

Durham Research Online

Deposited in DRO:

14 May 2020

Version of attached file:

Accepted Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Saunders, Joe and Sticker, Martin (2020) 'Moral education and transcendental idealism.', *Archiv für Geschichte der Philosophie.*, 102 (4). pp. 646-673.

Further information on publisher's website:

<https://doi.org/10.1515/agph-2020-1010>

Publisher's copyright statement:**Additional information:**

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

Moral Education and Transcendental Idealism¹

Joe Saunders and Martin Sticker

(Penultimate draft. [Final version available at the *Archiv für Geschichte der Philosophie*](#))

In this paper, we draw attention to several important tensions between Kant's account of moral education and his commitment to transcendental idealism. Our main claim is that, in locating freedom outside of space and time, transcendental idealism makes it difficult for Kant to both provide an explanation of how moral education occurs, but also to confirm that his own account actually works. Having laid out these problems, we then offer a response on Kant's behalf. We argue that, while it might look like Kant has to abandon his commitment to either moral education or transcendental idealism, there is a way in which he can maintain both.

Keywords: Kant, Education, Transcendental Idealism

Recently, there has been a surge of excellent work on moral education in Kant. In this paper, we draw attention to important tensions between Kant's commitment to the importance of moral education and his commitment to transcendental idealism. Our main claim is that transcendental idealism makes it difficult for Kant to both provide an explanation of how moral education occurs, but also to confirm that his own account actually works. While these problems might make it look like Kant has to abandon his commitment to either moral education or transcendental idealism, we argue that there is a way in which he can maintain both.

We begin by outlining moral education in Kant (§1). We then sketch the basic contours of transcendental idealism (§2), focussing on Kant's account of (transcendental) freedom, which is outside of time and experience. Having done this, we turn to the tensions between Kant's account of

¹ This paper was presented and received valuable feedback at the *UK Kant Society Annual Conference* in Cardiff, the *British Society for the History of Philosophy* conference in Sheffield, the workshop *Kant in Progress* in St Andrews, and the conference *Kant and Moral Psychology* at Bogazici University Istanbul. We wish to thank the organizers and participants of these events. Furthermore, we thank two anonymous *Archiv* referees for their feedback. Work on this paper was, on the part of Martin Sticker, supported by a two-year Irish Research Council fellowship.

freedom and his account of moral education (§3). We focus upon metaphysical and epistemic problems for Kant: The metaphysical problem is that, in locating freedom outside of time and experience, Kant is unable to offer a satisfying explanation of how education works. The epistemic problem is that, in locating freedom outside of experience, Kant is unable to confirm that his own account of moral education does in fact work. In the final section (§4), we then consider a way in which Kant could mitigate this. We claim that Kant could accept that transcendental idealism precludes him from providing an explanation of how education works and confirming that his own account works, but that nevertheless, still maintain that his is the only account that could work, and have faith and hope that it might.

I. Moral Education

Older Kant literature tends to neglect – or even dismiss – Kant’s conception of moral education.² Recently, however, moral education has received a surge of attention.³ In this section we firstly lay out the various ways in which education is important for Kant’s ethics. And secondly, we provides an overview of the different forms of education that we can find in Kant.

Education, moral and otherwise, is undoubtedly important to Kant. In the *Pedagogy*, he claims that education is the source of all good in the world (IX:448.12–17)⁴, and that education is what turns animals into human beings (IX:441.1–23, 443.19-20). Human beings need education to be in a position to make use of their rational capacities in the first place (see IX:441.1-23, 443.19-20), and they have to be “educated to the good” (VII:325.5–11). In the *Metaphysics of Morals*, Kant claims that parents are blameworthy and can even be punished if they fail to educate their children

² Relatively recently Herman (2007, 132 and 152-3) still has objected that Kant’s ethics lacks anything that could be understood as training for autonomous agency. Furthermore, Grenberg in her book on Kant and common moral experience only dedicates a single page to the issue of education. She acknowledges that “some basic education and some protection from external threats” is required “to realize one’s human capacities”. If these conditions are not fulfilled one could not hold “persons fully responsible for their actions” (ibid.282). We doubt that for Kant agents without any education will be rational enough to be subject of *any* deserved blame (or even of reduced blame).

³ Recent literature on Kant and education includes: Munzel (2002; 2012), Koch (2003), Baron (2009), Moran (2009; 2012, ch.3), Surprenant (2010), Loudon (2011, ch.11), Sweet (2011), Giesinger (2012), the essays in Roth, Surprenant (2012), and in Roth, Gustafsson, Johansson (2014), Vanden Auweele (2015), Sticker (2015), Cohen (2015), and Wehofsits (2016). For the historical context of Kant’s conception of education see Munzel (1999, 254-74; 2012, ch.1-2), Moran (2012, 130-42).

⁴ It should be noted that Kant’s *Pedagogy* is not a text that Kant himself wrote and published, but rather a text put together by another person, Friedrich Theodor Rink, on the basis of Kant’s notes for his lectures on the topic. Discussion of the authenticity of the *Pedagogy* can be found in Weisskopf (1970), Beck (1979, 14–8), and Stark (2000, 97–101). In what follows, we will mostly rely on other works.

(VI:281.13–19).⁵ Moreover, in the *Groundwork*, he motivates the introduction of different formulae of the Categorical Imperative by pedagogical concerns: the different formulae bring the Categorical Imperative closer to intuition, and help the moral law to obtain access to an agent’s mind (IV:437.1–4).

There is a second sense in which education is important for Kant, and this second sense is less obvious and often neglected in the literature, even though it is important for his practical philosophy and the way he engages his critics. Education is not only important for moral improvement, but it also plays a role in establishing Kant’s ethical theory and setting it apart from competing theories. We can see this when Kant addresses criticism by the popular philosopher Johann Georg Sulzer. Sulzer challenged Kant to explain why all practical philosophy so far has failed to morally improve agents. Kant replies (in a footnote in the *Groundwork*) that doctrines of virtue so far have achieved little, because moral educators present a multitude of purported good motives, or a mixed doctrine, instead of actions that abstract from all “advantage” and “alien incentives” (IV:411fn).⁶ He suggests that the “most common observation” shows that:

[...] when one represents an action of righteousness – as it was performed with a steadfast soul, without aiming at any advantage [...] it elevates the soul and stirs up the wish to be able to act like that too. Even children of intermediate age feel this impression (IV:411fn).

Presenting an example of action done for the sake of duty, Kant claims, will affect rational agents, and even children, and they will react in a way that nothing but duty presented in its full purity can bring about. Kant suggests that the way the child expresses or describes her feelings when presented with the moral law in its full purity will support his ethical system over popular contenders, since the child's reaction shows that the supreme principle of Kant’s theory can move agents in ways nothing else can and that this is the only basis for moral improvement. Kant proposes that his theory, if properly presented, affects and moves rational agents.

In the final section of this paper, we will return to consider this claim in more detail. There we will disambiguate between a weak and a strong reading of what Kant claims in the Sulzer footnote (and other similar passages). This disambiguation will turn out to make a big difference for the

⁵ See also V:151.8-12, 152.19-153.12, 162.24-163.35, VIII:26.22-36, IX:446.1-18, 450.27-451.2, 480.7-481.15, and XV:647-8, 779 for further passages that stress the importance of education.

⁶ See VIII:287.22-288.18. For Kant’s emphasis on a pure representation of the moral law as a foundation of education, see V:156.25-27, VI:48.17-33, 83.11-35, VI:217.9-27, VIII:402.21-403.10.

success of Kant's account of moral education. For now, however, we want to continue to say something about how Kant understands moral education.

Kant discusses four different kinds of education.

Firstly, there is *technical education* in which children are trained to achieve their ends in efficient ways, how to interact with people for their own good and the good of their community, and which prepares them to become "useful" members of society (see IX:441.1–23, 443.19–20, 448.12–7). This education is not *moral* education, since it does not aim at developing or training autonomy (in Kant's sense). In technical education, children are brought to develop general skills such as means-ends reasoning, as well as learning the habits and norms of the specific culture they live in. Some of this might prepare them for autonomous agency, but many of the culturally transmitted norms rather regulate the realm of the morally permissible.⁷

Secondly, a genuine form of *moral* education is the kind of education it takes to become a moral agent in the first place. We can call this *basic moral education*. Unfortunately, Kant says relatively little about it. In his ethics, he really only addresses agents who have already completed the transition to full agency and whose commitment to duty must now be strengthened.⁸ The reason for largely neglecting basic moral education might be that its aim is to facilitate a special self-relation, namely, the transition to a state of autonomy.⁹ Completing his transition is ultimately left to the agent herself, and at bottom, "inscrutable" (VI:138.19, see also VI:280.23–5).¹⁰ This appeal to inscrutability is important, and we will return to discuss it in more detail in section III.

