prohibition of suicide (VI:423.25-31). Ultimately, the common agent might need correction because morality can be complicated at times, and Kant was aware that even an uncorrupted grasp of morality might not always be sufficient to work out what is required.

Conclusion

I would like to close my paper with an issue for further inquiry. Here is one possible way of reacting to what I have argued: It is not the common agent who needs correction but *Kant*.^{xliv} I have been assuming that Kant got it right and that the common agent, if she deviates from Kant's philosophical position, must be corrected. Once we step back from Kant exegesis, this is, however, not clear at all. After all, the common agent seems to be able to avoid some of the problems that notoriously plague Kant's philosophy, such as that we are under no circumstances permitted to lie to the murderer at the door, or that our own happiness, even our life would count for nothing, if it could only be preserved via a transgression of strict duty. Kant, I believe, has certainly going for his approach that he provides a principle that systematizes our duties or that tells us why we have them and ranks them into a hierarchy of strict and wide duties and indirect prescriptions. This hierarchy offers advice in cases of seemingly conflicting duties. Furthermore, Kant offers an elaborate moral-psychology and metaphysics. This is no mean achievement, yet it might come with the cost of excessive systematicity and of losing track of some of the insights that the common agent maybe did get right after all, such as that it is not always the rational option to undergo every sacrifice for the sake of a strict duty. Maybe in this sense Kant is closer to his 'head-perplexing' academic colleagues than he thought (V:35.14-8): Even he himself can still learn from the common perspective.

This might be in particular true when it comes to the stringency of keeping promises. Kant, I believe, can account for much of what we think of as the normative status of our own happiness in his notion of the highest good, the notion that we ought to preserve our agency and that there are things we owe to ourselves. Yet, with regard to lying and promising it is easy to construe cases in which Kant's ethics is at odds with some of our deeply held

intuitions, and seems excessively rigoristic. This is in particular so given that Kant believes that the stringency of strict duties hinges on contradictions in thought, a notion that does not intuitively seem morally salient. Whilst the *Groundwork* parts that follow *Groundwork* I and the books that follow the *Groundwork* are philosophically richer and more inspiring than anything the common perspective has to offer, one cannot help but wonder as to whether the common perspective in *Groundwork* I offers a more naive but also more intuitive and in some cases more feasible conception of what to do when murderers knock at your door.

Literature

I quote Kant according to the standard *Academy Edition* (volume:page.line). Translations from the *Groundwork* are from Timmermann (2011). Other works by Kant are quoted, with occasional modifications, from the *Cambridge Edition of the Works of Immanuel Kant* edited by Paul Guyer and Allan Wood.

Allison, H. (2011), *Kant's Groundwork of the Metaphysics of Morals. A Commentary*. Oxford, New York: Oxford University Press.

Ameriks, K. (2000), *The Fate of Autonomy. Problems in the Appropriation of the Critical Philosophy*. Cambridge: Cambridge University Press.

– (2006), *Kant and the Historical Turn. Philosophy as Critical Interpretation*. Oxford: Oxford University Press.

– (2012), *Kant's Elliptical Path*. Oxford: Oxford University Press.

Beck, L. W. (1969), *Early German Philosophy. Kant and his Predecessors*. Cambridge, Mass.: Belknap.

Beiser, F. C. (1987), *The Fate of Reason. German Philosophy from Kant to Fichte*, Cambridge, Mass.: Harvard University Press.

Cohen, A. (2014), *Kant's Lectures on Anthropology. A Critical Guide*. Cambridge: Cambridge University Press.

Cummiskey, D. (1990), 'Kantian Consequentialism', *Ethics*, 100: 586-615.

Denis, L. (2001), *Moral Self-Regard. Duties to Oneself in Kant's Moral Theory*. New York, London: Garland.

Freyenhagen, F. (2011), 'Empty, Useless, and Dangerous? Recent Kantian Replies to the Empty Formalism Objection', *Bulletin of the Hegel Society of Great Britain*, 63: 95-118.