Whilst Kant says little about basic moral education, it is important to bear in mind that he is not committed to the bizarre view that human beings develop rational capacities automatically and regardless of all external factors. Of course, moral agency is not something that must be taught in

⁷ In the *Metaphysics of Morals*, Kant, however, declares that being "a useful member of the world" is a duty to self (VI:446.1–2). In this sense even non-moral education has an effect on morality. See Moran (2012, 143–54) for an overview of Kant's conception of non-moral education.

⁸ See, for instance, the *Metaphysics of Morals* Catechism in which a teacher engages a young pupil for the purpose of moral education. It is important that this catechism is "developed from ordinary human reason" (VI:479.7) and that the pupil is said to already be aware that duty is "an unconditional necessitation through a command (or prohibition) of reason, which I must obey; and in the face of it all my inclinations must be silent" (VI:481.31–36). Even the pupil who receives this relatively basic catechistic instruction is already an agent aware of the authority of duty and able to exercise her pure reason.

⁹ By a "state of autonomy" we mean one in which an agent is *under* the moral law, i.e., aware of the requirements of duty and in principle able to act on her moral judgements without external incentives. Such an agent might still act immorally.

¹⁰ Vanden Auweele (2015, 374) has recently argued that Kant believes that it is not "necessary for the moral law to be taught". This is correct, but, and this is an important "but", acquiring use of one's pure practical reason requires a basic form of education and the moral law is only accessible to those who can make use of pure practical reason. Hence the need for basic moral education.

universities or by experts. It is acquired as part of a normal upbringing that all agents, except those in extreme circumstances, undergo (see VI:281.13-19). At the beginning of the *Religion*, Kant expresses the belief that many of his philosophical distinctions are “in different words” contained in the “most popular instruction for children or in sermons, and easily understood” (VI:13.30–14.15). The common use of rational capacities is acquired through education and the educational means to help someone acquire it are relatively easily accessible and wide-spread.¹¹

Since basic moral education is a central aspect of our paper, we will attempt to reconstruct this crucial part of education as well as we can, based on what little Kant does say about the transition to moral agency. Since basic moral education starts with a pupil who can use practical reason only in empirical practical ways, no moral insights on behalf of the pupil can be presupposed. Education must therefore operate via an agent’s self-interest and fear of harm (V:152.21-2). Basic education initiates and guides a transformation from a pupil who first becomes aware that she has the capacity to use reason not only in an empirical practical way to an agent who can assess morally relevant cases according to universal standards and act on these assessments. For this purpose, external rules are enforced via rewards, threats, and punishments, and conformity to rules becomes a “habit” (IX:475.27).

The most comprehensive account of the process of basic education is provided in the pre-critical *Kaehler Lecture Notes*. A youth [Jüngling] of approximately 10 years is already able to employ their reason at least in a prudential (empirical practical) sense, and hence can, for self-interested reasons, conform to rules. Such a youth is subject to external rules, such as parental prohibitions and expectations. From subjection to and familiarity with external rules enforced by his parents and other adults, a youth gains more abstract concepts such as “decency”, “philanthropy”, and “principle” (Kaehler Lectures, p.363).

There is disagreement in the literature as to when Kant developed and introduced his conception of autonomy.¹² If it is introduced earlier than the *Groundwork*, as it is for instance argued by Karl Ameriks (2012, ch.2), then the *Kaehler* passage can be applied to Kant’s critical practical philosophy without accounting for substantial changes in Kant’s fundamental views. Even if the *Kaehler Lecture Notes* from the mid-1770s do not reflect Kant’s mature conception of autonomy, they are not without value, as their starting point is the same as in Kant’s critical period: a young person who begins to

11 See also Moran (2012, 162-3) who argues that the kind of education that is supposed to facilitate the development of moral agency is so basic that almost all human beings will undergo it.

12 The term “autonomy” is used by Kant the first time in lecture notes immediately preceding the publication date of the *Groundwork* (see XXVII:1326.16, XXIX:626.4).

use her reason and who has the potential to moral agency. Subjection to external constraints, experienced in basic education, makes a child familiar with abstract moral concepts.¹³ Through her subjection to universal standards, the agent learns to consider cases in the light of universal standards and not only in the light of what would be of personal benefit. This ultimately enables the agent to make non-self-interest-based decisions and shows her that her reason does not always have to be in the service of inclinations.¹⁴ In Kant's pre-autonomy conception, moral feeling plays the role of the moral incentive whereas in his developed conception respect for the moral law plays this role.

Before we move on to the kind of education that presupposes the capacity to feel respect and seeks to amplify it, it should be noted that there is an important gap between the capacity to defer immediate satisfaction of desires for the sake of a future greater benefit (practical freedom) and the ability to omit or commit actions based on something other than what is of personal benefit altogether (transcendental freedom). (We will discuss this in more detail in sec.III.¹⁵)

Thirdly, Kant warns that basic education might only produce a “*manner of the good*” (A/B:748/776) and even threatens to “corrupt” (IX:475.24) a pupil. It might lead a pupil to believe that external conformity to (parental or societal) rules in order to avoid punishment is all that is morally required. As seen in his reply to Sulzer, Kant sees it as one of the central problems of his time that educators treat rational moral agents like children who need to be bribed and scared into obedience to morality. *Motivational education* aims at strengthening an agent's “receptivity to a pure moral interest” (V:152.33-35, see also A/B:829-30/857-8) with the goal of “cultivating and founding genuine moral dispositions”.¹⁶ Once a pupil's reason is not solely in the service of inclinations

¹³ We stress the role of external coercion, since we are here talking about an individual who does not yet possess the capacity to be motivated by anything but external incentives. Kant himself clearly says that “coercion is necessary” for education (IX:453.30). Shapiro (1999) by contrast rather stresses the role of *play* for children's development on a Kantian framework. La Vaque-Manty (2006) argues that Kant with his pedagogical advice aims to respect the pupil as in some sense already rational and autonomous. We think that this is correct, but only for pupils who are already in some sense autonomous. Pupils must be brought into this state first.

¹⁴ In the third Critique, Kant calls this step “culture of training (discipline)”, which “consists in the liberation of the will from the despotism of desires, by which we are made, attached as we are to certain things of nature, incapable of choosing for ourselves” (V:432.4-8).

¹⁵ Shapiro (1999) makes use of elements of Kant's political philosophy for a Kantian account of becoming an adult. Her project differs from what we are talking about, since she is interested in how agents can “organize the fundamental constituents of [their] motivational world” or how they are able to live a self-determined life (ibid.730). This is rather a matter of empirical practical reason and her account is one that concerns the development of non-moral capacities, not of the transition to autonomy in the Kantian sense of the term. Since she, however, follows Korsgaard's (1996) constructivism, according to which there is a transition from a prudentially well-organized will to acknowledgment of moral commands, these matters are in a continuity for Shapiro. We believe, in contrast to Shapiro, that the step from empirical practical rationality to moral agency needs extra explaining. See also Ludwig (2014) who emphasizes that at least since the *Religion* Kant acknowledges that there can be instrumentally rational agents not under the moral law (see VI:26.26-37).

¹⁶ V:153.11. Kant stresses the motivational impact practical philosophy can have throughout his writings. See IV:405.2-5, 410.19-411 and fn., 412.10-4. V:117.21-4, 151.9-12, VI:217.9-27.

anymore, education should abandon the “leading string” (V:152.23) of self-interest and external enforcement of rules. Instead, the educator should appeal only to an agent’s newly discovered capacity to use reason independently of sensuous incentives (V.156.36–157.6). The “dignity” (IV:411.13) of morality lies in its motivational purity. It is therefore “of the greatest practical importance” (IV:411.18) to present the foundations of morality “pure and unmingled” (IV:411.19). Motivational education can take different forms: The educator can use real or fictional and stipulated examples to train the pupil’s assessment of cases. This helps the agent to distinguish sensuous from pure motives, and thus to gain a better understanding of what makes duty special as a motive (V:154.17-155.1, VI:48.17-33). Furthermore, an educator can present detailed examples of morally good conduct in adverse circumstances and discuss them with the pupil (V:155.12-157.6, VIII:286.8-287.21).

Presumably, this kind of education is not only of benefit to pupils who have just started to make use of their pure practical reason, but also to everyone else, since we, finite rational agents, do not do our duty automatically and always gladly. Motivational education is not simply a stage of one’s moral development that one can leave behind. It can and should be continuous and concurrent with the last form of education.¹⁷

Fourthly, there is the “sharpening” of moral appraisal in how to “distinguish different duties” (V:159.22-31). This is the training of an agent’s capacity of moral judgement or his capacity to recognize moral situations and to correctly prioritize perfect duties over imperfect ones and to balance different obligatory ends the promotion of which admit of latitude. Means for this purpose could be the casuistical questions in the *Doctrine of Virtue* as well as this work’s *Doctrine of Method*. Both offer exemplary cases to train a pupil’s capacity to apply duties to concrete cases. In the *Doctrine of Method*, the Catechism as the “first and most essential instrument for *teaching* the doctrine of virtue” (VI:478.28) is of particular importance for this. The education of moral judgment is again rather a matter of the phenomenal side of things and much like technical education, albeit that an agent here is not trained in pursuing any end but specifically *moral* ends.¹⁸

In summary, Kant has a developed and multi-faceted account of education. He writes of: 1) technical education; 2) education required to become a moral agent in the first place; 3) education

¹⁷ This also shows that Kant does not have in mind fixed stages for moral education. See also Guyer (2014, 5) who emphasizes that, contrary to how many read Kant and many of Kant’s own educational examples that concern children, Kant “like a true moral perfectionist, is clear that the process of moral education is never ending and never ended”. In addition, it is of course possible that the same means of education train both motivation and judgement at the same time.

¹⁸ There is, at least, one other important aspect to moral development in Kant other than the forms of moral education we discussed here, namely, his two-stage model of moral improvement that we find chiefly in the *Religion*. Agents must

that aims at strengthening motivation; 4) and training of moral judgment. Kant thinks that education is important, and that his framework can provide impulses for education and can even correct deficient educational practices; additionally, as we saw in the Sulzer footnote, he thinks that this helps reveal how his philosophy is superior to other approaches.

In the next section, we will sketch the basics of transcendental idealism, before turning to draw out some tensions between this and Kant's account of moral education in section III. In this, our discussion will not focus on the first and fourth kind of education, but the second and third – *basic moral* education and *motivational* education.

II. Transcendental Idealism

Kant views experience as determined by natural necessity. Transcendental Idealism allows him to maintain this conception of experience, but also to accept a libertarian conception of freedom.¹⁹ It does so by locating freedom (in some sense²⁰) outside of experience.²¹

This position has some clear advantages. Kant has found a way to insulate freedom from the empirical. No matter what the natural sciences reveals about our experience, they cannot threaten freedom. Transcendental idealism has thus carved out a safe space for freedom – and along with it, morality.

We are sympathetic to Kant's attempt to preserve a libertarian conception of freedom, but acknowledge that this attempt to *insulate* freedom from the empirical raises problems. In what follows, we want to set out two aspects of Kant's position, namely that: 1) we cannot *experience* freedom; and

undergo a *conversion* (stage one) or make a “single and unalterable decision” to prioritize the moral law over self-love and we can expect *reform* or progress in an agent's empirical conduct (second stage) (VI:47.1–48.16). The latter stage allows agents to make a “conjecture” about their goodness (VI:68.23-4). Even though many of the problems we bring up in section 3 also apply to this specific form of improvement, as here once again, we find the problem of interaction between the empirical and the noumenal, we bracket this model and focus on moral improvement in the form of moral *education*, as opposed to improvement that simply results from the agent's own decision without involvement of an educator. See instead, Papish (2018, ch.7) for a critical discussion of the two-stage model of moral improvement. She raises the problem of how “to explain how a person's intelligible character is meaningfully shaped by a consideration of her empirical character” (ibid.189). She proposes a new interpretation of the distinction between conversion and progress in terms of commitment and cognitive labour to tackle this problem. This interpretation is intended to honor Kant's commitment to transcendental freedom (ibid.194-5).

¹⁹ Throughout the rest of the paper, unless otherwise specified, we will use ‘freedom’ as shorthand for ‘transcendental freedom’. We understand this as a contra-causal power, and hence a libertarian conception of freedom.

²⁰ Spatial metaphors – such as *outside* – are somewhat cumbersome when it comes to discussing transcendental idealism. Nevertheless, Kant is clear that freedom is a property of things-in-themselves, and not a property of the world of appearances, and thus in this sense, freedom is *outside* of the empirical world.

²¹ See, for instance, A/B:557/585 and IV:457.16-9.

2) freedom occurs outside of time. After doing so, we will turn to possible tensions between these claims and Kant's commitment to transcendental idealism.

Before we begin, it is worth noting that in this paper, we want to remain relatively neutral on how to understand the distinction between the phenomenal and noumenal, as we think the tensions we are bringing out apply to any account of transcendental idealism. However, that being said, we will say something about both the *metaphysical* issues that moral education raises, as well as how our concerns also apply to a non-metaphysical *two-standpoint* reading of Kant.²²

No Experience of Freedom

In a footnote in the first *Critique*, after noting that we only have immediate cognition of the world of appearances, Kant claims that:

The real morality of actions (their merit and guilt), even that of our own conduct, therefore remains entirely hidden from us. Our imputations can be referred only to the empirical character. How much of it is to be ascribed to mere nature and innocent defects of temperament or to its happy constitution (*merita fortunae* [merit of fortune]) this no one can discover, and hence no one can judge it with complete justice. (A/B:551n/579n)

We see an echo of this in the start of *Groundwork* II, where Kant famously claims that:

In fact, it is absolutely impossible by means of experience to make out with complete certainty a single case in which the maxim of an action that otherwise conforms with duty did rest solely on moral grounds and on the representation of one's duty (IV:407.1-4)

It is important to differentiate between two claims here. The first is the plausible thought that we can never know with "complete certainty" or judge with "complete justice" what someone's maxim is. However, Kant seems to be committed to something more than this. On his set-up, it is not just that we cannot know *with complete certainty* whether someone acted for the sake of the moral law, but rather that we cannot know, with *any degree of confidence*, whether someone acted for the sake of the moral law. We can see this tension in the above two passages. Kant moves between the weaker claim that "it is absolutely impossible by means of experience to make out with *complete*

²² There is now an extensive scholarly and philosophical debate about how best to interpret transcendental idealism. See Allison (2004), Chiba (2012), and Allais (2015) as some recent monographies on the topic.

certainty” and the stronger claim that the “real morality of actions [...] remains *entirely hidden* from us”.

His conception of experience seems to commit him to the stronger claim. We can see this in the preceding claim: “imputations can be referred only to the empirical character” (A 551n/B 579n). When it comes to experience, all we encounter are empirical characters, which are determined by natural necessity.²³ Whether or not someone acted for the sake of the moral law is not going to show up *in experience*. This will turn out to be important for our discussion of the epistemic problem in the next section.

Before we continue, we should note that there is one special exception here, namely the fact of reason, where I feel the moral law pull against all of my sensible inclinations.²⁴ However, this is not technically *experience*, it is an exclusively first-personal awareness.²⁵ When it comes to third-person experience, we do not encounter morality resisting sensible inclinations – we only encounter the world of sense, determined by natural necessity.²⁶

Freedom outside of Time

There is an additional wrinkle here. Kant famously locates freedom outside of space and time. He views space and time as forms of intuition, and claims that *everything in time is determined*. In the second *Critique* for instance, he argues that everything in time is determined, and that freedom

²³See IV:407, V:100; VI:51; Stang (2013, 106n9). We should note that this applies to both others and ourselves. See Papish (2018, 159-64) for more extensive discussion of how self-cognition concerning maxims, motives and character can never be fully accomplished. Papish lists a number of psychological and epistemic reasons that underlie this opacity in Kant’s framework. She ultimately suggests that we can only make sense of the duty to cognize ourselves and of the epistemic limits of self-cognition, if we adopt a broader notion of self-cognition that she calls “self-interpretation” and that is not focused on motivation but on assessing oneself in the light of ethically rich concepts, such as modesty, cowardice, etc. (ibid.ch.6).

²⁴ See V:29.28 –30.30.

²⁵ See Grenberg (2013) for an account of the exclusively first-personal nature of the fact of reason.

²⁶ For an extended account of this epistemological problem, see Saunders (2016). See also Sticker (2016) for an account of the various ways in which Kant’s practical philosophy requires third-personal assessments.

therefore needs to be located outside of time (V:94.22-103.20).²⁷ He then refers back to what transcendental idealism accomplishes in the first *Critique*, and remarks:

Of such great importance is the separation of time (as well as space) from the existence of things in themselves that was accomplished in the *Critique* of pure speculative reason. (V:102.37- 103.2)

This has been quick, but hopefully two key points are clear. Transcendental idealism makes freedom possible for Kant, but also commits him to claiming that we cannot *experience* someone acting for the sake of the moral law, and that freedom is located *outside of time*.