- Geiger, I (2015), 'How Are the Different Formulas of the Categorical Imperative Related?', *Kantian Review*, 20: 395-419.
- Grenberg, J. (2013), *Kant's Defense of Common Moral Experience. A Phenomenological Account*. Cambridge: Cambridge University Press.
 – (2015), 'Response to Ware and Moyer', *Kantian Review*, 20: 313-30.
- Guyer, P. (2003), 'Kant on Common Sense and Scepticism', *Kantian Review*, 7: 1-38.
- Hill, T. (1992), *Dignity and Practical Reason*. Ithaca: Cornell University Press.
- Kant, I. (1900ff.), *Kant's gesammelte Schriften*, edited by the Königlich Preußische Akademie der Wissenschaften. Berlin: Georg Reimer.
- Kerstein, S. (2002), *Kant's Search for the Supreme Principle of Morality*. Cambridge: Cambridge University Press.
- Moran, K. (2012), *Community and Progress in Kant's Moral Philosophy*. Washington DC.: Catholic University of America Press.
- Parfit, D. (2011), *On What Matters*. Oxford: Oxford University Press.
- Schneewind, J. (1998), *The Invention of Autonomy. A History of Modern Moral Philosophy*. Cambridge: Cambridge University Press.
- Schönecker, D. (1997), 'Gemeine sittliche und philosophische Vernunftkenntnis. Zum ersten Übergang in Kants *Grundlegung*', *Kant-Studien*, 88: 311-33.
- Sticker, M. (2014), 'How can Common Human Reason confirm the Correctness of the Deduction in Groundwork III?', *Hegel Bulletin*, 35: 228-251.
 – (2015a), 'The Moral-Psychology of the Common Agent – A Reply to Ido Geiger', *British Journal for the History of Philosophy*, 23: 976-89.
 – (2015b), 'Educating the Common Agent. Kant on the Varieties of Moral Education', *Archiv für Geschichte der Philosophie*, 97: 358-87.
 – (forthcoming a), 'Kant on Engaging Other Agents and Observing Reason at Work', *History of Philosophy Quarterly*.
 – (forthcoming b), 'Experiments in Ethics? Kant on Chemistry, Means of Education and Methods of Practical Philosophy', *Idealistic Studies*.
 – (forthcoming c), 'Common Human Reason and the Fact of Reason', in: V. Waibel and M. Ruffing (eds.) *Proceedings of the 12. International Kant Congress Nature and Freedom*. Berlin: De Gruyter.
- Sweet, K. (2011), 'Philosophy and the Public Sphere: Kant on Moral Education and Political Critique', *Idealistic Studies*, 41: 83-94.
- Timmermann, J. (2004), *Grundlegung zur Metaphysik der Sitten. Herausgegeben, eingeleitet und erläutert von Jens Timmermann*. Göttingen: Vandenhoeck & Ruprecht.
 – (2006), 'Kant on Conscience, 'Indirect' Duty, and Moral Error', *International Philosophical Quarterly*, 46: 293-308.
 – (2007), *Kant's Groundwork of the Metaphysics of Morals. A Commentary*. Cambridge: Cambridge University Press.

- (2010), ‘Reversal or retreat? Kant’s deductions of freedom and morality’, in: A. Reath and J. Timmermann (eds.) *Kant’s Critique of Practical Reason. A Critical Guide*. Cambridge: Cambridge University Press.
- (2011), *Groundwork of the Metaphysics of Morals* (Translated by Jens Timmermann and Mary Gregor). Cambridge: Cambridge University Press.
- (2013), ‘Kantian Dilemmas? Moral Conflict in Kant’s Ethical Theory’, *Archiv für Geschichte der Philosophie*, 95: 36-64.

Uleman, J. (2010), *An Introduction to Kant’s Moral Philosophy*. Cambridge: Cambridge University Press.

Ware, O. (2014), ‘Rethinking Kant’s Fact of Reason’, *Philosopher’s Imprint*, 14: 1-21.

Wood, A. (2002), ‘Preface and Introduction (3-16)’, in O. Höffe (ed.), *Immanuel Kant. Kritik der praktischen Vernunft*. Berlin: Akademie Verlag.