III. Tensions

Having now laid out some of the very basics of Kant's account of moral education and two important elements of transcendental idealism, we want to now turn to some tensions between these commitments.

We think that Kant's locating freedom outside of time and experience causes a number of problems for his account of moral education.²⁸ These problems are specific cases of a much more general issue, namely, that it is difficult to reconcile transcendental idealism with anything that requires interaction between the empirical and the noumenal (cognition, synthesis, ethical decision making, etc.). Some of the problems we bring up might generalize and represent tensions between

²⁷ Here it is worth noting that the mental is in time (but not space). So conceived, deliberation, decisions, thought, experience and so on, are all subject to causal determination. This is important since Grenberg (2013) recently provided an extensive argument for the possibility of introspective moral experience of something that reveals a noumenal cause. We cannot discuss her proposal in detail here. We should keep in mind, though, that an introspective experience has to obey the same epistemic limitations as experiences of external objects. See Ware (2015) and Sticker (2016) for critical discussion of Grenberg.

²⁸ Another possible source of tension for Kant's conception of education is that moral education is something that is done to an agent by *external sources* such as parents, teachers or institutions. This raises the issue of how external sources can influence an agent in a way that does not undermine her autonomy. See for instance IX:453.28-37. Beck (1960, 235; 1979, 22), La Vaque-Manty (2006) and Formosa (2012, 163–77) provide discussions of this problem.

transcendental idealism and Kant's ethics more generally, but in this paper, we focus on tensions pertaining to moral education.

In what follows, we consider the metaphysical and epistemic problems that Kant's locating freedom outside of time and space causes.

The Metaphysical Problem

The metaphysical problem is relatively straightforward. Kant's idea of basic moral education suggests that things done in the empirical world can affect whether or not someone is transcendently free. This also suggests that we gain transcendental freedom *over time*, which would conflict with Kant's insistence that transcendental freedom is outside of time. The idea of motivational education in turn seems to imply that things done in the empirical world affect the noumenal subject, in particular, how effectively the moral law motivates an agent.

Here we can return to the basic issue we mentioned in section one, namely that transcendental idealism makes it mysterious how a child could transition to an agent. On this topic, Claudia Blöser claims that “[t]he transition from child to full person is beyond any empirical explanation” (Blöser 2015, 205). Of course though, we have *some* empirical explanations of how children develop into full persons or agents – we agree with Kant's account of basic moral education that it involves other agents (such as parents or educators), a certain amount of habituation and external pressure, socialization and learning abstract notions such as rules. The question is, however, how Kant could accommodate this, given that he locates transcendental freedom outside of experience, space and time.

There is both a general and a specific issue here. The *general* issue concerns the large and fundamental topic of interaction or affection between the empirical and the noumenal.

Kant believes to have addressed this issue sufficiently at least for moral *actions*. Morally worthy actions are events in the causally determined world, which are motivated by respect for the moral law. Respect for the moral law is a feeling and as a feeling part of the causal world, but it is a feeling that is self-wrought [“selbstgewirkt” (IV:401fn.)] and originates in reason or in the intelligible. In the Second Critiques' Incentive Chapter (V:71-89), Kant provides an extensive explanation for how reason can motivate agents to actions in the phenomenal world.

However, even if we accept Kant's conception of moral motivation, this does not yet solve the problem at hand. A child who is not yet a moral agent presumably lacks the feeling of respect (though it might, in some sense, already have a capacity to experience it). On Kant's framework the child is, so to speak, stuck entirely in the phenomenal world. The goal of basic moral education is to assist an agent in the crucial transition from an agent who can defer gratification of an immediate desire for

the sake of obtaining a greater good than satisfaction of this desire to someone who can abstract from *all* desires for the sake of duty alone. To understand moral education we would need a story about how the phenomenal can influence or affect the noumenal. And it is not clear how Kant could deliver such a story. In fact, it is telling that Kant dedicates almost 20 pages of his second Critique to a detailed discussion of how the noumenal or pure practical reason can become practically effective and considers this the most important question of this work, but does almost entirely neglect to explain how phenomenally imparted content can shape or influence the noumenal in any way. A satisfying account of moral education would require both: A story about how anything phenomenal can foster autonomous willing and how this willing can become effective in the world of experience.

At this point, there might be an easy reply to our criticism: Maybe educators or other agents are not needed at all for becoming a moral agent. After all, we thus far only spoke of possible ways education can *assist* the crucial transition to moral agency. Maybe this assistance is superfluous. Nothing we have said so far shows that it is impossible for children to become full moral agents and for fully developed rational agents to become morally better. Maybe in the case of basic education this is something that just happens (and we cannot understand how), and in the case of motivational education it happens as a result of an agent's work on her own character and without any external influence.²⁹

Whilst this would leave Kant's notion of agency intact, it would be a major problem for his conception of moral education, since it would leave Kant in a position where he cannot tell us why basic moral education is needed at all.³⁰ As the Sulzer footnote shows, Kant himself thinks that his

²⁹ As noted (fn.18), moral improvement can, according to Kant, occur simply as the result of a "single and unalterable decision" (VI:47.37-48.1), and could be inspired or fostered by other means than another agent or educator, for instance, through the experience of the feeling of the beautiful and sublime. These feelings, however, presumably already *presuppose* moral agency and can thus only serve for motivational education. See Guyer (2014, sec.3) for discussion of these aesthetic means of moral education. Guyer (idib.23) stresses that "the moral benefit of art cannot lie in teaching us moral ideas that we have not previously had, since the central ideas of morality are all *a priori* and available to us as soon as we reflect on a proposed action". Furthermore, Roth (2014, sec.2) suggests that moral education can also be a matter of "taking control of our conceptions of the world, others, and ourselves and our responses to them; and this we do when we comply with the principles of practical reason", i.e., it does not necessarily require someone who engages a pupil with the intention to morally educate her. Still, even if moral improvement were something not resulting from anyone doing something with the intention of educating, we can wonder how this is possible, according to a framework that distinguishes between a level of appearances, on which human beings can affect each other, and a level that is beyond the causal sphere. One conception that would avoid these problems is Papish's (2018, ch.8) proposal that the ethical community "addresses obstacles to moral progress" (ibid.205). In particular, the ethical community can contribute to the cognitive dimensions of moral improvement and "mitigate a problem regarding discordant moral judgments" (ibid.). This aid for moral improvement, however, only applies to what Papish identifies as the second, cognitive stage of moral improvement and leaves open whether and how other agents, educators or the community can influence an agent to commit himself to morality in the first place.

³⁰ It would be difficult to see then in what sense education could be the source of all good in the world (IX:448.12-17) and why human beings have to be "educated to the good" (VII:325.5-11).

account of education is to be preferred over other proposals. He should be able to explain how education works such that his account can do what competing accounts fail to achieve, namely, help to morally improve agents. The notion that important elements of his account are inscrutable seems inadequate given Kant's stress on the methodological role of education for his debate with his academic colleagues. We will say more about this shortly.

The more *specific* issue is that typically children become agents *over time*, and improve *gradually* as they develop into adults.³¹ Education presupposes the possibility of *gradual* improvement, which means that it occurs *over time*, whereas Kant locates (transcendental) freedom outside of time.³² Consider the following example: Immanuel Kind was not a moral agent in 2009 – he could not be brought to judge that he should forego any of his inclinations at all. However, things change. Now, in 2019, Immanuel Kind is a moral agent – he judges that he ought not to lie (when presented with scenarios in which it would be an option for him to further his own interest by means of a lie or false promise) and is convinced that it is at least possible for him to obey this prohibition. Immanuel Kind has changed over time. It seems that in 2009, Immanuel Kind was not under the moral law and now he is. This process and change in Immanuel Kind is not puzzling on an empirical level, since the empirical world constantly changes and we have observed the transition described in many individuals. On a noumenal level though, what happened to Immanuel Kind is very puzzling.

From the perspective of transcendental idealism, it is at least in one sense awkward to say that a child is not yet (transcendentally) free at some point in time and free at another, later, point in time. As far as the noumenal is concerned, there are no points in time. A noumenal self is always free, or rather: temporal terms such as 'always' are inapplicable to it. If we take this seriously, we might get a very different conception of the transition from a supposedly non-free child to a free adult, namely, one in which the patient of education is in one sense always fully free and only her phenomenal self needs to be moulded such that this freedom can be exercised in the phenomenal world.³³

This view however would have very revisionary consequences for how we would have to think about the dead and the yet to be born. If time does not matter for freedom and freedom is what matters for morality, we would have the same duties to respect past and future generations as we have to

³¹ Elsewhere, Saunders (2018) argues that transcendental idealism makes it hard for Kant to provide a plausible account of degrees of responsibility; see Frierson (forthcoming) for a response to this.

³² For a broader account of some of the problems caused by Kant's claim that transcendental freedom occurs outside of time, see Freyenhagen (2008).

³³ Guyer (2014, 21) at least implicitly endorses such a conception, when he claims that "the point of moral education is to convince the child of his own freedom". Freedom is here already presupposed.

respect our contemporaries. Furthermore, this view would also make it difficult to understand why we think that children below a certain age cannot be held morally or criminally responsible for their actions, and why, at a very young age, their behavior does not even constitute actions. The same is of course true for the elderly suffering from dementia, etc. It seems that in our everyday practices membership in the noumenal comes with a time index and that it starts at some (possibly extended) point in time and might end even before we die. There can also be interruptions, such as periods of severe mental illness.

A similar specific problem might apply to motivational education, depending on how we understand the idea of fostering motivation. Does this mean that the noumenal is monolithic and that only the sensuous agent is shaped by education? Even then the result is that the noumenal meets less resistance, can work more effectively on the agent's sensuous side. Is that a change that is in any sense a change of the noumenal, or affecting the noumenal? One could respond that the metaphoric manner of speaking about the noumenal we are employing here makes Kant's conception look more problematic than it is – for instance, we should certainly not think of the noumenal as *doing* anything at all (working on an agent or on the phenomena). However, if we want to understand how moral education exactly works and can be effective on Kant's framework we have to talk about these matters and transcendental idealism seems to preclude Kant from doing so.

At this point, Kant might respond that these metaphysical issues are beyond our grasp. The way we have discussed these problems so far violates the epistemic limits transcendental idealism imposes on us. Kant, as we saw, emphasizes that the ground of freedom is “inscrutable” (VI:138.19). According to Kant, we can cognize that we are free, since freedom follows from or can be deduced from morality. Freedom is hence cognizable (VI:138.16-9). The metaphysical ground of freedom, which could settle the question of how we come to be free, however, is “*not given to us for cognition*” (VI:138.20) and thus inscrutable.³⁴ The result of this for thinking about education is that “it is impossible to form a concept of the production of a being endowed with freedom through a physical operation” (VI:280.23-5) or, we should add, through *any* operation. Once we are under the moral law, we can cognize our freedom, but how we get to be under the moral law in the first place is inscrutable. Kant cannot tell a comprehensible empirical (or other) story about it.

Appealing to the inscrutability of freedom allows Kant to side-step the metaphysical tensions between transcendental idealism and his account of moral education. However, it also comes with a

³⁴ As an example of this inscrutability Kant suggests that we cannot know what, if any, role God plays for our moral agency (VI:139.1-12).

cost. For one, if we accept that these metaphysical issues are beyond our grasp, then we seem to renounce a claim to provide an explanation of how moral education works. When it comes to questions of how freedom develops over time or how the empirical world affects the noumenal in education, Kant cannot provide answers. This also introduces an epistemic tension between transcendental idealism and moral education, to which we will shortly turn.

Before we address this epistemic tension however, we should consider a further possible reply to our metaphysical objection. According to a certain reading of transcendental idealism, there is no metaphysical problem at all for Kant's account of moral education, since transcendental idealism is not a metaphysical position. The distinction between the phenomenal and the noumenal is not an ontological distinction between different worlds, but reflective of different *standpoints* on the world and self that rational agents can take up. On this picture, Kant cannot admit any *experience* of freedom, but that is fine. Experience concerns the *theoretical* standpoint, where when it comes to morality, we occupy the *practical* standpoint. And this allows us to overcome the metaphysical problems at hand, or to avoid them altogether.³⁵ The basic thought is that, from the *theoretical* standpoint, everything appears determined by inclinations, but from the *practical* standpoint, we can observe freedom.

However, we think that it is mistaken to assume that a two-standpoint reading is an easy fix for the metaphysical problem, even though this position presents a non-metaphysical reading of transcendental idealism. After all, what we need is a plausible story that links experience and freedom and there is nothing in the idea of two different standpoints (as opposed to different worlds) that would already on its own deliver such a story. Consider, for example, a simple case: I treat Immanuel Kind at age 25 differently than I do at age 4. At one of these times, Immanuel Kant is free, and at the other time, he is not. Now, Kant does not allow us to determine this through experience – in experience, we just encounter two beings determined by their inclinations. And it is not clear how the practical standpoint could help here. From the practical standpoint, how can we distinguish between Immanuel Kind at 4 and 25? It is not enough to say that we just do. This sidesteps the important issue of how Kant can vindicate these practices, given his conception of experience.³⁶ The appeal to the practical standpoint instead merely shifts the question to: How do we know which beings can take up, or have taken up, the practical standpoint? Kant's conception of experience precludes an easy answer to this question. The practical standpoint allows us, from the first-person, to conceive of ourselves as independent from the world of experience, but it does not give us a new third-personal conception of

³⁵ See Korsgaard (1996, x-xi) for a brief general account of this strategy.

³⁶ For another important objection to the 'two-standpoint' reading of Kant's theory of freedom, see Nelkin (2000).

experience or a new form of *intuition*. That would move us far beyond Kant and transcendental idealism.

Moreover, the introduction of two standpoints does not help us to understand the *interaction* between the practical and theoretical, and one of the important metaphysical problems that education poses is the issue of interaction. At this point, the standpoint theorist could agree, but insist we merely have two standpoints, and are unable to get beyond them to see how freedom might influence – or show up in – experience.³⁷ This takes us back to Kant’s claim that issues of interaction or affection are inscrutable. Again, this provides a way to side-step the metaphysical problems we are raising, but it comes at a cost. The more we insist upon the limits of our knowledge when it comes to freedom, the more we create epistemic problems for Kant’s position. On this note, we turn to the epistemic tensions, which affect both metaphysical and non-metaphysical readings of transcendental idealism.

The epistemic problem

Given Kant’s conception of experience, we cannot know whether other agents (or even we ourselves) are ever acting from duty. Since it is the point of moral education to make agents able to and better at acting from duty, it is difficult to determine whether any program of moral education is particularly successful or even adequate to the task. This problem also comes in a number of shapes:

1) We do not know *whom* to educate. Saunders (2016) has recently argued that transcendental idealism precludes knowledge of free agents *as* free agents, since this would require *experience* of (transcendental) freedom in experience.³⁸ If this is correct, the difficulty is even more pronounced for basic education, since here we must recognize those that *can* be free (but currently are not). If we cannot cognize noumenal properties, then how could we ever cognize *potential* noumenal properties? One thought might be that we just try to educate all sorts of entities and see if any are responsive. But this requires us to be able to tell from the phenomenal world who is responsive to moral education

³⁷A similar problem shows up in Allison’s (1996b, 127) criticism of Hudson (1994). Allison objects that, in offering compatibilist/anomalous monist readings of Kant’s theory of freedom, Hudson cannot provide a satisfying account of Kant’s talk of the “causality of reason”, or Kant’s worry that reason might not have causality. Allison is right that this counts against Hudson’s account, but it also counts against his own non-metaphysical reading of Kant. As he himself acknowledges elsewhere: “Although my interpretation denies that the claim that reason has a causal power can be taken literally without leading to absurdity ... it also leaves room for the notion that such a power can be attributed to the will (*Willkür*) of a rational agent, with the latter being itself noncausally determined (governed) by reason“ (Allison 1990, 80). But this only shows that it is not (conceptually or otherwise) impossible to speak of a causality of reason. What is still missing is a plausible account of how this is possible. For further discussion of this problem, see Ameriks (2003, 215-6).

³⁸Again, we should note that throughout this paper, unless otherwise specified, we use ‘freedom’ as a shorthand for ‘transcendental freedom’.

and who is not. Transcendental idealism precludes this knowledge. We could never know that an entity is able to or about to develop freedom, since, phenomenally, such an entity would not be any different from an entity that shares all phenomenal properties but not the capacity for freedom.

2) We do not know what kind of education *works*. Kant sets out an educational program, but on his framework it is difficult, maybe impossible, to tell whether his program is better than other ways of educating. Whether Kant's conception of education is better than competitors, or any good at all is settled by things which we cannot know. Furthermore, it weakens the claims he wants to make about which styles of education lead to moral agency if he cannot judge educational success. This ties into our third point.