ⁱ I am extremely grateful for valuable feedback on earlier versions of this paper to Joe Saunders, Matthias Hoesch and an anonymous EJP referee. Furthermore, I wish to thank Jens Timmermann, Marcia Baron, Ido Geiger, Karl Ameriks and Joe Saunders for discussions about common moral cognition. Moreover, I wish to thank the philosophy departments of the University St Andrews and the University of Göttingen for supporting my research as a PhD student and lecturer. Earlier versions of this paper were presented at the 2015 *Leuven Kant Conference* and in Bernd Ludwig’s Research Colloquium at the University of Göttingen.

ⁱⁱ Timmermann (2010) believes that Kant relies on the common perspective in the Second Critique even more than he does in the *Groundwork*. I am sympathetic to this view but my current investigation is focused on the *Groundwork*.

ⁱⁱⁱ See also the Kantians referenced in the following paragraph as examples for this interpretation.

^{iv} See IV:397.11-399.26, VI:273.15-7. Kerstein (2002: 97-8) argues that since it is, according to Kant, simple to know what duty requires, acting contrary to duty can only be due to lack of motivation. Timmermann (2013: 37) shares this view.

^v See e.g. IV:411, 454.9-455.9, V:155.12-157.6.

^{vi} See Sweet (2011: 86-8) and Wood (2002: 28) who claims that practical philosophy for Kant functions as an ‘antidote’ against rationalizing.

^{vii} The focus on misleading philosophical theory as a danger to common agents’ moral convictions is more prominent in the Second Critique than in the *Groundwork*. Kant there argues that only the ‘head-confusing speculations of the schools [...] are brazen enough to make oneself deaf’ to the voice of reason (V:35.16-18). He even claims that ‘only philosophers’, can make ‘doubtful’ the question ‘[w]hat, then, really is pure morality, by which as a touchstone one must test the moral content of every action?’ (V:155.12-18). Schönecker (1997: 330-2), however, argues that even in the *Groundwork* it is popular practical philosophy, not rationalizing, that corrupts common agents in the first place.

^{viii} Autonomy is revealed as the result of ‘a regressive analysis of what is implicit in a common, pre-philosophical conception of morality and of rational agency’ (Allison 2011: 1.fn.1). Schneewind (1998) emphasizes that autonomy is Kant’s main contribution to moral philosophy.

^{ix} IV:393.5-7, see also V:435.18-9 VI:181.33-4.

^x See also VIII:284.30-33, 403.23-5, IX:493.35-494.4.

^{xi} Kant does admit that with regard to *suicide* he revises the terminological conventions ‘adopted in the schools’, since there are no strict duties to self according to the traditional division of duties (IV:421.fn). Furthermore, Kant’s taxonomy, unlike the traditional one, does not contain any duties to God (see Timmermann 2007: 79). Both revisions, however, do not concern the common agent who knows nothing about the terminology of the schools.

^{xii} ‘[A]ct only according to that maxim through which you can at the same time will that it become a universal law’ (IV:421.7-9).

^{xiii} This of course raises the question: How did Kant know about common agents’ reasoning? In [...] I argue that whilst Kant certainly was no experimental philosopher in our sense of the term, he also should not be understood as a mere armchair thinker. He himself suggests that critical philosophers confront other agents with moral scenarios and this can confirm crucial elements of his theory (see for instance IV:411.fn.). Kant, I believe, was open to empirical input when it comes to learning about the common conception of morality. I cannot discuss this any further here. In [...], I provide textual evidence for my view and address important problems such as whether Kant’s approach to the common perspective endangers the purity of his ethics.

^{xiv} The way in which the notion of *humanity* as an end in itself is present in ordinary reasoning and how this conception finds its philosophical systematization in the Formula of Humanity is an issue I will not discuss here, since the moral significance of humanity is not explicitly mentioned in *Groundwork* I.

^{xv} Immediately after the philosophical formulation of the LNF in *Groundwork* II Kant sets out to ‘enumerate some duties’ (IV:421.21). There is debate as to whether these duties are supposed to be *derived* from the Categorical Imperative or merely supposed to *illustrate* the principle Kant had just introduced. For my argument nothing hinges on this.

^{xvi} There is debate on the exact nature of this contradiction. This need not concern us here. See instead Allison (2011: ch.7) for discussion.

^{xvii} Kant’s distinctions between perfect and imperfect (IV:421.fn), juridical and ethical duties (VI:218.11-221.3, 239.4-12), and duties of strict and wide obligation (VI:390.1-91.25) are intricate and cannot be given full treatment here. See instead Denis (2001: ch.2).