3) Kant cannot even be sure that education is of *any use* at all. If we cannot verify the success of education, it could be entirely redundant. This is perhaps less of a problem for basic moral education than it is for motivational education. After all, Kant, as we pointed out, says little about basic education and he seems to be confident that often or almost always this kind of education produces the right result: A rational agent under the moral law. Motivational education is, however, a matter of degree. It can be done better or worse (even so badly that its effect is counterproductive – as Kant suspects is the case with the proposals and practices of most of his academic colleagues and public educators) and Kant greatly cares about how well it is done. In fact, he thinks that a metaphysics of morals that presents the moral law in its full purity “is a desideratum of the highest importance” to encourage the “actual execution” of the prescriptions of duty (IV:410.19-25). The application of his theory to improve educational practices is one of the reasons why his critical practical philosophy is a desideratum.

In summary, the metaphysical problem poses an obstacle to basic moral and motivational education, since: 1) it is not clear how the world of sense and transcendental freedom *interact*, and 2) transcendental freedom occurs outside of time. We should re-iterate that the first part of this problem applies to interpretations of transcendental idealism that pose different worlds or realms of existence, but also to more metaphysically parsimonious readings, since they still incur the burden to explain how different standpoints *interact*.

The epistemic problem emerges from the epistemic limits that transcendental idealism imposes on finite agents like us, and this poses challenges independently of how metaphysically loaded one interprets transcendental idealism. For Kant, we can never know whether anyone acted

from duty, and as such, it is unclear that he can vindicate our knowledge of: 1) whom we educate; 2) What kind of moral education works; and 3) whether moral education is of any use at all.

IV. Resolving the tension

We will now discuss how Kant should react to the problems set out in the previous section. We will first discuss one way for Kant to respond that might seem promising but fails, and then present our own solution.

Here is one possible way to save Kant: Perhaps moral education makes agents better at *outwardly conforming* to duty. This would not be a meaningless achievement. If everyone always outwardly conformed to duty, no one's rights would be violated. Importantly, whether an agent outwardly conformed to duty can be cognized from observation of behaviour and hence is epistemically transparent to oneself and others.

However, it is a crucial feature of Kant's ethics that he is remarkably uninterested in mere conformity with the moral law.³⁹ When Kant speaks of morally improving agents, he clearly has in mind that educators foster agents' commitment to duty for the sake of duty and not for ulterior motives. This should encourage agents to act from the moral law and help them develop their freedom in the fullest sense. In addition to how Kant himself stresses that the point of moral education is not mere external conformity, we should bear in mind that many of the means that seem apt to further mere outward conformity to duty, such as the threat of punishment, differ considerably from the presentation of the moral law in all its purity, that Kant envisages in his conception of motivational education. An increase in outward conformity could therefore be interpreted as both a step in the right moral direction as well as that some pathological means to incentivize mere external conformity are very effective, and we cannot know with certainty which of these two interpretations is the correct one. Thus, education that effects outward conformity will not be enough for Kant (and might even be harmful if it is mere conformity that is incentivized).

A more promising way out for Kant becomes apparent when we revisit the Sulzer footnote. As we saw, in this footnote Kant seems to make a prediction that, if true, confirms his assumption that

³⁹ See IV:397.11-2 where Kant *en passant* grants that most actions are in conformity with duty, but does not seem to believe that this shows anything very interesting about the morality of actions. In what follows and for the rest of the *Groundwork* and *Second Critique*, he only focuses on what makes agents morally worthy, namely, duty as an action's incentive, not on how to achieve mere conformity. Matters are, however, different for Kant's *Doctrine of Right*. Here Kant is explicitly concerned with outward conformity.

an educational approach informed by a philosophy based on a pure principle cannot fail to morally improve agents. Presenting an example of action for the sake of duty alone, Kant predicts, will affect rational agents, and even children. Kant believes that the way a child expresses or describes his feelings, when confronted with the moral law in all its purity, will support his ethics over popular contenders, since it shows that the supreme principle of Kant's theory can, contra Sulzer, serve for moral improvement.

This claim is by no means a one-off. A similar thought occurs in the Second Critique's *Doctrine of Method*, where Kant discusses moral education and provides some practical indications of how an educator should go about helping pupils develop their disposition:

But if one asks: What, then, really is *pure* morality, by which as a touchstone one must test the moral content of every action? **I must admit that only philosophers can make the decision of this question doubtful [1]**, for it is long since decided in common human reason, not indeed by abstract general formulae but **by habitual use [2]**, like the difference between the right and the left hand. We will, accordingly, first show in an example the mark by which pure virtue is tested and, representing it as set-before, say, a ten-year-old boy for his appraisal, see whether he must necessarily **judge [3]** so of himself, without being directed to it by a teacher. One tells him the story of an honest man whom someone wants to induce to join the calumniators of an innocent but otherwise powerless person (say, Anne Boleyn, accused by Henry VIII of England). He is offered gain, that is, great gifts or high rank; he rejects them. This will produce mere **approval and applause [4]** in the listener's soul, because it is gain. Now threats of loss begin. [...] But, so that the measure of suffering may be full and he may also feel the pain that only a morally good heart can feel very deeply, represent his family, threatened with extreme distress and poverty, as *imploping him to yield* and himself, though upright, yet with a heart not hard or insensible either to compassion or to his own distress; represent him at a moment when he wishes that he had never lived to see the day that exposed him to such unutterable pain and yet remains firm in his resolution, without wavering or even doubting; then will my youthful hearer be raised gradually from mere approval to **admiration [5]**, from that to amazement, and finally to the greatest veneration, and a lively wish that he himself could be such a man (though certainly not in such circumstances). (V:155.12-156.27)

Once more Kant here maintains that his ethics can be defended against doubt and criticism by showing that presenting a pure incentive elicits a reaction from agents that nothing else can.⁴⁰ The way Kant introduces the example as a response to a possible challenge by an academic colleague ("But if one asks") suggests that Kant also hopes to gain confirmation for his *philosophy*.⁴¹ He

⁴⁰ Guyer (2014, 5) sees this passage as chief evidence that "examples of the possibility of actually living a moral life in the face of the inexorable imperfection of the human condition play a central role in moral education" for Kant. Furthermore, he stresses (ibid.18) that "on Kant's account, what both children and adults must learn is not the principle or content of morality—that, in Kant's view, everyone, even a child, already knows—but the possibility of their own freedom, in particular, their freedom to be moral". It is indeed striking that Kant's examples of moral education are hardly ever examples of teaching others about concrete duties they have, but rather appeals to their better nature or to the fact that it is always possible to do the right thing. See Sticker (2017) for critical discussion of the underlying methodological assumptions Kant makes here.

⁴¹ See Sticker (forthcoming) for an argument that Kant here seeks to support his own theory against other philosophers, particularly against the popular philosophers of his time. See also the *Common Saying*, where Kant explicitly suggests

recommends that when faced with a colleague who advocates a competing conception of morality, the critical philosopher involves a boy who has already gained use of his pure practical reason, but who is too young to have been subject to much corrupting external influence. Kant maintains that a criterion for morality (a “touchstone”) is already present and deeply rooted in how we ordinarily reason about morality [2]. Academic philosophers can overlook this and construct a system that leads away from this criterion ([1] see also V:35.16-17). The criterion is, however, exhibited in the boy’s *judgements* [3], as well as *reactions or attitudes* such as approval and applause [4] and *admiration* [5]. Common agents’ “habitual use” [2] of reason or their judgements [3] can inform the philosopher.

Shortly before Kant presents this example of moral education, he repeats the main idea of the Sulzer footnote:

We will prove, then, by *observations anyone can make*, that this property of *our mind*, this receptivity to a pure moral interest and hence the moving force of the pure representation of virtue, *when it is duly brought to bear on the human heart*, is the most powerful incentive to the good (V:152.32-153.1, our emphasis).

Kant believes that he can vindicate his claims concerning the affective force of duty from observation of “our mind” and “the human heart”. This observation is supposedly one “anyone can make” who is in doubt about the *de facto* effect the moral law has on rational agents.⁴² A philosopher, however, might be the only one who has the theoretical understanding and sophistication to know how to “duly” bring the moral law to “bear on the human heart” of a corrupted agent.

Kant believes that agents will feel admiration when they acknowledge (or think it likely) that an action is done from respect for the moral law and in the face of adversities. He thinks that it establishes the superiority of his ethics over all popular contenders that he can show that the supreme principle of his theory *de facto* affects and moves rational agents.

It is crucial to disambiguate two versions of this claim. The first is a *strong* reading of the above passages, where Kant claims that experience can confirm that his account of moral education works

that a scenario be presented to a boy of eight or nine years (VIII:286) in which someone is entrusted with a deposit and could embezzle it without negative consequences to anyone and with very positive consequences for himself and his family (ibid.17-29). The boy’s response, that this would be wrong independently of consequences, is supposed to answer a criticism by the popular philosopher and eudaemonist Christian Garve that he cannot find in his “heart” a notion of duty independent of happiness (ibid.284.9-285.22). If the critical philosopher can show that the boy or a (relatively) uncorrupted rational agent can be brought to acknowledge the value of pure morality, then this can function as a warrant for his theory.