^{xviii} In the *Groundwork*, Kant leaves open whether wide duties also admit of exceptions for inclinations (IV:421.fn.). In the *Metaphysics of Morals* he clarifies that this is not the case (VI:390.9-14). I accept the latter view, though it does not make a difference to my argument.

^{xix} See Cumiskey’s (1990: sec.7) ‘Spartan’ interpretation of Kantian beneficence as demanding to further overall happiness as much as one can, and Hill’s (1992: ch.8) latitudinarianism as the opposite poles of the debate.

^{xx} I am grateful to an anonymous referee for raising this issue.

^{xxi} Kant claims that agents become aware ‘soon’ (IV:403.10) that they cannot will a lie as universal. This means that it cannot require long and complicated reflection, but it also does not necessarily happen instantaneously. There is room for more or less moral reflection.

^{xxii} A reason for not mentioning contradiction in the discussion of suicide in *Groundwork* I could also be that the first four examples are only concerned with *motivation*, not with the question of how duties are *grounded*. This would leave the fifth example, the one I discussed in this section, as all the more significant for this question.

^{xxiii} The most famous case is Hegel’s empty formalism objection that draws on the notion that it is difficult to see how a quasi logical contradiction without additional normative assumptions can ground moral duties. See Freyenhagen (2011) for a recent discussion of the objection.

^{xxiv} For instance V:93.15-9. Even here he qualifies, however: It is only ‘in certain respects’ that one’s own happiness can be a duty.

^{xxv} For more on the conditioning relation between morality and happiness see A/B:808/836-815/843, V:130.11-28, VI:481.16-21, VII:326.11-14, XV:456.20-26, XIX:284.28-285.2, 288.12-14.

^{xxvi} In fact, Kant even indicates in *Groundwork* II after the four examples for the workings of the LNF that the four duties are ‘actual duties, or at least what we take to be such’ (IV:423.36-7). This might not be true for the duty to develop one’s talents. If Kant thinks that this duty is acknowledged by common agents with the same clarity as the other duties, then why does it make no appearance in *Groundwork* I and why does the indirect duty to preserve one’s happiness take its place? Kant might be using the ‘we’ here in the sense of: ‘critical philosophers’.

^{xxvii} That eudaimonists tend to rationalize their pursuit of happiness or look for pseudo-justifications is discussed by Kant in the preface to the *Doctrine of Virtue* (VI:375.1-378.31).

^{xxviii} The most striking problem with the Murderer at the Door is of course how my action would affect a person who trusts me to have his best interests in view. In this sense the Second Gallows Case in V:30 would be a better example for me, since here we have a clear conflict between the strict duty not to lie and one’s *own* happiness. Nonetheless, I think the Murderer at the Door is a more effective example, since it is one of the classical points of controversy regarding the viability of Kant’s ethics.

^{xxix} I switch from lying promises to lying *simpliciter* here. For those who think that the two are substantially different, I could stipulate a case in which the murderer will not believe me, unless I promise him to tell the truth.

^{xxx} Kant of course famously denies the possibility of moral dilemmas. See VI:224 and Timmermann (2013) for discussion.

^{xxxi} According to Schneewind (1998: 4), the ‘deepest and most pervasive difference’ between modern and pre-modern conceptions of morality is that modern conceptions assume ‘that people are equally competent as moral agents, unless shown to be otherwise’.

^{xxxii} Beck (1969: 489) even speaks of a ‘Rousseauistic Revolution’ in analogy to the Copernican Revolution.

^{xxxiii} In fact, we already find this notion in the *Pölitiz* lecture notes from the mid 1770s (XXVIII:261).

^{xxxiv} See also IV:404.2-4, V:35.14-5, VI:14.10, 84.13-4, 181.30-32, VI:273.16-7, VII:200.5-30, VIII:402.13-4. Kant seems to scale back his earlier optimistic statements in the *Metaphysics of Morals*’ casuistry in which he acknowledges that at least with regard to application of duties to concrete cases there can be epistemic challenges (see below).