⁴² See also Kant’s discussion of the humble common man (V:76.36-77.9). Kant here describes my reaction (“*my spirit bows*”) if I have reason to suspect that the humble common man is an example of good moral conduct. Kant even says that in this case “I see observance of that law and hence its practicability proved before me in fact”.

and is the only one that works. As we have seen though, transcendental idealism precludes this; transcendental freedom, and acting for duty's sake do not show up in experience. Thus, if Kant requires – or desires – that *experience* confirms that his account of moral education works, he would face a serious tension between his views on moral education and his commitment to transcendental idealism.

However, there is also a *weak* reading of these passages, where Kant instead claims that his account of moral education is the only account that *could* aid the development of a genuine moral disposition. This makes sense given what we know about Kant's theory of morality, namely, that we should act for duty's sake. As we saw in section one, Kant worries that other theories of moral education – or even his own theory of basic education – might only produce a “*manner of the good*” (A/B:748/776). Again, Kant sees it as one of the central problems of his time that educators treat rational moral agents like children who need to be bribed and scared into obedience to morality. However, if we are interested in a moral education that can cultivate the ability to act *for the sake of duty*, Kant can rule out any such accounts from the start, not because he can empirically confirm that they fail (or that his account succeeds).

Where does this leave Kant's account of moral education? We think that, given his account of moral motivation, he can claim that his is the only account of moral education that could work. However, given the epistemic limits of transcendental idealism, he can neither explain how it works, nor confirm *in experience* whether in fact his account does work. He thus lacks theoretical grounds for his account of moral education. As we saw in the first section of this paper though, Kant thinks moral education should not be regarded as separated from the more abstract task of philosophical enquiry into the supreme principle of morality. Advising on how to educate to moral agency and to the good is one of the most important functions a critical ethicist can fulfil. As such, Kant could maintain that while we lack the ability to explain how moral education works or to confirm that his account works in experience, we have important practical grounds for accepting it.

We suspect that this might be a one of those rare cases where a result pleases both (some) Kantians and (some of) his critics. Kant's critics can maintain that transcendental idealism makes it hard for Kant to account for the metaphysical and epistemic issues that moral education raises, and moreover, that he is unable to confirm in experience whether or not his theory works. Kantians can accept this, but claim that they cannot – and do not need to – address these issues. They can also claim

that the working of Kant's account of moral education ends up similar to a postulate, where we lack theoretical grounds for a proposition, but can rationally adopt it on practical grounds.

This raises an interesting question, namely, what the *appropriate attitude* is that Kant and Kantians are to adopt towards Kant's account of moral education. Before we end this paper, we want to briefly canvas some of the available options.

A postulate is a theoretical proposition, for which we lack theoretical grounds, but can adopt on practical grounds instead (V:122.23-5). The appropriate or warranted attitude for moral agents towards a postulate is "pure rational faith" [reiner Vernunftglaube] (V:126.10-1). However, Kant requires that for finite agents such propositions be *inseparably connected* (V:122.24-5) to the moral law and the highest good; and while education is important for Kant's practical philosophy, it is not clear that it stands in a *necessary or inseparable* connection to the moral law (or the highest good). After all, it seems that agents, at least once they have acquired autonomy, could in principle develop their character without external help, since this "only" requires unconditional commitment to duty. Agents can, without encouragement by an educator, make this commitment (V:36.28-37.5).

An attitude that is in some sense similar to this rational faith is *hope*. Is it appropriate to *hope* for the success of Kant's account of moral education? This question is complicated by the fact that Kant sometimes deploys 'hope' in a specific and narrow sense. In the *Critique of Pure Reason*, he claims that, "all hope concerns happiness" (A/B:805/833), and in this strict sense, hope is not the appropriate attitude to take towards moral education (which does not concern happiness but morality and worthiness to be happy). However, Kant also uses the term 'hope' in a broader sense,⁴³ closer to our everyday usage, namely an attitude of wishing or desiring that something will come about (where this is not impossible).⁴⁴ And this does seem like an appropriate attitude to take towards Kant's account of moral education. So conceived, we cannot *know* how his account of moral education works, we also cannot *know* whether or not it actually does work, but we can *hope* that it will. Indeed, this

⁴³ In the *Religion*, for instance, Kant discusses how we hope to make moral progress (VI:48.3-8, VI:51.16-9), and hope to be pleasing to God (VI:62.1-3). He also discusses hope at length in *Toward Perpetual Peace* (see for instance VIII:370-1), and none of these discussions focus solely on happiness.

⁴⁴ The Oxford Dictionary defines 'hope' as: A feeling of expectation and desire for a particular thing to happen; Grounds for believing that something good may happen. A more technical and classical philosophical definition is that "'A hopes that *p*' is true iff 'A wishes that *p*, and A thinks that *p* has some degree of probability, however small' is true" (Day 1969, 89).

hope can also be *rational*, in that we have practical grounds that count in favour of it (and it is not impossible⁴⁵).

This discussion of practical grounds that count in favour of something also connects to Kant's conception of faith/belief.⁴⁶ In a now famous passage in the first Critique's Doctrine of Method, Kant distinguishes knowing and faith as follows:⁴⁷

If taking something to be true is only subjectively sufficient and is at the same time held to be objectively insufficient, then it is called faith/belief [*Glauben*⁴⁸] [...] when taking something to be true is both subjectively and objectively sufficient it is called knowing [*Wissen*]. (A/B:822/850)

If we think that objective sufficiency requires an *object* or objective cognition, as Kant suggests it does (A 821/ B 849), and transcendental idealism precludes that we can have objective cognitions that would prove the working of Kant's account of moral education, then Kant cannot *know* whether his account works. Faith/Belief [*Glauben*] then, at least as Kant defines it here,⁴⁹ also seems like an appropriate attitude to take towards his account of moral education.⁵⁰

This account of faith/belief connects back to Kant's account of hope. Chignell draws this out well in looking at Kant's claim that we ought to will an ethical community in the *Religion*. Chignell writes that:

We can't know that an ethical community is really possible, but once we see that we ought to will it, we have to Believe on practical grounds that it is indeed possible. This in turn licenses Belief in

⁴⁵ Cf. Chignell (2013, 206), who claims that hope is rational only if one is in a position to know that what one hopes for is not really impossible.

⁴⁶ For a fuller account of practical grounds for belief, see Williams and Saunders (2018).

⁴⁷ For the classic treatment of this, see Chignell (2007).

⁴⁸ Here, we translate '*Glauben*' as 'faith/belief'. See Höwing (2016, 201n3) for a brief account of the complications of translating '*Glauben*'.

⁴⁹ It is worth noting that this definition of faith/belief is broader than Kant's definition of pure rational faith in the Second Critique.

⁵⁰ One complication concerns whether Kant's account of moral education is "only subjectively sufficient". For a full account of this clause, see Höwing (2016, 201-22). Höwing (2016, 204-5) acknowledges that Kant appears to be ambiguous in his use of this clause, but finds this unsatisfactory. He offers a reading of the clause, which looks to remove the seeming ambiguity. On this reading, "only subjectively sufficient" picks out cases where a particular agent is required to assent to a proposition, but *not every* rational agent (for whom it is available) is required to assent to that proposition. (Höwing 2016, 218). This is an interesting suggestion, and it certainly makes sense of Kant's claim in the *Jäsche Logik* (IX:66). However, it strikes us as? slightly ill-fitting for the case at hand. For one, it is not clear that anyone is *required* to assent to Kant's account of moral education. And secondly, it might make the attitude a little too personal, but there are no particular grounds that compel *me* to assent to Kant's account of moral education.

supersensible assistance, and also underwrites rational hope for a this-worldly but still inconceivable goal. (Chignell 2014, 116)

Kant could say something similar for moral education: We can't know that moral education is really possible, but once we see that we ought to will it, we have to have faith/believe on practical grounds that it is indeed possible.⁵¹ This in turn licenses faith/belief in the possibility of moral education being successful, and also underwrites rational hope for this this-worldly goal.

Of course, there is much more to say here. Kant's attitudes of hope and faith/belief are complicated in ways that we have not addressed here. Here, we take ourselves to have briefly canvassed the available options and to have suggested that both hope and faith/belief might be appropriate in this case.