^{xxxv} In his theoretical philosophy Kant warns of appeal to the masses and to ‘public rumors’ (IV:277.11, see also *ibid.*292.35-293.12). Kant is highly critical of any ‘quick and crude appeal’ (Ameriks 2006: 113-4) to the common perspective.

^{xxxvi} According to Allison (2011: 6, 52-3), Kant’s attacks against popular philosophy are particularly directed against Christian Garve and his *De officiis* translation and commentary (see also VI:206.4-28).

^{xxxvii} This is an appraisal of popular philosophers from a Kantian perspective. For a more positive evaluation see Beiser (1987: 165-9).

^{xxxviii} We get some indication of this popular presentation when Kant in *Groundwork* II illustrates the Categorical Imperative formulations with exemplary duties (see my section I and II) and explains that ‘*access* for the moral law’ can be obtained by presenting different formulations of the Categorical Imperative, since this brings duty ‘closer to intuition’ (IV:437.1-4).

^{xxxix} See IV:411.8-412.14, V:157.28-33, IX:493.35-494.4, 451.12-3, XXV:653.12-27.

^{xl} See X:230.28-231.2, XI:48.26-32, Moran (2012: 142) for Kant’s notion of the relation between the philosopher and the educator. Many essays in Cohen’s (2014) edited volume on Kant’s *Lectures on Anthropology* illustrate how Kant himself in his lectures on anthropology assumed the role of a public educator who teaches material that is ‘relevant, applicable and useful’ (*ibid.*1) for his students. For the importance of moral education for Kant see also IV:404.37-405.37, 410.19-411.7, V:151.8-12, 152.19-153.12, 162.24-163.35, VII:325.5, VIII:26.22-36, IX:441.1-4, 443.19, 446.1-18, 450.27-451.2, 480.7-481.15, XV:647-8, 779.

^{xli} Kant’s claim that the common agent has ‘no need of science and philosophy to know what one has to do’ (IV:404.5-6) is admittedly *prima facie* at odds with my interpretation. However, even on my interpretation the common agent has an understanding of some of the most important features of duty, such as what the status of morality is and how universality functions as a criterion for determining her duties. There are certainly cases for which common agents’ moral cognition is entirely adequate even without critical philosophy. I agree with Kant’s statement as long as it is only understood as a statement about *many* aspects of morality but not *all*. Even philosophers sympathetic to Kant wonder if he puts too much trust in universality as an easy and commonsensical guide to the morally good (see Timmermann 2004: xi, 105), and others dismiss claims such as that ‘what *duty* is, is plain of itself to everyone’ (V:36.31), as ‘overstatement[s]’ (Parfit 2011: 294). I hope to have shown in this paper that there are textual reasons why we should be sceptical of Kant’s claims that the common agent does not stand in need of any help. If we pay close attention to what Kant does in *Groundwork* I and II we see that he himself acknowledges that some help is needed. This, I take it, is a good thing, since it shows that Kant does not romanticize the common agent, that he does not conceive her as infallible, after all there is no argument why she should be (see the beginning of this section), and that he is aware that common agents can be confronted with difficult problems that go beyond their untrained rational capacities.

^{xlii} This, however, raises an interesting question: Would even an agent who is not caught in the natural dialectical tension between a rational and a sensuous nature require philosophical correction? According to my reading, the answer is: Yes. After all, that an entity is not tempted at all to violate morality for the sake of her inclinations does not imply that said entity is infallible when figuring out what the right thing is in a specific situation. There might still be difficult cases even for her. The difficulty would, however, not stem from the deceptive voice of her self-interest. Lack of self-interest does not guarantee a perfect understanding of morality. Of course, paradigms of infinite agents, such as God, are also omniscient and hence they do have a perfect grasp of morality – but this is a feature of their omniscience not of their lack of sensuousness.

^{xliii} It is not even clear that there is a dialectic within pure practical reason. In the Second Critique, Kant first claims that, in contrast to speculative reason, *pure* practical reason does not stand in need of a critique, since its use is always within proper limits (V:15.1-16.12). However, he later acknowledges that ‘[p]ure reason always has its dialectic, whether it is considered in its speculative or in its practical use’ (V:107.6-7).

^{xliv} I am grateful to Joe Saunders for pushing me on this point.