Conclusion

In conclusion, Kant has a multi-faceted account of moral education. However, in locating freedom outside of time and experience, transcendental idealism creates problems for this. We have considered what Kant can – and cannot – say about moral education, given his commitment to transcendental idealism. We have argued that he can neither explain how, nor confirm whether his account of moral

⁵¹ It appears that there is a link here between the limitations of our knowledge (that make faith possible), and Guyer's (2014, 5) earlier claim that education has to be open-ended and never-ending. After all, the impossibility of noumenal knowledge entails that there is no point in an agent's moral development at which she is warranted to judge that she is now good enough and does not need any further improvement. Even if an agent actually did not need to improve further (and Kant assumes that there is always room for improvement), she could not know this. We are grateful to an anonymous referee for noting this connection.

education works, but that he can still maintain that it is the only account that could work, and have hope and faith/belief that it does.

Literature

Kant's Works (with the exceptions of the First Critique) are quoted volume:page.line of the *Academy Edition*. Translations follow the *Cambridge Edition of the Works of Immanuel Kant*, edited by P. Guyer and A. W. Wood. The Kaehler lecture notes are quoted from Stark (2000).

Allais, Lucy (2015): *Manifest Reality: Kant's Idealism and his Realism*, OUP: Oxford.

Allison, Henry (1990): *Kant's Theory of Freedom*, CUP: Cambridge

(1996): *Idealism and Freedom*, CUP: Cambridge

(2004): *Kant's Transcendental Idealism: An Interpretation and Defense*, Yale University Press: New Haven, London.

Ameriks, Karl (2003): *Interpreting Kant's Critiques*, OUP: Oxford.

Baron, Marcia (2009): "Kantian Moral Maturity and the Cultivation of Character", in: Siegel (ed.): *The Oxford Handbook of Philosophy of Education*, OUP: Oxford, pp.227-244.

Beck, Lewis White (1960) *A Commentary on Kant's Critique of Practical Reason*, The University of Chicago Press: Chicago.

– (1979): "Kant on Education", in: Browning (ed.): *Education in the 18th Century* garland: New York, pp.10-24.

Blöser, Claudia (2015): "Degrees of Responsibility in Kant's Practical Philosophy", in: *Kantian Review*.

Chiba, Kiyoshi (2012): *Kants Ontologie Der Raumzeitlichen Wirklichkeit: Versuch Einer Anti-Realistischen Interpretation Der Kritik Der Reinen Vernunft*, de Gruyter: Berlin.

Chignell, A. (2007): "Belief in Kant", in: *Philosophical Review* 116: 323-360.

(2013). Rational Hope, Moral Order, and the Revolution of the Will. In Eric Watkins (ed.), *Divine Order, Human Order, and the Order of Nature*.

(2014). Rational hope, possibility, and divine action. In Gordon E. Michalson (ed.), *Religion within the Bounds of Mere Reason: A Critical Guide*. Cambridge University Press. pp. 98-117.

Cohen, Alix (2015): "The Role of Feelings in Kant's Account of Moral Education", in: *Journal of Philosophy of Education*, 50(4): 511-23.

Day, J.P. (1969): "Hope", in: *American Philosophical Quarterly*, 6(2): 89–102.

- Formosa, Paul (2012): “From Discipline to Autonomy: Kant’s Theory of Moral Development”, in: Roth, Surprenant (eds.): *Kant and Education. Interpretations and Commentary*, Routledge: New York, pp.163-177.
- Freyenhagen, Fabien (2008): ‘Reasoning Takes Time: On Allison and the Timelessness of the Intelligible Self’ in *Kantian Review*, 13.2, pp. 67-84.
- Frierson, Patrick (forthcoming): “Another Look at Kant and Degrees of Responsibility” in *Con-Texts Kantianos*.
- Giesinger, Johannes (2012): “Kant’s Account of Moral Education”, in: *Educational Philosophy and Theory*, vol.44, pp.775–86.
- Grenberg, Jeanine (2013): *Kant's Defense of Common Moral Experience. A Phenomenological Account*, CUP: Cambridge.
- Guyer, Paul (2014): “afterExamples of Perfectionism“ in: *The Journal of Aesthetic Education*, 48.3, pp. pp. 5-27
- Herman, Barbara (2007): *Moral Literacy*, Harvard University Press: Cambridge, Mass.
- Höwing, Thomas (2016): Kant on Opinion, Belief and Knowledge, in: *The Highest Good in Kant's Philosophy*, ed. by Thomas Höwing, Berlin: de Gruyter 2016, pp. 201-222).
- Hudson, Hud (1994): *Kant’s Compatibilism*, Cornell University Press
- Koch, Lutz (2003): *Kants ethische Didaktik*, Ergon Verlag: Würzburg.
- Korsgaard, Christine (1996): *Creating the Kingdom of Ends* (Cambridge University Press).
- La Vaque-Manty, Mika (2006): “Kant’s Children”, in: *Social Theory and Practice*, vol.32(3), pp.365-88.
- Louden, Robert (2011): “Becoming Human: Kant and the Philosophy of Education”, in: Loudon (ed.): *Kant’s Human Being. Essays on his Theory of Human Nature*, OUP: Oxford, pp.136-50.
- Moran, Kate (2009): “Can Kant Have an Account of Moral Education?”, in: *Journal of Philosophy of Education*, vol.43(4), pp.471-484.
- (2012): *Community and Progress in Kant's Moral Philosophy*, Catholic University of America Press: Washington DC.
- Munzel, Felicitas (1999): *Kant’s Conception of Moral Character. The “Critical” Link of Morality, Anthropology, and Reflective Judgement*, The University of Chicago Press: Chicago.
- (2002): “‘Doctrine of Method’ and ‘Closing’ (151-163)“, in: Höffe (ed.): *Immanuel Kant. Kritik der praktischen Vernunft*, Akademie Verlag: Berlin, pp.203-219.
- (2012): *Kant’s conception of pedagogy. Toward education for freedom*, Northwestern University Press: Evanston.

Nelkin, Dana (2000): 'Two Standpoint and the Belief in Freedom' in *The Journal of Philosophy*, 97, pp. 564-576.

Roth, Klas, Surprenant, Chris (eds.) (2012): *Kant and Education. Interpretations and Commentary*, Routledge: New York.

Roth, Klas, Gustafsson Martin, Johansson Viktor (eds.) (2104): „Introduction: Perfectionism and Education—Kant and Cavell on Ethics and Aesthetics in Society” in *The Journal of Aesthetic Education*, 48.3, pp.1-4.

Roth, Klas (2014): „Making Ourselves Intelligible—Rendering Ourselves Efficacious and Autonomous, without Fixed Ends“, in: *The Journal of Aesthetic Education*, 48.3, pp. 28-40.

Saunders, Joe (2016): 'Kant and the Problem of Recognition: Freedom, Transcendental Idealism, and the Third-Person' in: *International Journal of Philosophical Studies*, 24(2), 164-182.

– (2018): “Kant and Degrees of Responsibility” in: *Journal of Applied Philosophy*

Stang, Nicolas (2013): 'Freedom, Knowledge and Affection: Reply to Hogan' in *Kantian Review*, 18.1, pp. 99-106.

Stark, Werner (2000): “Vorlesung – Nachlass – Druckschrift? Bemerkungen zu Kant über Pädagogik”, in: *Kant-Studien*, vol.91, pp.94-105.

Sticker, Martin (2015): “Educating the Common Agent. Kant on the Varieties of Moral Education”, in: *Archiv für Geschichte der Philosophie*, vol.97(3), pp.358–87.

– (2016): “Kant on Engaging Other Agents and Observing Reason at Work”, in: *History of Philosophy Quarterly*, vol.33(4), pp.347-73.

– (2017) “Kant’s Criticism of Common Moral Rational Cognition”, in: *European Journal of Philosophy*, vol.25(1), pp.85-109.

– (forthcoming): “Kant, Eudaimonism, Act-Consequentialism and the Fact of Reason”, in: *Archiv für Geschichte der Philosophie*.

Surprenant, Chris (2010): “Kant’s Contribution to Moral Education: The Relevance of Catechistics”, in: *Journal of Moral Education*, vol.39, pp.165–74.

Sweet, Kristi (2011): “Philosophy and the Public Sphere: Kant on Moral Education and Political Critique”, in: *Idealistic Studies*, vol.41, pp.83-94.

Vanden Auweele, Dennis (2015): “Kant on Religious Moral Education”, in: *Kantian Review*, vol.20(3), pp.373-94.

Ware, Owen (2014): “Accessing the Moral Law through Feeling”, in: *Kantian Review*, vol.20(2), pp.301-11.

Weisskopf, Traugott (1970): *Immanuel Kant und die Pädagogik: Beiträge zu einer Monographie*, Editio Academic: Zürich.

Williams, Neil. William., & Saunders, Joe (2018): “Practical Grounds for Belief: Kant and James on Religion” in: *European Journal of Philosophy*.